

STRUCTURE-ACTIVITY RELATIONSHIPS OF MEMBRANE PROTEINS:
THE NCS1 FAMILY OF TRANSPORTERS AND
SENSOR KINASES OF TWO-COMPONENT SYSTEMS

Pikyee Ma

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy



The University of Leeds
Institute of Membrane and Systems Biology
Faculty of Biological Sciences

JUNE 2010

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 4 – Section 4.4

The work in Section 4.4 of this thesis is solely the candidate's own work and part of it (Figure 4.5) has been published in the following publication:

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., **Ma, P.**, Saidijam, M., Patching, S.G., Hope, R.J., Norbertczak, H.T., Roach, H.T., Iwata, S., Henderson, P.J.F. & Cameron, A.D. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science*, 322, 709-713.

Chapter 5 – Section 5.1

The work in Section 5.1 of this thesis has been published in the following paper:

Ma, P., Yuille, H.M., Blessie, V., Göhring, N., Iglói, Z., Nishiguchi, K., Nakayama, J., Henderson, P.J., Phillips-Jones, M.K. (2008). Expression, purification and activities of the entire family of intact membrane sensor kinases from *Enterococcus faecalis*. *Molecular Membrane Biology*, 25(6-7), 449-473.

The work contribution to the above published paper is as follows:-

Sensor kinases name (work carried out by)

EF3197, EF2912, EF1051, EF0373, EF1209 and EF1820 (the candidate); EF2219, EF1704, EF3290, EF1261, EF1863, EF0927, EF2298, EF0570 and EF1335 (Dr Hayley Yuille); EF1194 (Dr Victor Blessie, Miss Nadine Göhring and Miss Zofia Iglói).

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2010 The University of Leeds and Pikyee Ma

Acknowledgements

First of all, I would like to thank my supervisor Professor Peter Henderson for his encouragement, support and invaluable guidance during all these years of my research study. His enthusiasm for and dedication to science has been an inspiration for me, and I would like to express my greatest gratitude to him for giving me the opportunity to accomplish the research in his laboratory.

I would also like to thank Dr. Mary Phillips-Jones for her guidance and advice on the histidine kinases work.

I express my gratitude to Dr Kim Bettaney, Dr Gerda Szakonyi, Dr Hayley Yuille, Dr Victor Blessie, Dr Simon Patching, Dr Massoud Saidijam, Dr Nicholas Rutherford and Miss Halina Norbertczak for their scientific advice and for teaching me some essential scientific techniques.

Thanks are also due to the past and present members of the Henderson, Phillips-Jones and Baldwin labs for providing a wonderful working environment, and for their friendship; including the names mentioned above, I would also like to thank Dr Peter Roach, Mr David Sharples, Miss Preethi Sukumar, Mr Scott Jackson and Miss Lynsey Jones.

Importantly, I thank my family Kwanyiu, Feiyin, Pikwah and Pikki; especially my beloved parents Fongkiu and Kwaiying who have brought me up the way I am and for providing me with their support and encouragement. I also express my gratitude to Mr Yuming Chen for his love, patience and encouragement during the time that I was busy with my work.

Finally, I would like to acknowledge Faculty of the Biological Sciences (University of Leeds), BBSRC, EMeP and Ajinomoto Company for funding. Last, but by no means least, I acknowledge and thank the encouragement received from an endless list of friends.

Abstract

Membrane proteins perform in many critical physiological processes, constituting 25 – 30% of prokaryotic and eukaryotic genomes, and represent up to 70% of current drug targets. However, the number of three dimensional structures of membrane proteins elucidated has been modest in comparison with those of soluble proteins, and this is attributable to three major bottlenecks: expression, purification and structure determination of these hydrophobic proteins.

A genomic approach was conducted to overcome these bottlenecks for two families of membrane proteins: (1) the NCS1 family of transporters and; (2) the histidine kinases of two-component signal transduction systems of *E. faecalis*. The genes encoding these membrane proteins were cloned separately into plasmid pTTQ18-His₆ and subjected to expression trials in *Escherichia coli*. Purification of the recombinant intact proteins from *E. coli* membranes was undertaken using Ni-NTA chromatography, and for proteins that were produced in sufficient quantities crystallisation trials and investigations of structural characteristics were performed. The activities of membrane-bound and/or purified proteins were also tested, and in the case of the histidine kinases, these activity assays were expanded to successfully identify modulating signals. This is the first time that direct *in vitro* activity assays using intact proteins have proven successful for identification of environmental ‘signals’ in enterococci.

In this study, thirteen NCS1 transporters were cloned. Twelve of the thirteen cloned proteins were expressed in *E. coli* membranes, seven were successfully purified, and substrates transported by five of these were confirmed. Three of these transporters, were further characterised, including; (1) a cytosine transporter – CodB, (2) an allantoin transporter – PucI, and (3) a uracil transporter – PA0443, and their substrate specificities and kinetics were also determined.

All sixteen membrane histidine kinases of *E. faecalis* were successfully cloned (six were assigned to me in a joint project with Dr Hayley Yuille). Fifteen of the sixteen histidine kinases were successfully expressed in *E. coli*, and thirteen were purified. Eleven of the fifteen membrane-bound histidine kinases and twelve of the thirteen purified proteins were active in autophosphorylation activity assays. Three histidine kinases were studied further, including; (1) EF3197, which responded to

increasing concentrations of a reducing agent, suggesting a possible role in redox-sensing, (2) EF1051, which was further purified by size exclusion chromatography and showed promising results in crystallisation trials for future structural determination, and (3) EF1820 (FsrC), which was activated in *in vitro* assays by its pheromone signal, gelatinase biosynthesis-activating pheromone (GBAP). This study showed GBAP-induced activity of FsrC was inhibited by the anti-HIV inhibitor, siamycin I, identifying unequivocally and for the first time, the target protein for siamycin I inhibition in the Fsr pathway.

This study presents a successful genomic approach for the production of milligram quantities of membrane proteins leading to characterisation and crystallisation studies. The methods can now be applied to further membrane proteins.

Research outputs

Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., **Ma, P.**, Saidijam, M., Patching, S.G., Hope, R.J., Norbertczak, H.T., Roach, H.T., Iwata, S., Henderson, P.J.F. & Cameron, A.D. (2008). Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science*, 322, 709-713.

Ma, P., Yuille, H.M., Blessie, V., Göhring, N., Iglói, Z., Nishiguchi, K., Nakayama, J., Henderson, P.J., Phillips-Jones, M.K. (2008). Expression, purification and activities of the entire family of intact membrane sensor kinases from *Enterococcus faecalis*. *Molecular Membrane Biology*, 25(6-7), 449-473.

Contents

	Page
Acknowledgements	iii
Abstract.....	iv
Research outputs.....	vi
Contents	vii
Lists of Figures, Tables & Appendices.....	xv
Abbreviations	xx
Chapter 1: Introduction	1
1.1. The bacterial cell membranes	2
1.2. The cytoplasmic membrane of bacteria	4
1.2.1. Functions and properties of the cytoplasmic membrane	4
1.2.2. Proteins in the cytoplasmic membrane	5
1.2.3. Membrane transport	6
1.2.3.1. Passive diffusion	6
1.2.3.2. Facilitated diffusion	6
1.2.3.3. Active transport.....	7
1.2.3.4. Group translocation.....	8
1.2.4. Responses to environmental stress.....	8
1.2.4.1. Two-component signal transduction systems (TCS)	8
1.2.4.2. Extracytoplasmic function σ factors	8
1.3. Membrane transport proteins.....	9
1.3.1. The classification of membrane transport proteins	9
1.3.2. Transporter families for nucleobases and nucleosides.....	9
1.3.2.1. The NCS2/NAT family	10
1.3.2.2. The NCS1/PRT family	10
1.3.2.3. The PUP family.....	10
1.3.2.4. The ENT and CNT families	10
1.3.3. Role of nucleobases in organisms.....	11
1.3.3.1. Nomenclature of nucleobases, nucleosides and nucleotides	11
1.3.3.2. Role of transport in salvaging	12
1.3.4. The NCS1 family of transporters	12

1.3.5. The <i>Microbacterium liquefaciens</i> hydantoin permease (Mhp1)	14
1.3.5.1. Structure of Mhp1	14
1.3.5.2. Substrate and cation binding sites	17
1.3.6. Mechanism of alternating access transport	19
1.3.6.1. Secondary active transporters	19
1.3.6.2. Previous models of alternating access transport	19
1.3.6.3. Emergence of structures of other secondary active transport proteins	20
1.3.6.4. Structures of sodium-coupled transporters	20
1.3.6.5. Six-state model for sodium-coupled transporters with inverted repeats	24
1.4. Two-component signal transduction systems of <i>Enterococcus faecalis</i>	26
1.4.1. Enterococci and <i>Enterococcus faecalis</i>	26
1.4.2. The two-component signal transduction systems	27
1.4.3. Sensor histidine kinase	28
1.4.3.1. Protein phosphorylation in bacteria	28
1.4.3.2. Histidine kinase	29
1.4.3.3. Histidine kinase domains and conserved regions (homology boxes)	30
1.4.4. Histidine kinases of <i>E. faecalis</i>	30
1.4.4.1. Membrane-bound sensor histidine kinases of <i>E. faecalis</i>	30
1.4.4.2. Predicted biological functions	33
1.4.5. Mechanisms of stimulus perception	36
1.4.5.1. Periplasmic-sensing	38
1.4.5.2. Transmembrane-sensing	40
1.4.5.3. HAMP domain	41
1.4.5.4. Cytoplasmic-sensing	41
1.5. Aims and objectives of this work	42
1.5.1. The NCS1 family of transporters	42
1.5.1.1. Importance of the NCS1 proteins	42
1.5.1.2. Aims and research objectives	43
1.5.2. The sensor histidine kinases of <i>Enterococcus faecalis</i>	44
1.5.2.1. Importance of sensor histidine kinases of <i>E. faecalis</i>	44
1.5.2.2. Aims and research objectives	44
1.6. Outline of research	45
Chapter 2: Methods	47
2.1. Sources of Material	48
2.1.1. Antibiotics	48

2.1.2. Bacterial growth media	48
2.1.3. Filters, membranes and chromatography paper	48
2.1.4. Protein assay, membrane preparation, purification, phosphorylation assay, fluorimetry	48
2.1.5. Protein standards	48
2.1.6. Radiolabelled compounds	48
2.1.7. Recombinant DNA technology	49
2.1.8. Scintillation fluid and vials	49
2.1.9. Substrates	49
2.1.10. SDS-PAGE and Western blotting	49
2.2. General Microbiology	50
2.2.1. Bacterial strains used in this study	50
2.2.2. Growth and storage media	50
2.2.3. Plasmid pTTQ18	51
2.3. DNA preparation and manipulation	52
2.3.1. Cloning strategy - overview	52
2.3.2. Primer design and amplification of membrane protein genes	53
2.3.3. Preparation of plasmid pTTQ18-His ₆	55
2.3.4. Agarose gel electrophoresis and restriction digestion of PCR product	56
2.3.5. Quantification of DNA	57
2.3.5.1. The Lambda DNA marker method	57
2.3.5.2. The absorbance method	57
2.3.6. Ligation of gene into plasmid pTTQ18-His ₆	57
2.3.7. Preparation of <i>E. coli</i> competent cells and transformation	58
2.3.7.1. Chung <i>et al.</i> (1989) method	58
2.3.7.2. Inoue <i>et al.</i> (1990) method	58
2.3.8. Transformation of XL-10 Gold cells	59
2.3.9. PCR screening for the presence of the correct gene in a recombinant clone	59
2.3.10. Extraction of plasmid	59
2.3.11. Restriction digestion analysis	59
2.3.12. Automated DNA sequencing	60
2.3.13. Transformation of plasmid into <i>E. coli</i> BL21(DE3) cells	60
2.4. Protein preparation and manipulation	61
2.4.1. Testing for the expression of cloned protein	61
2.4.2. Preparation of total <i>E. coli</i> membranes	61
2.4.3. Determination of protein concentration	61
2.4.4. Analysis of proteins by SDS-PAGE	62

2.4.5. Drying of SDS-PAGE gels onto cellophane	64
2.4.6. Detection of His ₆ -tagged proteins by immunoblotting & chemiluminescent	64
2.4.7. Optimisation of protein expression	65
2.5. Biochemical techniques	65
2.5.1. Preparation of inner membranes using sucrose gradients	65
2.5.2. Solubilisation trials of membrane proteins using detergents	65
2.5.3. Purification of hexahistidine-tagged proteins	66
2.5.3.1. Purification of NCS1 family proteins	66
2.5.3.2. Purification of sensor kinases.....	67
2.5.3.3. Size exclusion chromatography	67
2.5.4. Uptake of radiolabelled substrates by the NCS1 family of transporters	68
2.5.5. Autophosphorylation assays for sensor kinases and signal testing	69
2.5.5.1. Assays using inner membranes	69
2.5.5.2. Assays using purified sensor kinases	69
2.5.5.3. Signal testing.....	69
2.6. Biophysical techniques	70
2.6.1. Circular dichroism studies of secondary structure of purified membrane proteins ...	70
2.6.2. Electrospray ionisation mass spectrometry.....	70
2.6.3. Fluorimetry studies of purified proteins	71
2.6.3.1. Titration by substrate	71
2.6.3.2. Titration by salt	71
2.6.3.3. Titration by HgCl ₂	72
2.6.3.4. Titration by ATP.....	72
2.6.3.5. Titration by GBAP.....	72
2.6.4. N-terminal sequencing	73
2.6.5. Removal of N-terminal formyl group by acid-methanol treatment	73
2.6.6. Crystallisation trials	73
Chapter 3: Overexpression, purification and characterisation of the NCS1 family of transport proteins.....	74
3.1. Cloning, expression and activities of the NCS1 family proteins.....	75
3.1.1. Introduction – A genomic approach.....	75
3.1.2. An homology search and <i>in silico</i> analysis of Mhp1 homologues.....	77
3.1.3. Cloning of genes encoding NCS1 proteins	80
3.1.4. Expression of the NCS1 proteins.....	83
3.1.5. Analysis of the codon usage in the NCS1 proteins	85

3.1.6. Identification of substrates for the NCS1 proteins	86
3.1.7. Discussion: cloning, expression and activities of NCS1 proteins	89
3.2. Characterisation of an <i>Escherichia coli</i> cytosine transporter – CodB (b0336)	91
3.2.1. Introduction – the <i>codBA</i> operon of <i>E. coli</i>	91
3.2.2. The <i>E. coli</i> cytosine transporter (CodB): topology analysis and conserved residues	92
3.2.3. Cloning and expression of CodB	94
3.2.4. Uptake of radiolabelled cytosine by CodB	95
3.2.5. Substrate specificity of CodB	96
3.2.6. Kinetic analysis of cytosine transport by CodB	98
3.2.7. Purification of CodB from the inner membrane	99
3.2.8. Testing the integrity of CodB by circular dichroism	102
3.2.9. Discussion: characterisation of CodB	102
3.3. Characterisation of a <i>Bacillus subtilis</i> allantoin transporter – PucI (Bsu36470)	105
3.3.1. Introduction – nitrogen metabolism & the <i>puc</i> genes of <i>B. subtilis</i>	105
3.3.2. The <i>B. subtilis</i> allantoin transporter (PucI): topology analysis and conserved residues .	109
3.3.3. Cloning and expression of the <i>B. subtilis</i> PucI	110
3.3.4. Time course of uptake of radiolabelled allantoin by PucI	112
3.3.5. Optimisation of uptake of allantoin by PucI	113
3.3.6. Assessing the effect of sodium ions on the uptake of allantoin by PucI.....	115
3.3.7. Substrate specificity of PucI	115
3.3.8. Kinetic analysis allantoin transport by PucI	117
3.3.9. Purification of PucI from the inner membrane	118
3.3.10. Testing the integrity of PucI by circular dichroism	121
3.3.11. Discussion: characterisation of PucI	122
3.4. Characterisation of a <i>Pseudomonas aeruginosa</i> uracil transporter (PA0443)	124
3.4.1. Introduction – <i>Pseudomonas aeruginosa</i> and PA0443	124
3.4.2. The <i>P. aeruginosa</i> uracil transporter (PA0443): topology analysis and conserved residues	125
3.4.3. Cloning and expression of PA0443	126
3.4.4. Optimisation of expression of PA0443	126
3.4.5. Uptake of radiolabelled uracil by PA0443.....	130
3.4.6. Substrate specificity of PA0443.....	131
3.4.7. Kinetic analysis of uracil transport by PA0443	132
3.4.8. Purification of PA0443 from the inner membrane.....	133
3.4.9. Testing the integrity of PA0443 by circular dichroism.....	135
3.4.10. Discussion: characterisation of PA0443	136

Chapter 4: Purification and substrate binding studies of *Microbacterium liquefaciens* hydantoin permease (Mhp1) 140

4.1. Introduction – the hydantoin permease of <i>M. liquefaciens</i>	141
4.2. Assessing the effect of sodium ions on the uptake of 5-L-benzylhydantoin by Mhp1 ...	142
4.3. Purification of Mhp1 from the inner membrane	143
4.4. Fluorimetry studies of Mhp1	145
4.5. Discussion: substrate binding studies of Mhp1	149

Chapter 5: Overexpression and purification of the intact membrane sensor kinases of *Enterococcus faecalis* V583 for activity, structural and ligand binding studies 151

5.1. Evaluation of the expression, purification and activities of all membrane-bound sensor kinases from <i>Enterococcus faecalis</i> V583.....	152
5.1.1. Introduction – the histidine kinases of <i>Enterococcus faecalis</i>	152
5.1.2. Cloning and expression of sensor kinases of <i>E. faecalis</i>	153
5.1.3. Autophosphorylation of <i>E. faecalis</i> histidine kinases in the inner membrane	154
5.1.4. Solubilisation and purification of <i>E. faecalis</i> histidine kinases	158
5.1.5. Integrity of histidine kinases determined by Western blotting, N-terminal sequencing and/or mass spectrometry	159
5.1.6. Autophosphorylation of purified <i>E. faecalis</i> histidine kinases	160
5.1.7. Discussion: expression, purification and activities of the entire family of intact membrane histidine kinases from <i>E. faecalis</i>	163
5.2. Trials to identify the signalling ligands for histidine kinase EF3197	164
5.2.1. Introduction.....	164
5.2.1.1. Histidine kinase EF3197: a putative LytS homologue in <i>E. faecalis</i>	164
5.2.1.2. Bacterial cell lysis	165
5.2.1.3. The <i>lrgAB</i> and <i>cidABC</i> genes.....	166
5.2.1.4. The histidine kinase EF3197 of <i>E. faecalis</i>	167
5.2.2. Cloning, expression and purification of EF3197	170
5.2.3. Circular dichroism of purified EF3197	173
5.2.4. Autophosphorylation activities of EF3197	174
5.2.5. Autophosphorylation of purified EF3197 is not affected by penicillin.	174
5.2.6. Signal testing of EF3197.....	175
5.2.7. Changes in redox potential affect EF3197 autophosphorylation	177
5.2.8. Discussion: towards identification of a signaling ligand for EF3197	178
5.3. Structural studies: crystallisation trials of histidine kinase EF1051.....	182
5.3.1. Introduction.....	182

5.3.1.1. A similar search of histidine kinase EF1051 (EtaS) of <i>E. faecalis</i>	182
5.3.1.2. The EF1051-EF1050 system of <i>E. faecalis</i>	184
5.3.2. Cloning, expression and purification of EF1051	186
5.3.3. Size exclusion chromatography of EF1051	192
5.3.4. Autophosphorylation activities of EF1051	193
5.3.5. Mass spectrometry and circular dichroism spectroscopy studies of EF1051.....	194
5.3.6. Crystallisation trials of EF1051	196
5.3.7. Discussion: crystallisation trial of EF1051	198
5.4. Characterisation of signal interactions by the quorum sensor EF1820 (FsrC).....	201
5.4.1. Introduction.....	201
5.4.1.1. Quorum sensing in bacteria.....	201
5.4.1.2. Comparison of quorum sensing in Gram-negative and Gram-positive bacteria....	201
5.4.1.3. The Fsr quorum sensing pathway of <i>E. faecalis</i>	204
5.4.2. The EF1820 protein (FsrC) of <i>E. faecalis</i>	205
5.4.3. Cloning, expression and purification of FsrC	208
5.4.4. Autophosphorylation activities of FsrC	210
5.4.5. GBAP specifically increases autophosphorylation of FsrC in <i>in vitro</i> autophosphorylation assay.....	210
5.4.6. Siamycin I inhibits GBAP-induced autophosphorylation of FsrC	212
5.4.7. Time course of autophosphorylation of FsrC in the presence and absence of GBAP and/or siamycin I	214
5.4.8. Mass spectrometry of synthetic GBAP	215
5.4.9. Circular dichroism spectroscopy of FsrC.....	216
5.4.10. Fluorimetry study of FsrC.....	218
5.4.11. Cloning of the putative GBAP transporter: EF1821 (FsrB).....	220
5.4.12. Discussion: study of the Fsr quorum sensing pathway of <i>E. faecalis</i>	221
 Chapter 6: Conclusions and future perspectives	 225
6.1. Summary of results – towards structural and functional determination of bacterial transport and sensory membrane proteins	226
6.2. Impact of this research and potential applications	227
6.3. Limitations and future perspectives	229
6.3.1. Future studies related to expression and purification.....	229
6.3.2. Future studies related to determination of function	230
6.3.3. Future studies related to structural determination.....	231
6.4. Concluding remarks	232

References.....	233
Appendices.....	273

Figures, Tables & Appendices

List of Figures

- Figure 1.1 General cell wall structures of bacteria.
Figure 1.2 Diagram of the cytoplasmic membrane.
Figure 1.3 Membrane transport across the bacterial cytoplasmic membrane.
Figure 1.4 Structure of purine and pyrimidine.
Figure 1.5 Structural elements of nucleotide.
Figure 1.6 Topology of Mhp1.
Figure 1.7 Structure of Mhp1.
Figure 1.8 Outward- and inward- facing cavities in Mhp1.
Figure 1.9 Substrate and cation binding sites of Mhp1.
Figure 1.10 Structures of AcrB, GlpT and LacY.
Figure 1.11 Structures of sodium-coupled transporters.
Figure 1.12 Topology of sodium-coupled transporters with inverted repeats of five helices.
Figure 1.13 Common substrate and sodium binding sites of sodium-coupled transporters.
Figure 1.14 A six-state model for sodium-symporters.
Figure 1.15 Two-component signal transduction system.
Figure 1.16 The kinase cores of histidine kinases PhoQ and HK853.
Figure 1.17 Schematic diagram of the sensor histidine kinases of *E. faecalis*.
Figure 1.18 Mechanisms of stimulus perception by sensor kinases.
Figure 1.19 Structures of histidine kinase domains.
Figure 1.20 Periplasmic sensing domains.
Figure 1.21 Cytoplasmic sensing domains.
Figure 2.1 The plasmid pTTQ18.
Figure 2.2 Cloning strategy for membrane proteins using plasmid pTTQ18-His₆.
Figure 2.3 A typical calibration curve using BSA standards.
Figure 2.4 A typical log Mwt-distance curve to determine size of protein on a SDS-PAGE.
Figure 2.5 Procedures for solubilisation trial of membrane proteins.
Figure 3.1 Pipeline of genomic approach for the study of NCS1 family proteins.
Figure 3.2 Phylogenetic tree of Mhp1, its homologues and previously studied NCS1 proteins.
Figure 3.3 NCS1 family genes amplified by PCR.
Figure 3.4 PCR screening for positive colonies.
Figure 3.5 Restriction digestion analysis to confirm the correct size gene in plasmid pTTQ18-His₆.
Figure 3.6 Expression of thirteen NCS1 genes in *E. coli*.
Figure 3.7 Chemical structures of possible substrates for the NCS1 family of transporters.
Figure 3.8 Map of the *E. coli codBA* region.
Figure 3.9 Topology model of *E. coli* cytosine permease (CodB).
Figure 3.10 Growth and expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(*codB*)-His₆.
Figure 3.11 Time course of cytosine uptake by CodB.
Figure 3.12 Pyrimidine catabolic pathways in bacteria.
Figure 3.13 Substrate specificity of CodB.

- Figure 3.14 Common structural features of CodB substrates.
- Figure 3.15 Michaelis-Menten kinetics of ³H-cytosine uptake by CodB.
- Figure 3.16 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) expressing CodB.
- Figure 3.17 SDS-PAGE and Western blot of purification of *E. coli* CodB.
- Figure 3.18 Circular dichroism spectrum of purified CodB of *E. coli*.
- Figure 3.19 Map of the *B. subtilis puc* genes region.
- Figure 3.20 Pathway of nitrogen metabolism in *Bacillus subtilis*.
- Figure 3.21 Topology model of *B. subtilis* allantoin permease (PucI).
- Figure 3.22 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ in LB, 2TY and minimal media.
- Figure 3.23 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆.
- Figure 3.24 Time course of allantoin uptake by PucI.
- Figure 3.25 Effect of induction time with IPTG on allantoin uptake by PucI.
- Figure 3.26 SDS-PAGE and Western blot analysis of total membranes of BL21(DE3) pTTQ18(*pucI*)-His₆ cells induced for different periods of time.
- Figure 3.27 Effect of sodium ions on allantoin uptake by cells expressing PucI.
- Figure 3.28 Substrate specificity of PucI.
- Figure 3.29 Common structural features of PucI substrates.
- Figure 3.30 Michaelis-Menten kinetics of ¹⁴C-allantoin uptake by PucI.
- Figure 3.31 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) cells expressing PucI.
- Figure 3.32 SDS-PAGE of total and solubilised inner membranes of *E. coli* BL21(DE3) cells expressing PucI.
- Figure 3.33 SDS-PAGE and Western blot of purification of *B. subtilis* PucI.
- Figure 3.34 Circular dichroism spectrum of purified PucI of *B. subtilis*.
- Figure 3.35 Map of the *P. aeruginosa* PA0443 gene region.
- Figure 3.36 Topology model of *P. aeruginosa* uracil permease (PA0443).
- Figure 3.37 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ in LB, 2TY and M9 minimal media.
- Figure 3.38 Expression of *P. aeruginosa* uracil permease (PA0443) in cells grown in 2TY and M9 minimal medium.
- Figure 3.39 Growth and expression of *P. aeruginosa* uracil permease (PA0443) in a range of IPTG concentrations, in minimal medium.
- Figure 3.40 Time course of uracil uptake by PA0443.
- Figure 3.41 Substrate specificity of PA0443.
- Figure 3.42 Common structural features of PA0443 substrates.
- Figure 3.43 Michaelis-Menten kinetics of ³H-uracil uptake by PA0443.
- Figure 3.44 SDS-PAGE and Western blotting of fractions obtained from the preparation of inner membranes of *E. coli* expressing PA0443.
- Figure 3.45 SDS-PAGE of purification of *P. aeruginosa* PA0443.
- Figure 3.46 Circular dichroism spectrum of purified PA0443 of *P. aeruginosa*.
- Figure 4.1 Topology diagram of *Microbacterium liquefaciens* hydantoin permease (Mhp1).
- Figure 4.2 Effect of sodium ions on L-benzylhydantoin uptake by cells expressing Mhp1.
- Figure 4.3 SDS-PAGE of purification of Mhp1.
- Figure 4.4 Electrospray ionisation mass spectrometry of Mhp1.
- Figure 4.5 Quenching of tryptophan fluorescence in Mhp1 with benzylhydantoin.
- Figure 4.6 Quenching of tryptophan fluorescence in Mhp1 with NaCl.

- Figure 4.7 Quenching of tryptophan fluorescence in Mhp1 by different cations.
- Figure 4.8 Quenching of tryptophan fluorescence in Mhp1 by different anions.
- Figure 4.9 Quenching of tryptophan fluorescence in Mhp1 by benzylhydantoin in the presence and absence of HgCl₂.
- Figure 4.10 Quenching of tryptophan fluorescence in Mhp1 by HgCl₂.
- Figure 5.1 Expression of the 16 intact membrane-bound histidine kinases of *E. faecalis* in *E. coli* BL21(DE3).
- Figure 5.2 Autophosphorylation activities of intact histidine kinases in inner membrane vesicles.
- Figure 5.3 One-step purification of the expressed intact membrane histidine kinases of *E. faecalis*.
- Figure 5.4 Autophosphorylation activities of purified intact histidine kinases.
- Figure 5.5 LytSR two-component regulatory system of *S. aureus*.
- Figure 5.6 Model of the holin and endolysin function.
- Figure 5.7 Map of the LytST two-component regulatory system in *E. faecalis*.
- Figure 5.8 Topology model of *E. faecalis* histidine kinase EF3197.
- Figure 5.9 Phylogenetic tree of EF3197, LytS and redox sensors.
- Figure 5.10 Domain analysis of EF3197 and LytS of *S. aureus* and *Bacillus* strains.
- Figure 5.11 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF3197)-His₆.
- Figure 5.12 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) cells expressing EF3197.
- Figure 5.13 SDS-PAGE of fractions obtained from the purification of *E. faecalis* EF3197.
- Figure 5.14 Circular dichroism spectrum of purified EF3197.
- Figure 5.15 Autophosphorylation of purified EF3197 in response to the presence and absence of penicillin.
- Figure 5.16 *In vitro* signal screening for EF3197.
- Figure 5.17 Effect of reducing agent on autophosphorylation activity of EF3197 *in vitro*.
- Figure 5.18 A model of the *lrg*- and *cid*- regulated cell wall metabolism.
- Figure 5.19 Phylogenetic tree of EF1051, LisK, CsrS and ArlS.
- Figure 5.20 Domain analysis of EF1051 and its homologues.
- Figure 5.21 Map of the EF1051-EF1050 two-component system in *E. faecalis*.
- Figure 5.22 Topology model of *E. faecalis* sensor kinase EF1051.
- Figure 5.23 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆ in LB, 2TY and minimal media.
- Figure 5.24 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆.
- Figure 5.25 Growth and expression of *E. faecalis* EF1051 using a range of IPTG concentrations in LB medium.
- Figure 5.26 SDS-PAGE and Western blot of membrane fractions of *E. coli* BL21(DE3) cells expressing EF1051.
- Figure 5.27 SDS-PAGE of total and solubilised inner membrane of *E. coli* BL21(DE3) cells expressing EF1051.
- Figure 5.28 SDS-PAGE and Western blot of fractions from purification of *E. faecalis* EF1051.
- Figure 5.29 Size exclusion chromatography of purified EF1051.
- Figure 5.30 Calibration of the S200 column.
- Figure 5.31 Effect of buffer composition and detergent on autophosphorylation activities of purified EF1051.

Figure 5.32	Electrospray ionisation mass spectrometry of EF1051.
Figure 5.33	Circular dichroism spectrum of purified EF1051.
Figure 5.34	Microscopic images of crystals from the crystallisation trial of EF1051.
Figure 5.35	The LuxIR quorum sensing circuit in Gram-negative bacteria.
Figure 5.36	Quorum sensing regulation in enterococci.
Figure 5.37	Topology model of <i>E. faecalis</i> histidine kinase FsrC.
Figure 5.38	Phylogenetic tree of FsrC and Agr family kinases.
Figure 5.39	Domain analysis of FsrC and its homologues.
Figure 5.40	SDS-PAGE and Western blot of membrane fractions of <i>E. coli</i> BL21(DE3) cells expressing FsrC.
Figure 5.41	SDS-PAGE of total and solubilised inner membranes of <i>E. coli</i> BL21(DE3) cells expressing FsrC.
Figure 5.42	SDS-PAGE of fractions isolated during the purification of <i>E. faecalis</i> FsrC.
Figure 5.43	Structures of GBAP and siamycin I.
Figure 5.44	Effect of the GBAP peptide pheromone on autophosphorylation activity of FsrC and other purified histidine kinases of <i>E. faecalis</i> .
Figure 5.45	Inhibition of GBAP-induced autophosphorylation of FsrC by siamycin I.
Figure 5.46	Time course of autophosphorylation of FsrC in the presence and absence of GBAP and/or siamycin I.
Figure 5.47	Electrospray ionisation mass spectrometry of GBAP.
Figure 5.48	Circular dichroism spectrum of purified FsrC.
Figure 5.49	Circular dichroism spectrum of purified FsrC with ATP and GBAP.
Figure 5.50	Circular dichroism spectrum of purified FsrC with GBAP.
Figure 5.51	Tryptophan fluorescence measurements of FsrC titrated by ATP.
Figure 5.52	Tryptophan fluorescence measurements of FsrC titrated with GBAP.
Figure 6.1	Structure determination pathway.

List of Tables

Table 1.1	The nucleobase-cation-symporter-1 (NCS1) family.
Table 1.2	Histidine kinases of <i>Enterococcus faecalis</i> V583.
Table 2.1	<i>E. coli</i> strains used for cloning and expression.
Table 2.2	Growth media for bacterial cultures.
Table 2.3	Primer sequences for PCR amplification of NCS1 family genes.
Table 2.4	Primer sequences for PCR amplification of <i>E. faecalis</i> sensor kinase genes.
Table 2.5	DNA molecular weight markers used for determination of DNA concentration and size.
Table 2.6	Ligation reactions using various plasmid:insert ratios.
Table 2.7	Restriction digestion reactions to verify the present of target gene.
Table 2.8	Primers used for sequencing of plasmid constructs.
Table 2.9	Composition of SDS-PAGE gel with 15% resolving and 4% stacking gels.
Table 2.10	SDS-PAGE molecular weight markers (Sigma).
Table 2.11	High-range rainbow molecular weight markers (Amersham Biosciences).
Table 2.12	Composition of mini-gels (12% resolving and 4% stacking gels).
Table 2.13	Titration by substrate.
Table 2.14	Titration by salts.
Table 2.15	Titration by HgCl ₂ .

Table 2.16	Titration by ATP.
Table 2.17	Titration by GBAP.
Table 3.1	Proteins of the NCS1 family studied in this project.
Table 3.2	Summary of the topology predictions determined by TMHMM (Krogh <i>et al.</i> , 2001) for the selected NCS1 proteins.
Table 3.3	Expression level, predicted and observed size of the NCS1 transporters.
Table 3.4	Substrate screening of the NCS1 family transporters.
Table 3.5	Current status of the selected NCS1 family proteins.
Table 3.6	The regulatory proteins involved in nitrogen metabolism in <i>B. subtilis</i> .
Table 5.1	Expression conditions and solubilisation of the 16 intact histidine kinases of <i>E. faecalis</i> .
Table 5.2	N-terminal sequencing and ESI-MS of histidine kinases of <i>E. faecalis</i> .

Appendices

Appendix 1	Topology prediction by TMHMM of the selected NCS1 proteins.
Appendix 2	Sequence alignment of Mhp1 and its homologues.
Appendix 3	DNA sequences of the recombinant NCS1 clones.
Appendix 4	Amino acid sequences of the recombinant NCS1 clones.
Appendix 5	Comparison of the codon usage of the NCS1 family of transporters and the <i>E. coli</i> expression host.
Appendix 6	GC contents and rare codons in the NCS1 family of transporters.
Appendix 7	DNA sequences of the recombinant histidine kinases clones.
Appendix 8	Amino acid sequences of the recombinant histidine kinases clones.
Appendix 9	Sequence alignment of EF3197 and LytS.
Appendix 10	Sequence alignment of EF1051, LisK, CsrS and ArlS.
Appendix 11	Sequence alignment of EF1820 (FsrC) and AgrCs.
Appendix 12	DNA and amino acid sequences of recombinant <i>fsrB</i> .

Abbreviations

a.a.	amino acid
APC	amino acid, polyamine and organo cation
APS	ammonium persulphate
ATP	adenosine 5-triphosphate
BCCT	betaine/carnitine/choline transporter
<i>bla</i>	β -lactamase
BSA	bovine serum albumin
CD	circular dichroism
CNT	concentrative nucleoside transporter
C-terminus	carboxyl terminus
DAACS	dicarboxylate/amino-acids:cation symporter
DDM	<i>n</i> -dodecyl- β -D-maltoside
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	2-deoxy nucleoside-5-triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
<i>endA</i>	endonuclease deficient
ECF	extracytoplasmic function
ENT	equilibrative nucleoside transporter
GBAP	gelatinase-biosynthesis activating peptide
HAMP	histidine kinses, adenylyl cyclases, methyl binding proteins and phosphatases
HAWP	hydrophobic acetate white plain (filters)
HK	histidine kinase
IPTG	isopropyl- β -D-thiogalactopyranoside
<i>lacI</i> ^q	LacI repressor gene
<i>lacZ</i> α	β -galactosidase alpha fragment gene
LB	Luria Bertani culture media
LP	lipoprotein
LPS	lipopolysaccharides
LTA	lipoteichoic acid
MCS	multiple cloning site
MFS	Major Facilitator Superfamily
MES	2-(N-morpholino)ethanesulphonic acid
Mhp1	<i>Microbacterium liquefaciens</i> hydantoin permease-1
Mwt	molecular weight
Nac	nitrogen assimilation control
NAG	<i>N</i> -acetylglucosamine
NAM	<i>N</i> -acetylmuramate
NAT	nucleobase-ascorbate transporter
NCS1	nucleobase-cation-symporter-1
NCS2	nucleobase-cation-symporter-2
Ni-NTA	nickel-nitriloacetic acid
NSS	neurotransmitter-sodium-symporter
N-terminus	amino terminus
ORFs	open reading frames

PAS	Per-Arnt-Sim
PCR	polymerase chain reaction
PDC	PhoQ-DcuS-CitA
PEP	phosphoenolpyruvate
PRT	purine-related transporter
psi	pounds per square inch
<i>ptac</i>	<i>tac</i> promoter
PUP	purine-related permease
PVDF	poly (vinylidene) difluoride
<i>recA</i>	recombination deficient
RND	resistant nodulation cell division
rpm	revolution per minutes
RR	response regulator
<i>rrnB</i> t12	transcription terminator
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	standard error mean
SSS	solute-sodium-symporter
<i>tac</i>	trp-lac
TAE	tris acetate EDTA
Taq	thermal polymerase
TBSTT	Tris Buffer Saline with Triton-X100 and Tween 20
TCA	trichloroacetic acid
TCDB	transport classification database
TCS	two-component signal transduction system
TEMED	tetramethylethylenediamine
Triton-X100	nonaethylene glycol octylphenol ether
TWEEN 20	polyoxyethylenesorbitan monolaurate
v/v	volume by volume
w/v	weight by volume
w/w	weight by weight

Chapter 1

Introduction

Chapter 1 – Introduction

The first bacteria were observed by Anton van Leeuwenhoek in 1674 using a single-lens microscope that he designed (Porter, 1976). Bacteria are single cell organisms classified as prokaryotes and are the most abundant organism on Earth. There are approximately five nonillion (5×10^{30}) bacteria on Earth, forming much of the world's biomass (Whitman *et al.*, 1998). In the human body, there are approximately ten times as many bacteria cells than there are human cells, with the majority residing on the skin and in the gut (Sears, 2005). Most of the human microflora resides harmlessly in the body, some are beneficial, and a few species are pathogenic and could cause infectious diseases. The bacteria being able to live in a wide variety of conditions, consequently must have evolved mechanisms for their survival. One form of protection for the bacteria is their cell membrane, which forms a barrier against the harmful external environment. But due to impermeability of the cytoplasmic membrane to hydrophilic molecules, transport of nutrients into the cell or the extrusion of toxins and waste products wholly rely upon integral membrane proteins. In addition, bacteria also have two-component signal transduction systems, a mechanism that enables the bacteria to sense and respond to their environment (Stock *et al.*, 2000). This thesis focuses on two different families of membrane proteins: 1) nucleobase-cation-symporter-1 (NCS1) family of transporters; and 2) the sensor histidine kinase family that forms part of the two-component signal transduction system.

1.1. The bacterial cell membranes

Microorganisms are grouped into either prokaryotes or eukaryotes according to their cellular structure. Prokaryotic cells do not have subcellular organelles separated from the cytoplasm by phospholipid membranes like eukaryotes do; thus they do not have a nucleus, mitochondria or endoplasmic reticulum. Most bacteria are classified into Gram-positive or Gram-negative based on their reaction with an iodine-based stain discovered by Hans Christian Gram that allows visualisation of bacteria in tissues under ordinary bright-field microscope (Gram, 1884).

The Gram-positive cell wall is characterised by the presence of a thick layer of peptidoglycan that comprises as much as 90% of the compact cell wall (Figure 1.1). Embedded in the Gram-positive cell wall are polyalcohols called teichoic acids, some of which are attached to lipid to form lipoteichoic acids (LTA), which anchors the

peptidoglycan layer to the cytoplasmic membrane (Figure 1.1). The peptidoglycan is comprised of alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramate (NAM) saccharides (Meroueh *et al.*, 2006). The main difference between the two classes of bacteria is that the Gram-positive lacks an outer membrane, and therefore is more sensitive to hydrophobic antibiotics, ionophores, lysozyme, and hydrolytic enzymes. This also has less constraint on the diffusion of solutes to the cytoplasmic membrane.

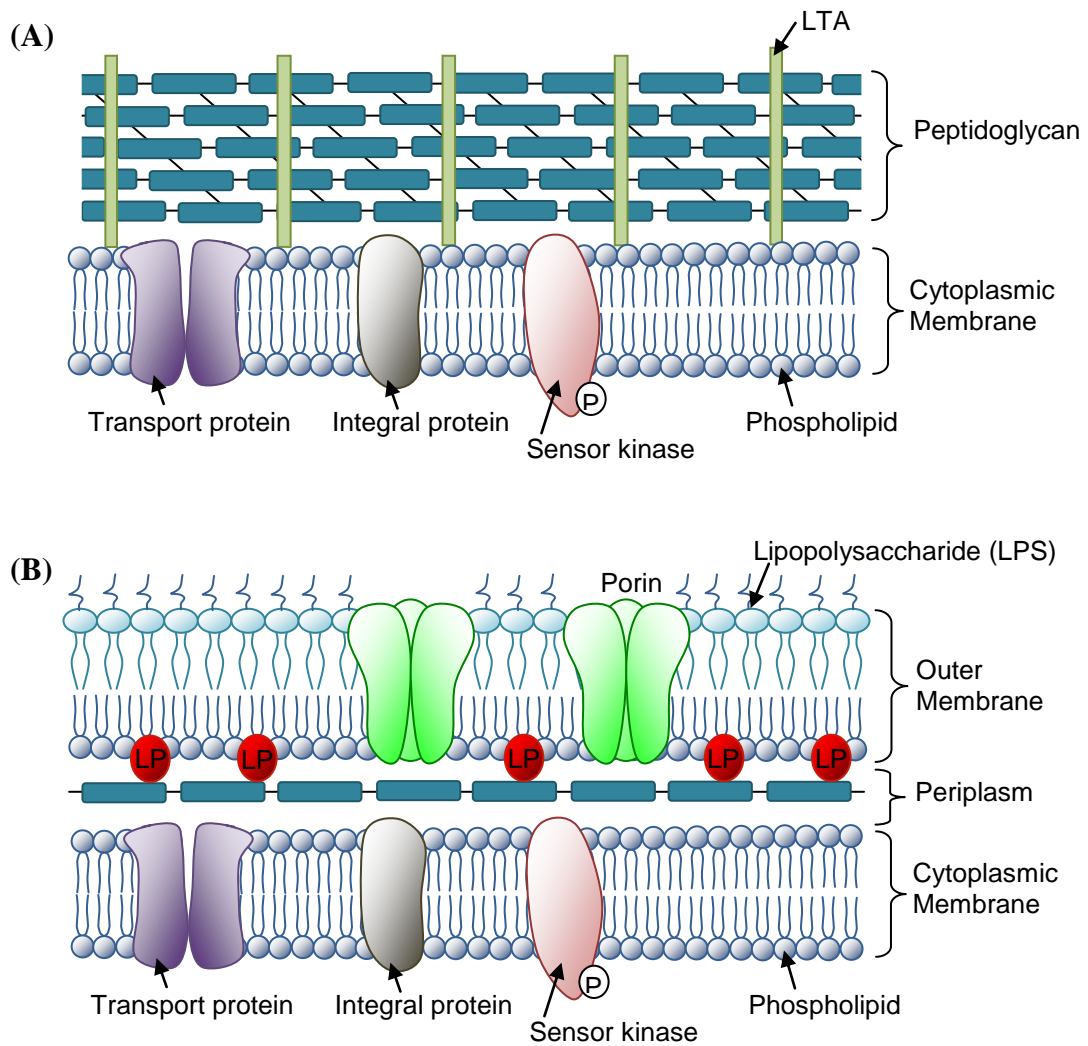


Figure 1.1 General cell wall structures of bacteria. (A) Gram-positive bacteria. Lipoteichoic acids (LTA) are present only in Gram-positive bacteria, and these polysaccharides extend through the entire peptidoglycan layer and emerge on the cell surface. (B) Gram-negative bacteria. The outer membrane is comprised of a layer of lipopolysaccharide (LPS), found only in Gram-negative bacteria, which is hydrophilic in nature and provides a barrier against hydrophobic compounds; and a layer of lipoprotein (LP) that stabilises the cell wall structure. The lipid part of LP is embedded in the outer membrane, and the sugar part is covalently bound to murein.

In contrast to the Gram-positive cell wall, the Gram-negative cell wall has a thin layer of peptidoglycan, comprising only 5 – 20% of the cell wall thickness, and this is responsible for its inability to retain the crystal violet stain upon decolourisation with ethanol during Gram-staining. In addition, the Gram-negative cell wall has an outer membrane consisting of a layer of phospholipids and a layer of lipopolysaccharide (LPS), with high levels of porins, which function as a molecular sieve through which molecules with a molecular mass ≥ 600 -1000 Da cannot penetrate (Figure 1.1) (Nikaido & Vaara, 1985; Nikaido, 2003). Porins allow for passive transport of many ions, sugars and amino acids across the outer membrane. These molecules enter into the periplasm then undergo binding with substrate binding proteins, hydrolysis by periplasmic enzymes and/or translocation across the cytoplasmic membrane by specific carriers that are integral proteins. Both types of bacteria have an overall negative charge derived from the phosphodiester bonds between teichoic acid monomers in Gram-positive, and the LPS in Gram-negative organisms.

1.2. The cytoplasmic membrane of bacteria

1.2.1. Functions and properties of the cytoplasmic membrane

The cytoplasmic membrane is made of approximately 35 – 50% phospholipid and 50 – 65% protein, forming a ‘sea’ of phospholipids with ‘islands’ of proteins, named the fluid mosaic model (Figure 1.2) (Singer & Nicolson, 1972). The phospholipid bilayer acts as the selective permeability barrier for most molecules. It isolates the cell contents (cytoplasm) from the external environment, but at the same time nutrients and wastes are transported into and out of the cell (Section 1.3). Other important physiological processes associated with the cytoplasmic membrane include solute transport, maintenance of electrochemical gradients and ATP synthesis, signal transduction (Section 1.4) (Facey & Kuhn, 2010), and synthesis of cell surface structures and excretion of proteins.

Phospholipids are amphipathic molecules consisting of hydrophobic hydrocarbon tails and a hydrophilic head group (Kim & Gadd, 2008). In water, the hydrophilic head interacts with water and meanwhile the hydrophobic tail hides away, forming micelles or liposomes. Due to this property, broken membranes can spontaneously restore their structure to form vesicles. Membrane vesicles are often prepared in laboratories for study. This includes using French press (Vanderheiden *et al.*, 1970) or sonication for inside-out vesicles, or osmotic lysis of protoplast by

enzymes for right-side-out vesicles (Altendorf & Staehelin, 1974). The phospholipid bilayer is permeable to hydrophobic solutes and water but not to charged solutes. Water can diffuse slowly across the phospholipid bilayer but rapid flux is mediated by a protein called aquaporin.

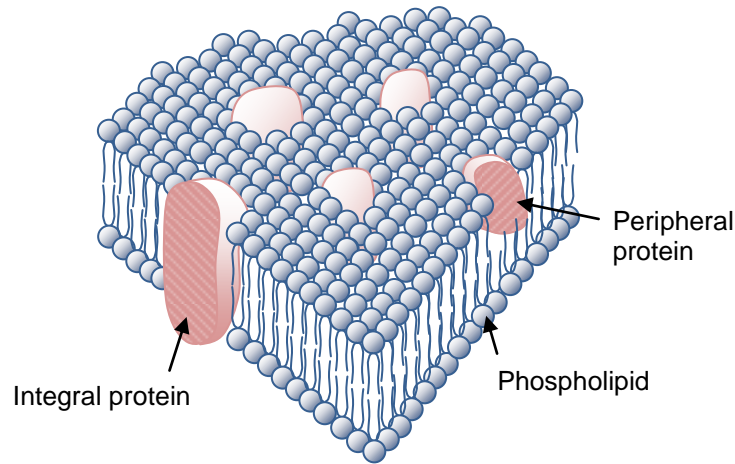


Figure 1.2 Diagram of the cytoplasmic membrane. The proteins float around in the phospholipid bilayer forming a structure called the fluid mosaic model.

1.2.2. Proteins in the cytoplasmic membrane

Proteins in the cytoplasmic membrane have many biological functions and they are either peripheral or integral according to their location in the membrane (Figure 1.2). Integral proteins are hydrophobic and are therefore embedded in the membrane via hydrophobic interaction with the hydrophobic tails of the phospholipids, and they could be removed by use of detergents. However, the peripheral proteins could be removed by washing with salt solutions as they are attached to the membrane by ionic interactions. There are various important functions carried out by the proteins in the cytoplasmic membrane and examples include: the water channels called aquaporin that are involved in osmoregulation of the cell (Tanghe *et al.*, 2006); enzymes such as phospholipase, protease, and peptidase for the renewal and synthesis of the cell wall (Scheffers & Pinho, 2005); enzymes involved in DNA replication since the bacterial DNA is attached to the cytoplasmic membrane (Ganesan & Lederberg, 1965; Smith & Hanawait, 1967); proteins that maintain the energy status of the cell (Hellingwerf *et al.*, 1985; Konings *et al.*, 1994), i.e. proteins of the electron transport system, and ATPase enzyme complex involved in creating the proton motive force. Many of the membrane proteins have a role in protein secretion or solute transport (Section 1.2.3), and there are membrane-bound sensor histidine kinases of the two-component signal transduction system (Section 1.2.4) involved in sensing and responding to environmental stimuli.

1.2.3. Membrane transport

Genes encoding membrane proteins constitute 25 – 35% of prokaryotic and eukaryotic genomes and have critical physiological roles in the cells. The importance of transport systems in bacteria was recognised more than half a century ago (Gale & Taylor, 1947). As mentioned earlier in the chapter, the cytoplasmic membrane forms a permeability barrier to hydrophilic solutes. Therefore membrane transport processes are mediated by integral membrane proteins, and often they function in conjunction with extracytoplasmic receptors or receptor domains together with cytoplasmic energy-coupling and regulatory proteins or protein domains (Saier, 2000). Membrane transport processes were best understood in Gram-negative bacteria *Salmonella typhimurium* and *Escherichia coli* because of the ease of their genetic manipulation. Nevertheless, many of the transport processes first identified in these organisms are also found in higher organisms, including man (Henderson, 1998).

The data generated by whole-genome sequencing projects show that approximately 5 – 12% of the complete bacterial genome codes for proteins involved in membrane transport (Paulsen *et al.*, 1998; Saier, 1999; Paulsen *et al.*, 2000; Ren *et al.*, 2004; Saier *et al.*, 2006; Ren & Paulsen, 2007). In order to survive and function, cells must exchange substances such as nutrients, ions and metabolites with their environment. Transport processes involved with nutrient uptake are associated with the processes of oxidative and photosynthetic ATP synthesis (Figure 1.3) (Mitchell, 1961; Mitchell, 1966) and are conventionally classified as: passive diffusion, facilitated diffusion, primary active transport, secondary active transport, and group translocation (Figure 1.3), and these are defined below.

1.2.3.1. Passive diffusion is the translocation of solute across a membrane down its electrochemical gradient without the need of a transport protein, following Fick's law ($v = Pac$, where P is the permeability coefficient for a particular solute, a is the area, and c is the difference in solute concentration across the membrane) and where velocity has a linear relationship to the concentration of the solute.

1.2.3.2. Facilitated diffusion involves the translocation of a solute across a membrane down its electrochemical gradient catalysed by a transport protein, and follows Michaelis-Menten relationship that relates the initial rate of transport (v) to initial substrate concentration ($[S] = c$ at zero time). Where V_{\max} is the maximum velocity and K_m is the concentration of substrate when $v = (V_{\max} / 2)$.

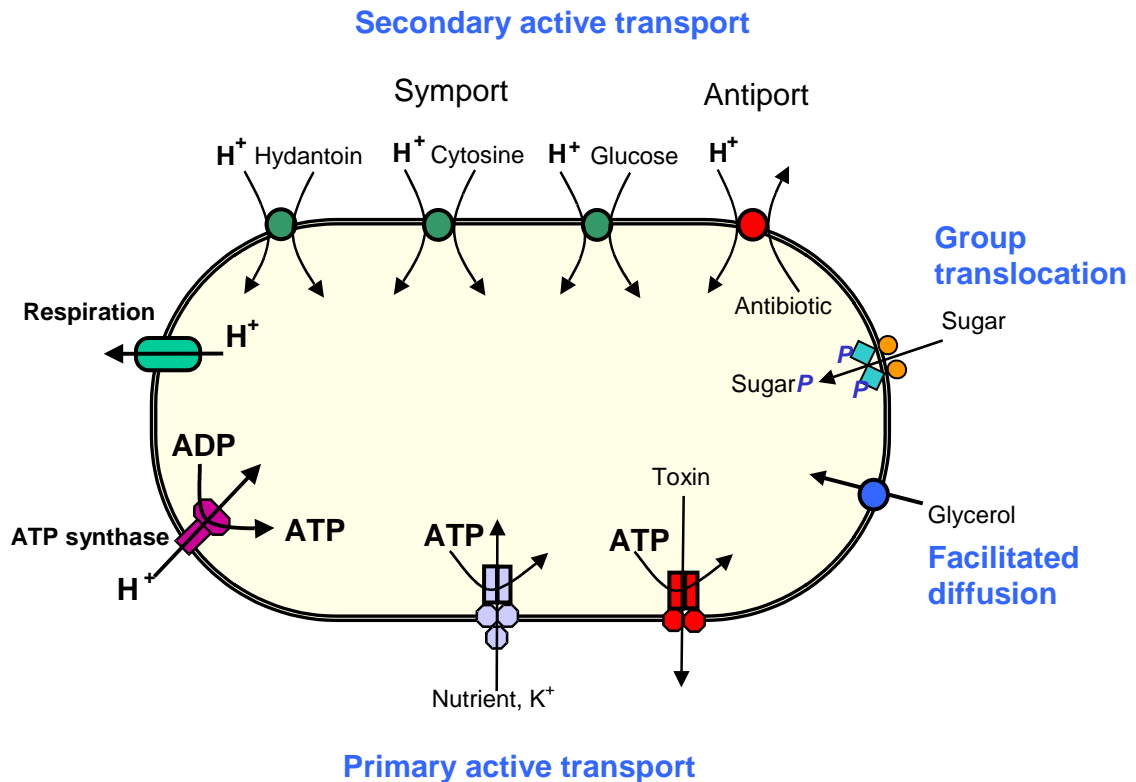


Figure 1.3 Membrane transport across the bacterial cytoplasmic membrane. (diagram modified from one of Peter J.F. Henderson). The large oval represents the cytoplasmic membrane of the microorganism. ATP hydrolysis or respiration of the cell generates a transmembrane electrochemical gradient of protons (top left). This can be utilised by the proton-nutrient symport or proton-substrate antiport system (along the top). An ATP-dependent primary transport system for uptake of K^+ or efflux or toxin is shown (bottom). A phosphotransferase mechanism involving phosphoenolpyruvate (PEP) for accumulation of sugars is shown (upper right), and uniport system for transport of glycerol (lower right).

1.2.3.3. Active transport involves the net transport of a solute across a biological membrane from a low to a high electrochemical potential. Energy expenditure is coupled to transmembrane substrate translocation without chemical modification, where the energy may be derived from light, ATP hydrolysis, or pre-existing solute gradients. There are two type of active transport, namely primary and secondary. **Primary active transport** involves the direct conversion of energy into an electrochemical potential of solute across the cell membrane. Consequently, the transport involves hydrolysis of ATP, light, or redox energy (Figure 1.3), with the translocation of protons or sodium ions driven by oxidation of respiratory substrates. Many nutrient uptake systems in bacteria involve direct utilisation of ATP and binding proteins. **Secondary active transport** entails the conversion of a pre-existing electrochemical gradient (Na^+ or H^+

ion), into a new electrochemical gradient of the transported species. The source of this energy is the ion gradient set up by primary chemical or photochemical conversion. In *E. coli*, H⁺ gradients are found to energise many transport reactions, and Na⁺ gradients are relatively few, though Na⁺ gradients are found predominantly in organisms living in salt environments.

1.2.3.4. Group translocation is another form of active transport and involves chemical modification of the solute during its transport across the cell membrane so that once the solute is inside, it becomes impermeable and it remains within the cell. One example includes the phosphotransferase system; a high-energy phosphate group from phosphoenolpyruvate (PEP) is transferred by a series of enzymes to glucose.

1.2.4. Responses to environmental stress

Bacteria are able to live and survive in unique and/or fluctuating environments, in which nutrient and toxin levels, acidity, osmolarity, temperature, and many other conditions can change unexpectedly and rapidly. To survive, cells must constantly sense and respond to changes in their environment. To achieve this, cells sense environmental change and transmit the information from the cell surface to the cytoplasm, a process that requires signal transduction across biological membranes. The two most common transmembrane signal-transducing mechanisms in bacteria are two-component systems (TCS) and extracytoplasmic function (ECF) σ factors, where both systems consist of two proteins.

1.2.4.1. Two-component signal transduction systems (TCS). TCS are widely distributed amongst bacteria and constitute the major mechanism of signal transduction. Signal transduction involves phosphorylation as a mean of information transfer and requires a sensor kinase, often located in the cytoplasmic membrane, that monitors changes in the environment, and a cytoplasmic response regulator that mediates an adaptive response, often causing a change in gene transcription (Hoch & Silhavy, 1995; Stock *et al.*, 2000). The mechanisms of stimulus perception by sensor kinase include: extracellular or periplasmic-sensing, transmembrane region-sensing and cytoplasmic sensing; and will be described in more detail in Section 1.4.5.

1.2.4.2. Extracytoplasmic function σ factors. Extracytoplasmic function σ factors (ECF) signal transduction involves a transmembrane sensory protein (anti- σ factor) that

detects a specific stimulus, and a cognate cytoplasmic effector protein (σ factor) that mediates the cellular response (Brooks & Buchanan, 2008). This system differs from the TCS in the way the two proteins communicate and their output. TCS entails a phosphorylation and dephosphorylation event of both histidine kinase and response regulator. In contrast, ECF signal transduction occurs via protein-protein interaction. For example, in the absence of a signal the anti- σ factor binds tightly to the σ factor and holds it in an inactive state. When a specific stimulus is present, the σ factor is released and binds to promoter sequences upstream of its target genes. The ECF σ factor can only function as an activator of expression, whereas the TCS is able to inhibit as well as induce gene expression. After the one- and two-component systems (Stock *et al.*, 2000; Ulrich *et al.*, 2005), ECF represents the third fundamental mechanism of bacterial signal transduction, with an average of six such regulators per bacterial genome (Staroń *et al.*, 2009).

1.3. Membrane transport proteins

1.3.1. The classification of membrane transport proteins

The transporter classification database (TCDB) is a comprehensive classification system for membrane transport proteins and is freely accessible from the web (Saier *et al.*, 2006). Approximately 3000 distinct proteins from all varieties of known organisms are classified into over 550 transporter families based on the transport classification system. The transport systems are classified based on five components and these are V, W, X, Y and Z. Firstly, each category of transporter is assigned a number in relation to its molecular mode of transport (V), and energy coupling mechanism (W). Next, the protein is further classified on the basis of its amino acid sequence and phylogenetic family and subfamily (X and Y), and its substrate specificity (Z).

1.3.2. Transporter families for nucleobases and nucleosides

Nucleobase and nucleoside transporters are classified into five distinct families (de Koning, 2000). These include nucleobase-ascorbate transporter (NAT), microbial purine-related transporter (PRT), plant family of purine-related permeases (PUP), equilibrative nucleoside transporter (ENT) and concentrative nucleoside transporter (CNT), all of which are electrochemical potential-driven transporters (secondary active transporters).

1.3.2.1. The NCS2/NAT family. This family is classified as 2.A.40 according to the TCDB system (Saier *et al.*, 2006) and is the largest and most conserved family for nucleobases. It includes members from eubacteria, archaea, fungi, plants, insects, nematodes and mammals, including human. Some of the described transporters of the NAT family include: purine transporter of *Bacillus*, PbuX (Christiansen *et al.*, 1997); purine transporter of *Aspergillus*, UapA and UapC (Gorfinkiel *et al.*, 1993; Diallinas *et al.*, 1995); uracil transporter of *Escherichia coli* (*E. coli*) and *Bacillus*, UraA (Andersen, 1995) and PyrP (Ghim & Neuhard, 1994; Turner *et al.*, 1994) respectively; and the mammalian (human and rat) ascorbate transporters, SVCT1 and SVCT2 (Faaland *et al.*, 1998; Daruwala *et al.*, 1999; Rajan *et al.*, 1999; Tsukaguchi *et al.*, 1999). An alternative designation for the NAT family is the nucleobase-cation-symporter-2 (NCS2) family.

1.3.2.2. The NCS1/PRT family. This family is not as broadly distributed compared to the NAT family, and is described in further detail in Section 1.3.4.

1.3.2.3. The PUP family consists of members found only in plants. The members studied so far, includes the 21 *Arabidopsis* PUPs, several of which are known to transport adenine, cytosine or secondary compounds such as cytokinins and caffeine (Gillissen *et al.*, 2000; Bürkle *et al.*, 2003). This family does not share any similarity of amino acid sequence with the NAT or PRT proteins.

1.3.2.4. The ENT and CNT families. These families are nucleoside transporters mainly described in mammals, and are classified as belonging to 2.A.57 and 2.A.41, respectively (Saier *et al.*, 2006). Homologues of the ENT family have also been sequenced from yeast, protozoa and nematodes. Several members of the ENT have been functionally characterised, including: monoamine transporter, PMAT (Engel *et al.*, 2004); nitrobenzylthioinosine-insensitive ENT transporter, *eiENT* (Griffiths *et al.*, 1997); nucleoside transporter of *Tyranosoma brucei*, TbAT1 (Mäser *et al.*, 1999); hENT1 and rENT1 (Sundaram *et al.*, 1998); and nucleoside transporter of *Leishmania*, LdNT1 (Vasudevan *et al.*, 1998). Members of the CNT family include proteins derived from Gram-negative and Gram-positive bacteria as well as yeast and animals. In bacteria and yeast, members of the CNT family are energised by H⁺ symport, whereas in mammals they are energised by Na⁺ symport. These transporters also exhibited different specificities for nucleosides. Some members of the CNT family include:

NupC and NupG of *E. coli* (Craig *et al.*, 1994; Westh Hansen *et al.*, 1987); NupG of *B. subtilis* (Johansen *et al.*, 2003); CaCNT of *Candida albicans* (Loewen *et al.*, 2003); hCNT1 of human (Ritzel *et al.*, 1997); and hCNT3 of human (Smith *et al.*, 2005).

1.3.3. Role of nucleobases in organisms

Nucleobases are essential components of the genetic material in all living organisms and play a crucial role in cell metabolism, contributing to base pairings in DNA and RNA. They form part of nucleotides, which are the building blocks for RNA and DNA, and precursors for essential coenzymes (NAD⁺, FAD and CoA) as well as ATP and GTP that are the main forms of chemical energy of the cell. In addition, they serve as nitrogen sources in plants and many microorganisms.

1.3.3.1. Nomenclature of nucleobases, nucleosides and nucleotides. Nucleobases are nitrogen heterocycles (*N*-heterocycles) and are either purine (adenine and guanine) or pyrimidine (cytosine, thymine and uracil), abbreviated as A, G, C, T and U, respectively (Figure 1.4). Their structures are also shown in Figure 3.7. Nucleosides are glycosylamines comprised of a nucleobase bound to a pentose sugar moiety, either ribose or deoxyribose via a beta-glycosidic linkage. This includes for example adenosine, guanosine, cytidine, thymidine, uridine and inosine. Nucleotides are phosphate esters of nucleosides and contain one to three phosphate groups (Figure 1.5).

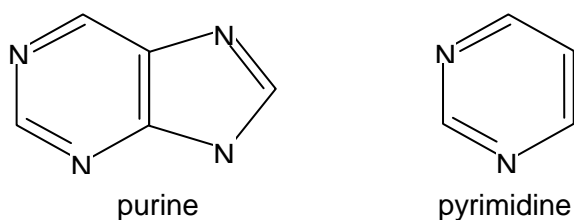


Figure 1.4 Structure of purine and pyrimidine. Purine and pyrimidine are nucleobases and are nitrogen heterocycles.

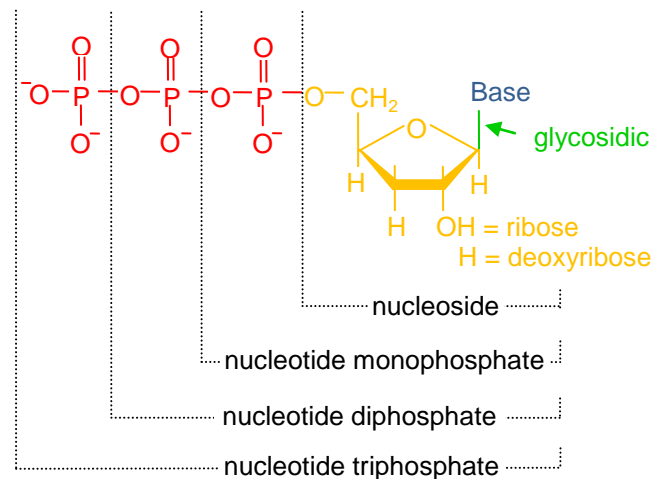


Figure 1.5 Structural elements of nucleotide. A nucleobase (blue) together with a pentose sugar (orange) forms a nucleoside. A nucleotide is composed of a nucleoside and one to three phosphate groups.

1.3.3.2. Role of transport in salvaging

Degradation of nucleotides rarely occurs inside the cells at site of high demand, for example replication of DNA for growth of cells. Therefore, the acquisition of nucleobases or nucleosides from adjacent cells or extracellular sources relies on nucleobase transport proteins (Section 1.3.2). In microorganisms, the uptake of purines and pyrimidines fulfils two main functions. The first is to scavenge exogenous bases for nucleotide biosynthesis and the second for catabolism as many purines can serve as nitrogen sources (de Koning & Diallinas, 2000).

1.3.4. The NCS1 family of transporters

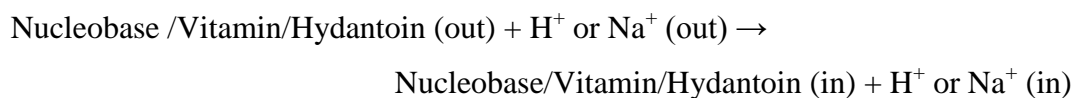
The NCS1 family, also named the PRT family (Section 1.3.2), is classified as 2.A.39 according to the TCDB classification (Saier *et al.*, 2006). This family of transporters consists of more than 1000 currently sequenced proteins derived from Gram-negative and Gram-positive bacteria, archaea, yeast, fungi and plants. Transporters of this family in yeast and bacteria are widely divergent and do not cluster closely on the NCS1 family phylogenetic tree (Figure 3.2). There are currently eight proteins that have been studied from yeast, three of which (Dal4, Fur4 and Fui1) cluster tightly together, another two (Thi10 and Ntr1) also cluster closely together, and there are the three other yeast proteins (Fcy2, Fcy21 and Tpn1) which cluster quite closely together with a fungal protein (FcyB). Proteins of this family from bacteria were also

studied (Table 1.1), including *E. coli* (CodB), *Pseudomonas putida* (CytX), *Bacillus subtilis* (PucI), *Arthrobacter aurescens* (HyuP), and *Microbacterium liquefaciens* (Mhp1). These NCS1 proteins are classified according to the TCDB (Saier *et al.*, 2006) (Table 1.1).

TC number	Name Accession no.	Organism Type	Example	Reference
2.A.39.1.1	Cytosine permease P0AA82	Bacterium	CodB of <i>Escherichia coli</i>	Danielsen <i>et al.</i> (1992)
2.A.39.1.2	The putative hydroxyl- methylpyrimidine transporter B1JAG2	Bacterium	CytX of <i>Pseudomonas putida</i>	Rodionov <i>et al.</i> (2002)
2.A.39.2.1	Cytosine-purine permease P17064	Yeast	Fcy2 of <i>Saccharomyces cerevisiae</i>	Pinson <i>et al.</i> (1997) Rodriguez <i>et al.</i> (1995)
2.A.39.2.2	The vitamine B ₆ :H ⁺ symporter P53099	Yeast	Tpn1 of <i>Saccharomyces cerevisiae</i>	Stolz & Vielreich (2003)
2.A.39.2.3	The hypoxanthine/ adenine/ guanine (purine) transporter Q708J7	Yeast	Fcy21 of <i>Candida albicans</i>	Goudela <i>et al.</i> (2006)
2.A.39.2.4	The cytosine-purine- scavenging protein B1PXD0	Fungus	FcyB of <i>Asperillus nidulans</i>	Vlanti & Diallinas (2008)
2.A.39.3.1	Allantoin permease Q04895	Yeast	Dal4 of <i>Saccharomyces cerevisiae</i>	Yoo <i>et al.</i> (1992)
2.A.39.3.2	Uracil/uridine permease P05316	Yeast	Fur4 of <i>Saccharomyces cerevisiae</i>	Smits <i>et al.</i> (1994)
2.A.39.3.3	Urindine:H ⁺ symporter YBL042	Yeast	Fui1 of <i>Saccharomyces cerevisiae</i>	Wagner <i>et al.</i> (1998) Zhang <i>et al.</i> (2006)
2.A.39.3.4	Allantoin permease P94575	Bacterium	PucI of <i>Bacillus subtilis</i>	Schultz <i>et al.</i> (2001)
2.A.39.3.5	The probable hydantoin permease Q9F467	Bacterium	HyuP of <i>Arthrobacter aurescens</i>	Wiese <i>et al.</i> (2000)
2.A.39.3.6	Hydantoin permease 2JLN_A	Bacterium	Mhp1 of <i>Microbacterium liquefaciens</i>	Suzuki & Henderson (2006) Weyand <i>et al.</i> (2008) Shimamura <i>et al.</i> (2010)
2.A.39.4.1	Thiamine permease Q05998	Yeast	Thi10 of <i>Saccharomyces cerevisiae</i>	Enjo <i>et al.</i> (1997)
2.A.39.4.2	The nicotinamide riboside transporter Q08485	Yeast	Nrt1 of <i>Saccharomyces cerevisiae</i>	Belenky <i>et al.</i> (2008)

Table 1.1 The nucleobase-cation-symporter-1 (NCS1) family. This table shows some current members of the NCS1 family in the TCDB. Table adapted from the TCDB website (Saier *et al.*, 2006).

To date, there is only one structure of a NCS1 family transporter elucidated and that is the hydantoin permease of *Microbacterium liquefaciens* (Mhp1) (Weyand *et al.*, 2008; Shimamura *et al.*, 2010). This transporter is described in further detail (Section 1.3.5). The NCS1 transporters are essential components of the salvage pathways for nucleobases and related metabolites and catalyse the generalised transport reaction:



Transporters of the NCS1 family are 419 – 635 amino acid residues long and are predicted to possess twelve putative transmembrane helices. Some of the members of this family were shown to function in substrate-proton symport, but one example, Mhp1, was shown to be sodium-dependent (Weyand *et al.*, 2008). Symporters are membrane proteins that couple energy stored in electrochemical potential gradients to mediate the cotransport of ions and molecules into the cell (Poolman & Konings, 1993). In sodium-coupled symporters, sodium is the driving force that causes the conformation change. In humans, the malfunctioning of sodium-coupled transporters has been implicated in diseases such as glucose-galactose malabsorption, familial renal glucosuria (Wright *et al.*, 2007). Others, such as the sodium-iodide transporter, involved in active iodide transport into the thyroid follicular cell have medical significance (Dohan *et al.*, 2003).

1.3.5. The *Microbacterium liquefaciens* hydantoin permease (Mhp1)

Mhp1 was first discovered by Suzuki *et al.* (2005b) in a study of proteins of the *hyu* operon of *M. liquefaciens*. His study of the soluble protein, hydantoin racemase (Suzuki *et al.*, 2005a; Suzuki *et al.*, 2005b) led to the identification of Mhp1, which was characterised and shown to transport 5-substituted hydantoin (Suzuki & Henderson, 2006). The stereoselective hydrolyses of DL-5-monosubstituted hydantoin compounds have been widely studied for the production of optically pure D- or L-amino acids (Owaga *et al.*, 1997; Altenbuchner *et al.*, 2001), which have long been of interest in industry for production of various drugs and pharmaceuticals (Syldatk *et al.*, 1990; Syldatk *et al.*, 1992; Bommarius *et al.*, 1998; Syldatk *et al.*, 1999).

1.3.5.1. Structure of Mhp1. The first structures of a NCS1 protein, a secondary active transporter, Mhp1 was solved to 2.85 Å and 4.0 Å of resolution (Weyand *et al.*, 2008), providing an insight into the alternating access mechanism. Mhp1 mediates the uptake

of indolylmethyl- and benzyl-hydantoin into *M. liquefaciens* (Section 1.3.5) (Suzuki & Henderson, 2006). The structures of the outward-facing open and substrate-bound occluded conformations of Mhp1 were solved, revealing twelve transmembrane helices, ten of which are arranged in two inverted repeats of five helices (Figure 1.6), highlighting how the outward-facing cavity closes upon binding of substrate (Figure 1.7 and 1.8). This structure was compared to the leucine transporter of *Aquifex aeolicus* (LeuT) (Yamashita *et al.*, 2005; Singh *et al.*, 2007; Zhou *et al.*, 2007) and the galactose transporter of *Vibrio parahaemolyticus* (vSGLT) (Faham *et al.*, 2008), revealing that the outward- and inward-facing cavities are symmetrically arranged on opposite sides of the membrane. The reciprocal opening and closing of these cavities is synchronised by the inverted repeat helices 3 and 8, providing the structural basis of the alternating access model for membrane transporter (Weyand *et al.*, 2008; Shimamura *et al.*, 2010).

Mhp1 has between 17 – 31% identity and 52 – 65% similarity to the NCS1 proteins studied in this work (Table 3.1). Meanwhile, it shares 16% identity and 50% similarity to vSGLT, and 15% identity and 48% similarity to LeuT, falling just below the range with NCS1 proteins. Despite the low similarity of the amino acid sequence of Mhp1 to the solute-sodium-symporter (SSS) and neurotransmitter-sodium-symporter (NSS) families of transporters, the X-ray structure of Mhp1 (Weyand *et al.*, 2008) revealed similarities between them, suggesting that the NCS1 proteins are closely related to the SSS and NSS family proteins.

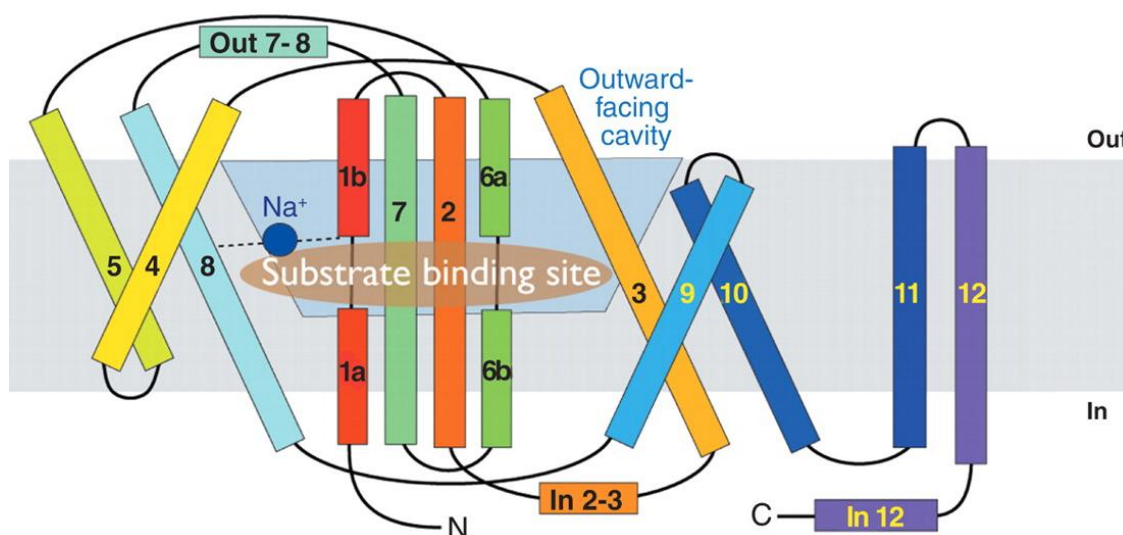


Figure 1.6 Topology of Mhp1. The positions of the substrate- and cation-binding sites are indicated as brown ellipsoid and a blue circle, respectively. The phospholipid bilayer is shown in grey, and the outward-facing cavity of the structure is highlighted in light blue rhombus. In the three dimensional structure, helices 3 and 8 lie very closely onto each other. Diagram taken from Weyand *et al.* (2008).

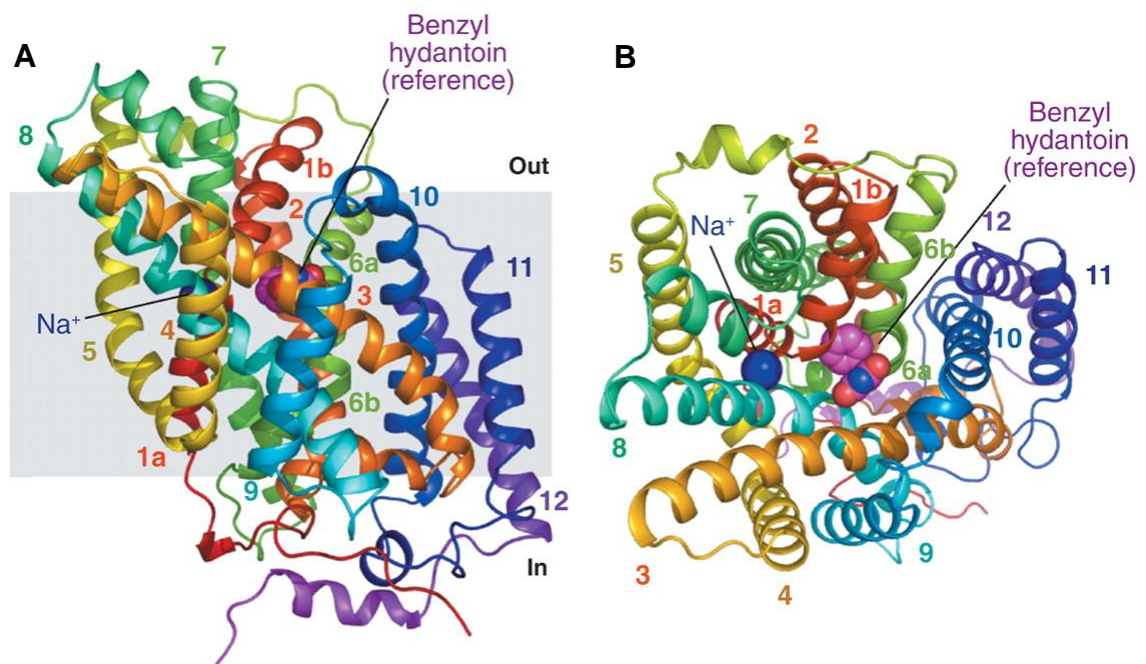


Figure 1.7 Structure of Mhp1. The Mhp1 structure viewed: (A) in the plane of the membrane and (B) from outside of the membrane. These images are based on the high-resolution structure of Mhp1 with benzylhydantoin. The position of the benzylhydantoin in the structure is shown as reference. A sodium ion is also shown and labelled. Diagram taken from Weyand *et al.* (2008).

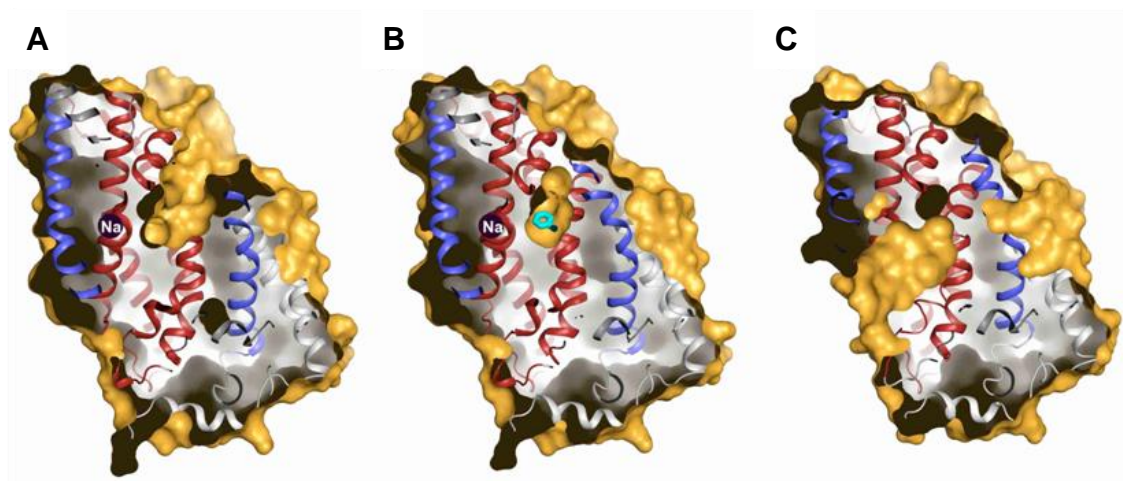


Figure 1.8 Outward- and inward-facing cavities in Mhp1. A slice through the Mhp1 structures, viewed parallel to the membrane is shown. (A) Mhp1 outward-facing structure (substrate-free). (B) Mhp1 substrate-occluded with benzylhydantoin bound. (C) Mhp1 substrate-occluded with the inward-facing cavity shown. The benzylhydantoin is shown in cyan. Diagram taken from Shimamura *et al.* (2010).

The SSS family proteins catalyse the solute- Na^+ symport. The solutes transported could be sugars, amino acids, choline, nucleosides, inositols, vitamins or urea (Saier *et al.*, 2006). Meanwhile the NSS family proteins catalyse uptake of a variety of neurotransmitters, amino acids, osmolytes and related nitrogenous substances by a Na^+ symport mechanism; and sometimes Cl^- or K^+ is cotransported (Saier *et al.*, 2006).

The inverted repeat units intertwine and forms a central four-helix bundle consisting of two broken helices (1 and 6) associated with helices 2 and 7 (Figure 1.6). This four-helix bundle comprises the central block between the V-shaped structures formed by helices 4 and 5 and helices 9 and 10, and the core for substrate and cation binding (Figure 1.6 and 1.7) Helices 11 and 12 are poorly conserved and their role remains unclear.

1.3.5.2. Substrate and cation binding sites. Benzylhydantoin interacts with amino acids W117 (helix 3) and W220 (helix 6) of Mhp1; the hydantoin moiety of benzylhydantoin forms a π -stacking with the indole ring of W117 and hydrogen bonds to the N318 and Q121, while the side chain of W220 forms π -stacking with the benzyl moiety (Figure 1.9A), forming a L-shaped density in the electron density map (Figure 1.9B) (Weyand *et al.*, 2008). Residues W117 and N318 are conserved in the NCS1 family (Appendix 2). N314 is also conserved, which could hold the N318 side chain in position to interact with the substrate (Weyand *et al.*, 2008).

A mutant of Fcy2, a member of the NCS1 family that transports cytosine into *S. cerevisiae*, with substitutions in the segment (residues 371 – 377 IANNIPN) of Fcy2 showed altered kinetics of substrate uptake (Chevallier *et al.*, 1975; Bloch *et al.*, 1992; Brèthes *et al.*, 1992). These residues correspond to 311 – 318 in Mhp1. Site-directed mutagenesis on these hydrophilic residues in Fcy2 further suggested their role in substrate binding, and are equivalent to N314 and N318 in Mhp1 (Ferreira *et al.*, 1997). When benzylhydantoin is bound to Mhp1, helix 10 moves into the outward-facing cavity (Figure 1.9, C and D), a position which occludes the substrate-binding site from the outside space of the membrane (Figure 1.8B and 1.9C), possible triggered by the re-positioning of W220 (helix 6) (Weyand *et al.*, 2008). This outward-facing substrate-occluded conformation prevents leakage of molecules across the membrane.

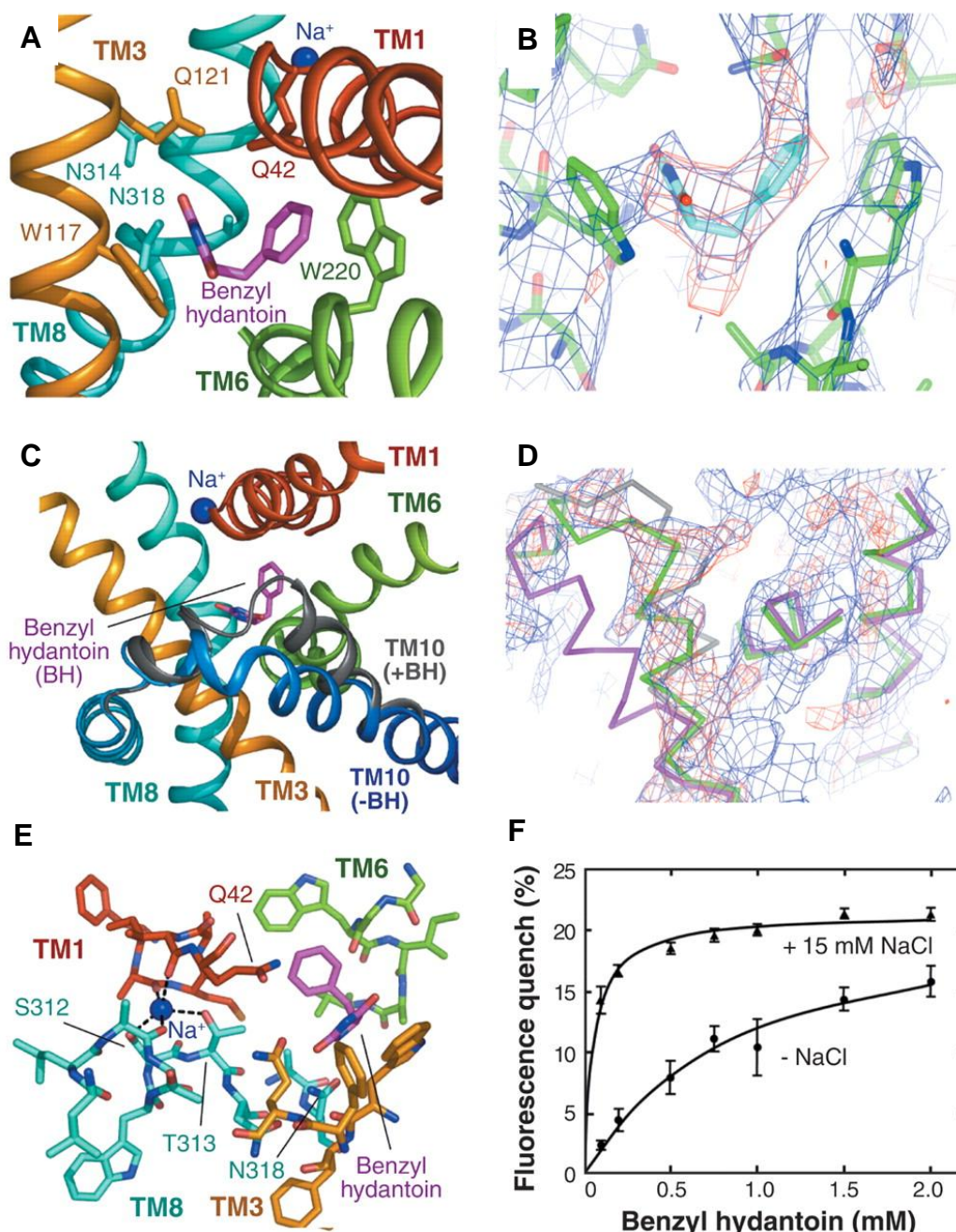


Figure 1.9 Substrate and cation binding sites of Mhp1. This Figure is taken from Weyand *et al.* (2008). (A) Benzylhydantoin and the key residues in the binding site are shown. The benzylhydantoin and sodium are shown in magenta and blue respectively. (B) Electron density associated with benzylhydantoin. (C) Conformational change upon substrate binding. A shift in helix 10 is shown, which occludes the bound benzylhydantoin from the outside of the membrane. Helix 10 with no benzylhydantoin (dark blue) and with benzylhydantoin bound (grey). (D) Electron density and model for helix 10. The benzylhydantoin free structure is shown in magenta and benzylhydantoin bound structure in green. The vSGLT structure (PDB accession code 3DH4) is superimposed onto Mhp1 (grey trace). (E) Helices and key residues in the cation and substrate binding sites. (F) Tryptophan fluorescence quenching of Mhp1 by benzylhydantoin (This graph forms part of the work in Chapter 4). Mhp1 was titrated with benzylhydantoin and the decrease in tryptophan fluorescence at 348 nm was monitored. The titration was performed with and without sodium. Error bars represent the mean of triplicates.

There is one sodium binding site in Mhp1. At this site, the carbonyl-oxygen-atoms of A38 and I41 of helix 1 and the carbonyl-oxygen-atom of A309, and the hydroxyl-oxygen-atoms of the side chains of S312 and T313, together form the cation-binding pocket (Figure 1.9E) (Weyand *et al.*, 2008). The dipole moment at the C-terminus of helix 1a possibly contributes to the binding, as seen in other transporters (Dutzler *et al.*, 2002; Yernool *et al.*, 2004; Hunte *et al.*, 2005; Yamashita *et al.*, 2005; Faham *et al.*, 2008; Ressler *et al.*, 2009). Mhp1 was confirmed to be sodium-dependent by a tryptophan fluorescence quenching study (Figure 1.9F – forms part of the work in Chapter 4). An increase in affinity of Mhp1 for benzylhydantoin was shown in the presence of sodium.

1.3.6. Mechanism of alternating access transport

1.3.6.1. Secondary active transporters. Secondary active transporters are present in all species throughout the kingdoms of life (Sobczak & Lolkema, 2005). They participate in many physiological processes in the human, including the uptake of nutrients in the intestines (Wright & Turk, 2004), transport of Na⁺ and Cl⁻ in the kidney (Hebert *et al.*, 2004), and the reuptake of neurotransmitters from the synaptic cleft (Sonders *et al.*, 2005). They also form the fundamental translocation unit for solutes and ions in bacteria (Nikaido & Saier Jr, 1992). Secondary active transporters are biochemically well characterised, including the lactose permease, LacY (Abramson *et al.*, 2003; Mirza *et al.*, 2006; Guan *et al.*, 2007; Smirnova *et al.*, 2008; Zhou *et al.*, 2008), the best-studied representative of the Major Facilitator Super family (MFS). Other members of the MFS, such as multidrug transporter (EmrD) and glycerol-3-phosphate transport (GlpT) (Huang *et al.*, 2003; Yin *et al.*, 2006) are also biochemically and structurally characterised, though the mechanism of secondary transport has not yet been fully established.

1.3.6.2. Previous models of alternating access transport. The mechanism of alternating access, a central concept for membrane transport could be explained by models. Peter Mitchell proposed an early theory for the mechanism of secondary transporters, suggesting two alternating structural states are involved. One in which the substrate binding site is accessible to the extracellular solution (open-to-out), and another wherein the binding site is accessible to the cytoplasm (Mitchell, 1957). In 1960, Crane discovered sodium-glucose cotransport in intestinal glucose absorption, the first proposal of flux coupling in biology (Crane *et al.*, 1960). Later, Jardezy (1966),

proposed a substrate binds to one face of the membrane protein and an energy input causes a conformational change to expose the substrate to the opposite face of the membrane, followed by release. Tanford (1983) proposed a similar mechanism, suggesting there are at least two states and binding of substrate could cause a change in conformation.

1.3.6.3. Emergence of structures of other secondary active transport proteins.

Models for transport mechanisms were previously based only on biochemical data, but with the emergence of membrane protein structures, models based on atomic structure are now available. The first crystal structure of a secondary transporter was reported in 2002, the proton-multidrug efflux pump AcrB of *E. coli*, of the resistant nodulation cell division (RND) family (Murakami *et al.*, 2002). In 2003, two further MFS structures were solved, GlpT (Huang *et al.*, 2003) and LacY (Abramson *et al.*, 2003). Their structures are shown in Figure 1.10.

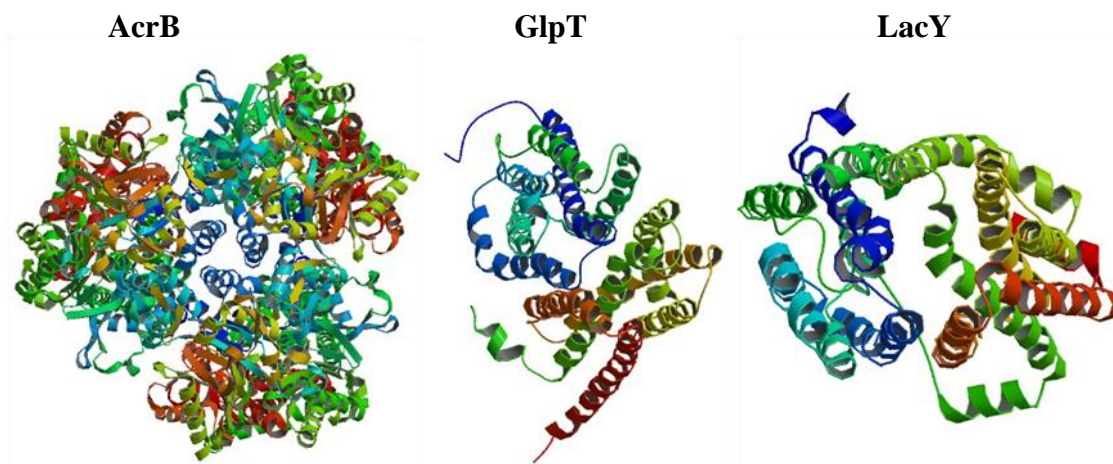


Figure 1.10 Structures of AcrB, GlpT and LacY. AcrB at 3.5 Å (PDB code 1IWG), GlpT at 3.3 Å (PDB code 1PW4), and LacY at 3.5 Å (PDB code 1PV6).

1.3.6.4. Structures of sodium-coupled transporters. In 2004, the first atomic structures of a sodium-coupled secondary transporter, GltPh, of the dicarboxylate/amino-acid:cation symporter (DAACS) family, that transported glutamate were reported (Yernool *et al.*, 2004); followed by two other structures in 2005, the Na⁺/H⁺ antiporter NhaA of *E. coli* (Hunte *et al.*, 2005) and the Na⁺-leucine transporter LeuT (Yamashita *et al.*, 2005). These structures of sodium-coupled transporters belonging to different membrane protein families revealed unique folds in their structure (Figure 1.11), though they all shared an internal two-fold structural pseudo-symmetry and discontinuous transmembrane helices (Screpanti & Hunte, 2006). The

recent structures of three more sodium-coupled secondary transporters belonging to different families; vSGLT, Mhp1 and BetP surprisingly showed the same protein fold seen in LeuT (Figure 1.11 and 1.12).

Membrane protein classification has been based on sequence homology and functional properties, but the crystal structures of four sodium-symporter proteins of unrelated sequences revealed a common architecture of inverted repeats of five transmembrane helices (Figure 1.12) (Yamashita *et al.*, 2005; Faham *et al.*, 2008; Weyand *et al.*, 2008; Ressler *et al.*, 2009). LeuT was the first high-resolution crystal structure of a sodium symporter with two inverted repeats of five helices (Yamashita *et al.*, 2005), having two sodium binding sites and solved in the outward-facing substrate occluded conformation. The second structure with the inverted five helices motif is the sodium-galactose transporter vSGLT (Faham *et al.*, 2008); the protein was in an inward facing conformation with a hydrophobic gate blocking the sugar exit into the hydrophilic channel leading to the cytoplasm. The structures of the outward facing LeuT and inward facing vSGLT hinted at a structural basis for the alternating access mechanism of these sodium-symporters. Mhp1 was the third reported transporter with inverted repeats of five helices (Weyand *et al.*, 2008). Two structures of Mhp1 were solved, as mentioned earlier (Section 1.3.5), one in the outward-facing conformation without substrate and the other in the outward-facing occluded conformation with substrate. A third structure of Mhp1 in the inward-facing state was recently solved (Shimamura *et al.*, 2010) and provides a unique conformation. Another quite recent structure of a sodium symporter with the same structural motif, the sodium-betaine transporter BetP of the betaine/carnitine/choline transporter (BCCT) family, in an intermediate between the outward and inward facing occluded conformations showed a unique conformation in the sodium-coupled transport mechanism (Ressler *et al.*, 2009).

These four sodium-symporters with inverted repeats belong to different gene families, but they have the same structural core. They have an inverted topology where helices 1 – 5 are related to helices 6 – 10 by a two-fold symmetry and the central helices (helices 1, 2, 3, 6, 7, 8 and 10) are intertwined to make the substrate and ion binding sites, forming the structural hallmark for these sodium-symporters. This could be an alternative for classification of transporters, rather than classification by primary sequence alone. Though the relative substrate-binding sites in these sodium-symporters are similar (Figure 1.13), the specific interactions vary consequently to accommodate the different substrates (Abramson & Wright, 2009). For vSGLT, all OH-moieties of

galactose are coordinated by hydrogen bonds from polar side chains in the central helices (helices 1, 6, 7 and 10). Meanwhile for BetP, betaine interacts with three tryptophan and one tyrosine residues in helices 2 and 6. For Mhp1, benzylhydantoin sits between two tryptophan residues in helices 3 and 6, and is coordinated by conserved glutamine and asparagine residues in helices 1 and 8. Lastly, LeuT coordinates leucine through the main chain in the unwound segments of helices 1 and 6. These transporters are highly conserved in terms of their 3D structure, especially in helices 1, 6 and 8, which are involved in their substrate and cation binding.

Other structures with inverted repeat of five helices include: the Na⁺ independent amino acid transporter ApcT (Shaffer *et al.*, 2009) of the Amino acid, Polyamine and organo-Cation (APC) family; the arginine-agmatine antiporter AdiC (Fang *et al.*, 2009) of the APC family; the L-carnitine/ γ -butyrobetaine exchanger CaiT of BCCT family (Tang *et al.*, 2010).

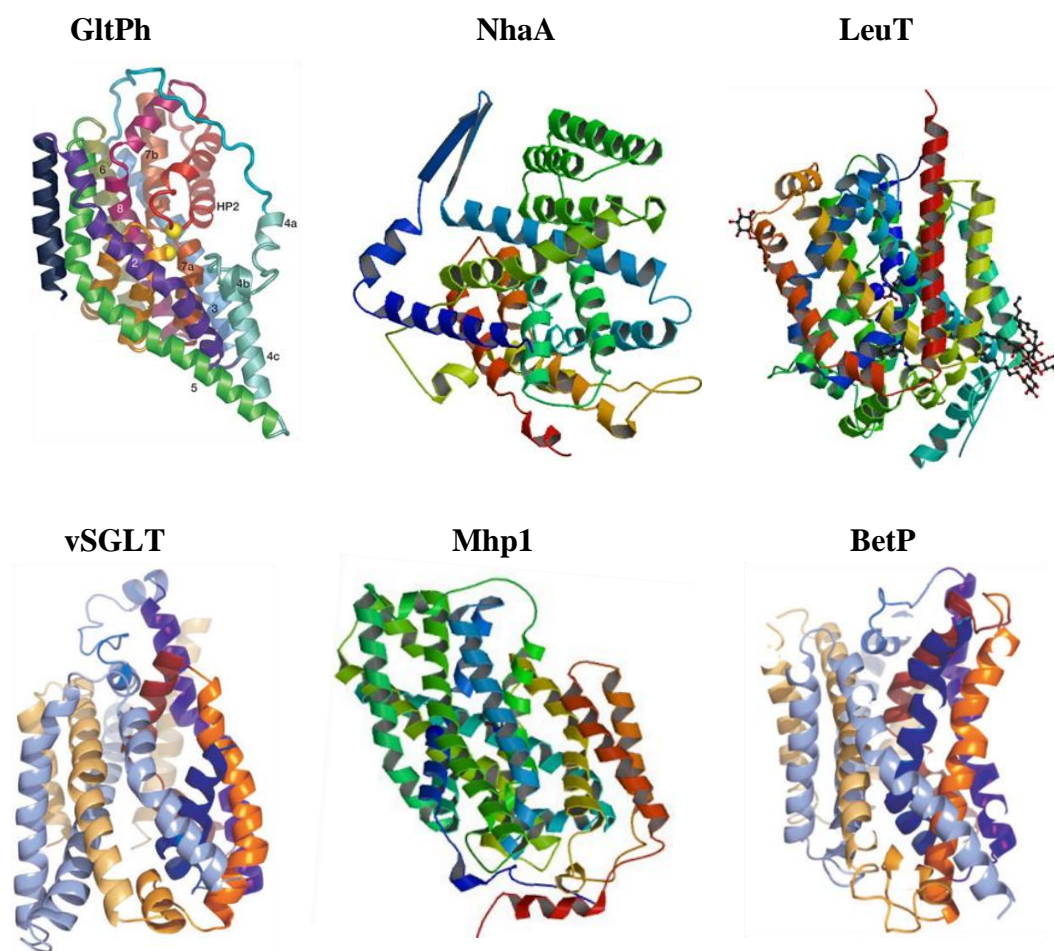


Figure 1.11 Structures of sodium-coupled transporters. GltPh at 3.5 Å (PDB code 1XFH) (Yernool *et al.*, 2004), NhaA at 3.45 Å (PDB code 1ZCD), LeuT complex with L-leucine, sodium and desipramine at 1.9 Å (PDB code 2QB4), vSGLT with bound galactose at 2.7 Å (PDB code 3DH4), Mhp1 with bound benzylhydantoin at 4 Å (PDB code 2JLO), BetP with bound glycine betaine at 3.35 Å (PDB code 2WIT).

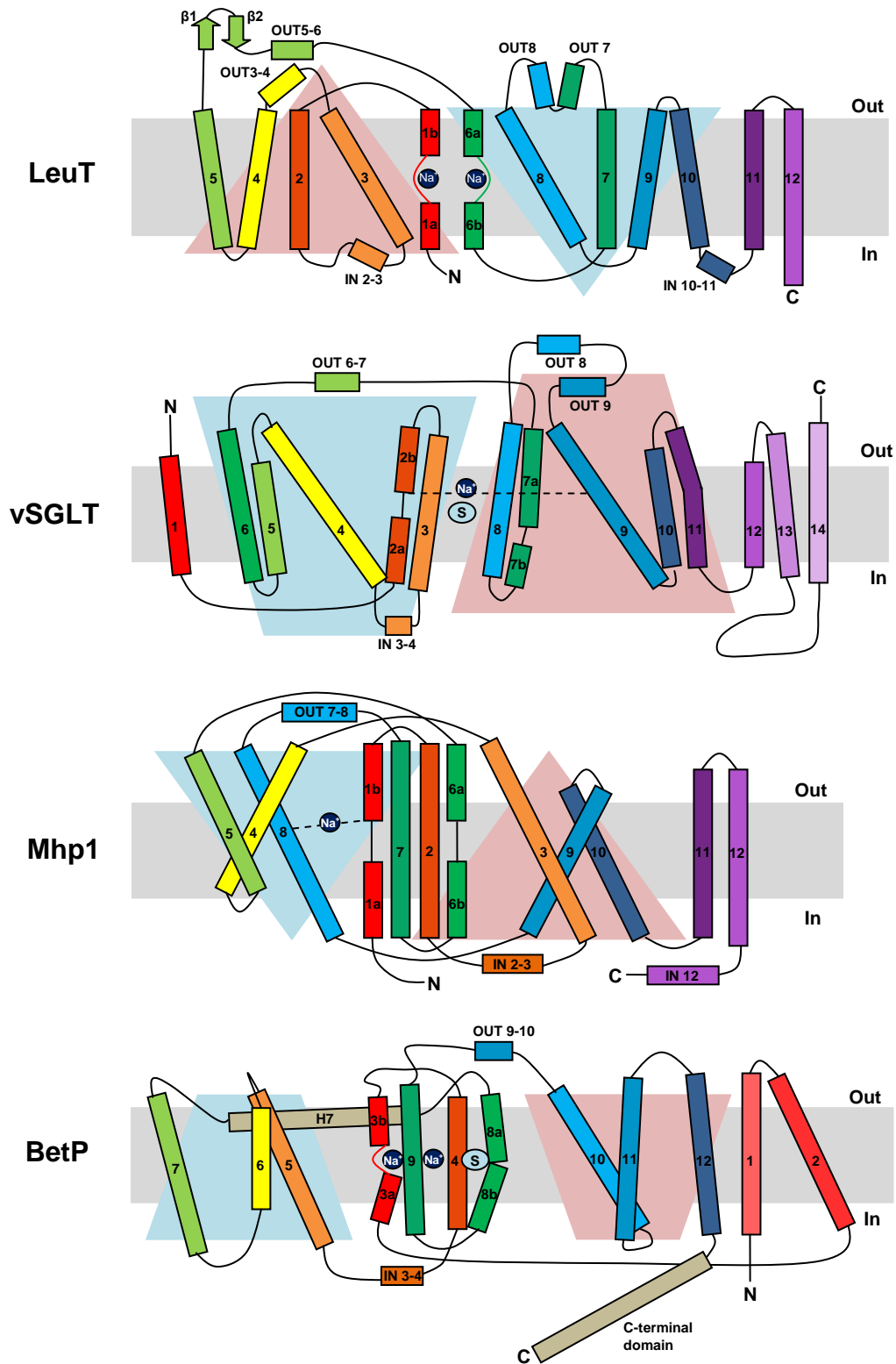


Figure 1.12 Topology of sodium-coupled transporters with inverted repeats of five helices. The positions of the substrate- and sodium-binding sites are indicated by light and dark blue circles, respectively. The phospholipid bilayer is shown in grey. The helices are coloured according to LeuT for comparison.

A distinctive feature of the two inverted repeats of five helices is that two of the helices, 1 and 6, are discontinuous. Discontinuous helices have been known to correlate with ion translocation function (Screpanti & Hunte, 2007), they create a local polar environment for the binding of ions and substrates, as in NhaA and GltPh (Yernool *et al.*, 2004; Hunte *et al.*, 2005). Interestingly, a recent structure of an amino acid transporter (ApcT) also revealed an inverted structural repeats of five helices (Shaffer *et al.*, 2009), but is not sodium-dependent, demonstrating common principles in proton- and sodium- coupled transporters.

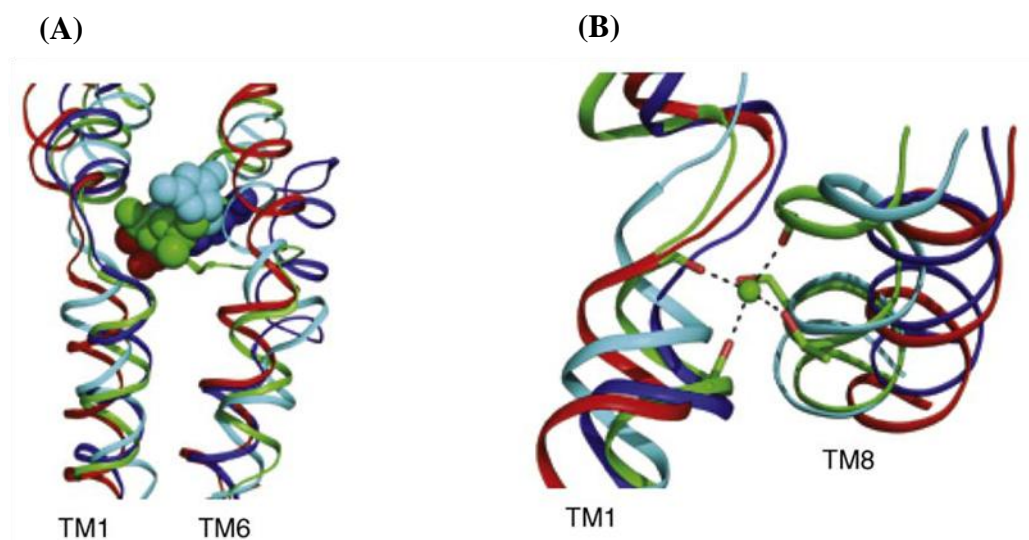


Figure 1.13 Common substrate and sodium binding sites of sodium-coupled transporters. (A) Structural alignment of helices 1 and 6, and substrate binding site for vSGLT (red), BetP (blue), LeuT (green), and Mhp1 (cyan). (B) Structural alignment of helices 1 and 8, and the conserved sodium-binding site. Coloured as for (A). Diagram taken from Abramson & Wright (2009). TM = transmembrane/helix.

1.3.6.5. Six-state model for sodium-coupled transporters with inverted repeats.

Structures of sodium-coupled symporters in their different conformations provide insights into the mechanism of transport. Based on these structures a six-state model could be explained (Figure 1.14) (Abramson & Wright, 2009). Firstly, sodium ion(s) binding causes opening of the gate allowing substrate to bind (outward-facing open conformation of Mhp1) (state 2). In Mhp1 and human SGLT (Weyand *et al.*, 2008; Hirayama *et al.*, 2007), the extracellular part of helix 10 bends towards the extracellular facing cavity forming the outward facing occluded conformation (state 3) upon substrate binding. For LeuT, this is achieved by stacking of the bulky residues F253 and Y108 on the substrate-binding site at the extracellular face (Yamashita *et al.*, 2005). State 4 represents the intermediate state for the conversion from the outward

facing to inward facing conformation. In this state, there is no extracellular or intracellular cavity, as shown by the structure of the BetP complex with both substrate and sodium bound (Ressl *et al.*, 2009).

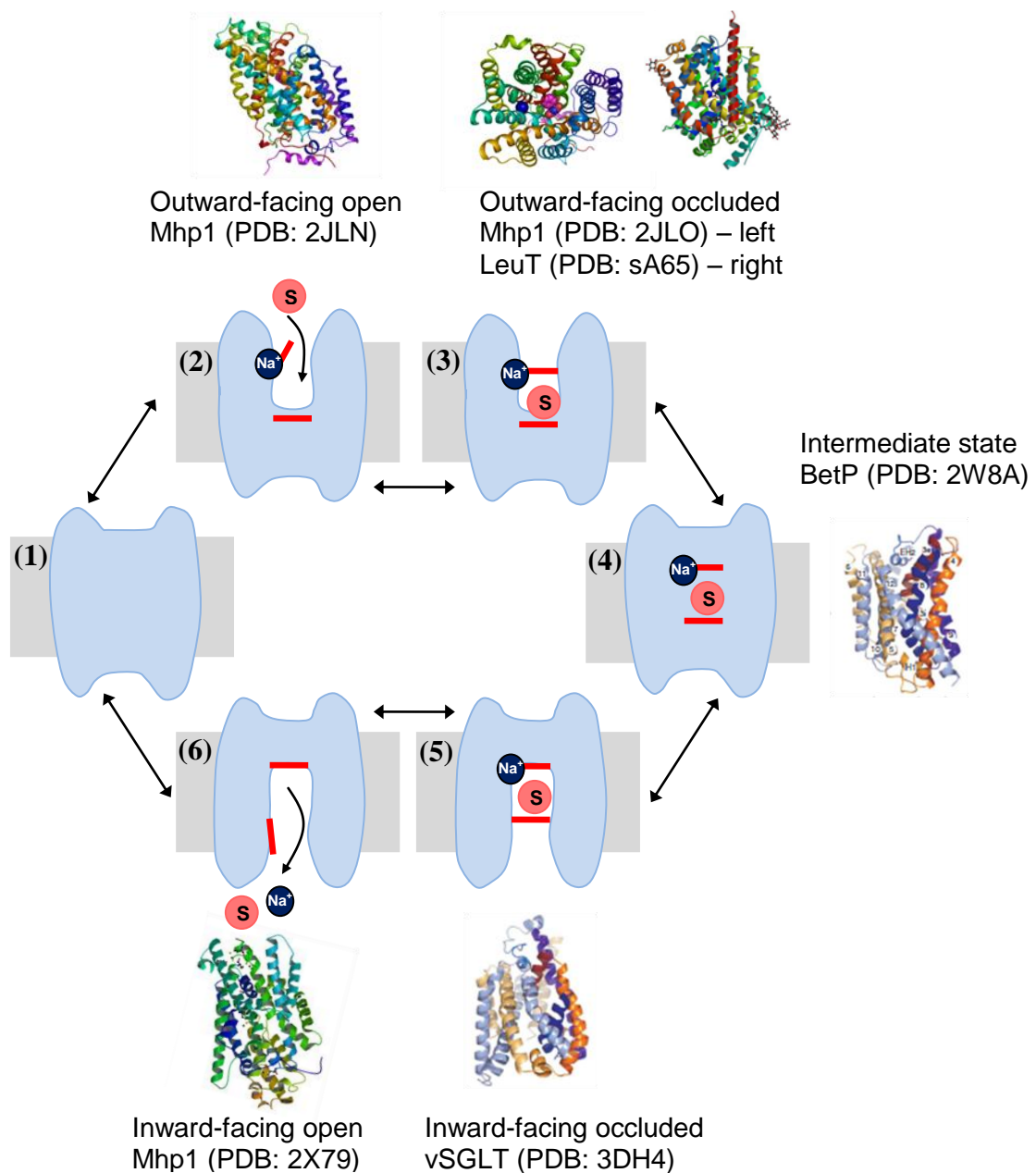


Figure 1.14 A six-state model for sodium-symporters. Diagram modified from Abramson & Wright (2009). (1) No ligand is bound. (2) One or two sodium ions bind to the external surface and external gate opens. (3) Substrate is then able to bind to its binding site and closing the external gate. (4) The external vestibule closes. (5) Opening of the internal vestibule. (6) The internal gate opens and substrate and sodium ion(s) dissociate and enter the internal face of the membrane. The internal vestibule closes to complete the cycle and returns to state (1). The crystal structures of Mhp1 correspond to state 2 and 3 (PDB code 2JLN and 2JLO), LeuT to state 3 (PDB code 2A65 and 3F3A), BetP to state 4 (PDB code 2W8A), vSGLT to state 5 (PDB code 3DH4), and Mhp1 to state 6 (PDB code 2X79).

Next, the transporter forms the inward-occluded conformation as indicated by the structure of vSGLT. In vsGLT, galactose is packed in between hydrophobic residues (Y263 on the inside; M73, Y87 and F424 on the outside), which makes the substrate inaccessible to both faces of the membrane. The intracellular exit pathways of these transporters consist of a large hydrophilic cavity formed from portions of helices 1, 2, 3, 6, 8 and 10. Very recently, the structure of an inward facing conformation of Mhp1 was revealed showing a dramatic change in the structure from its outward-occluded state to the inward structure (Shimamura *et al.*, 2010). In order to release the sodium and the substrate, Mhp1 achieves this by opening the intracellular thick gate with a rigid body rotation of helices 3, 4, 8 and 9 relative to helices 1, 2, 6 and 7 (the bundle) (Shimamura *et al.*, 2010). It is presumed the intracellular thin gate (helix 5) also opens to allow the substrates to exit toward the cytoplasm (Shimamura *et al.*, 2010).

The aims and objective of this work are detailed in Section 1.5.1.

1.4. Two-component signal transduction systems of *Enterococcus faecalis*

1.4.1. Enterococci and *Enterococcus faecalis*

Enterococci are natural inhabitants of the oral cavity, normal intestinal microflora, and female genital tract of both human and animals (Murray, 1990; Jett *et al.*, 1994), and are recognised as opportunistic pathogens. They are now acknowledged to be the second most important cause of hospital-acquired infections in the US and fourth in the UK (Schaberg *et al.*, 1991; National Nosocomial Infections Surveillance, 2004; Oncu *et al.*, 2004). The two most commonly isolated species of enterococci associated with nosocomial infections are *Enterococcus faecalis* (*E. faecalis*) and *Enterococcus faecium* (*E. faecium*), both of which are capable of producing biofilms (Mohamed & Huang, 2007). *E. faecalis* is responsible for approximately 80 – 90% of human enterococcal infections (Jett *et al.*, 1994; Jones *et al.*, 2004), whilst *E. faecium* accounts for the remainder of enterococcal infections (Jett *et al.*, 1994).

The genome of a vancomycin-resistant clinical strain of *E. faecalis* (strain V583) was sequenced (Paulson *et al.*, 2003), revealing a total of 3182 open reading frames (ORFs) with 1760 of these showing similarity to known proteins and 221 of unknown functions. Approximately 25% of the genome is made up of mobile and exogenously acquired DNA, a unique feature of this genome, which includes a number of conjugative and composite transposons, a pathogenic island of integrated plasmid genes

and phages (Giridhara Upadhyaya *et al.*, 2009). *E. faecalis*, like other enterococci, exhibits some unusual but characteristic features. It is able to grow in 6.5% NaCl and survive temperatures ranging from 10°C to 45°C, and it is tolerant to acidic and alkaline conditions, commensurate with its intestinal environment. These observations led to the question of how enterococci have evolved to permit such hardy survival. It is possible that sensing of environmental signals plays a very important role in this bacterium, and therefore that the two-component signal transduction systems confer important selective advantages for the survival of enterococci.

1.4.2. Two-component signal transduction systems

Signal transduction in bacteria and plants is often mediated by His-Asp phosphorelay systems. Two-component signal transduction systems (TCS) (Figure 1.15) enable bacteria to sense, respond, and adapt to the changes in their environment (Hoch & Silhavy, 1995; Stock *et al.*, 2000).

Bacteria possess many distinct TCS pairs, as many as 150 in some species depending on genome size (Barakat *et al.*, 2009). Each TCS consists of a sensor histidine kinase (HK) and a response regulator (RR) (Figure 1.15). Upon detection of a specific signal(s), the sensor kinase autophosphorylates on a conserved histidine residue, and subsequently transfers the phosphoryl group to a conserved aspartate residue on a cognate response regulator (Stock *et al.*, 2000). The activated response regulator can then effect changes in response to the environment stimulus, often by regulating gene expression. There are currently ~50,000 TCS proteins identified from the sequenced genomes (Barakat *et al.*, 2009) and over 200 structures of two-component proteins have been determined. TCSs regulate a wide range of cell activities, including fundamental processes such as motility and metabolism as well as virulence and development. TCSs are the most abundant signal transduction mechanism utilised by bacteria and they have been extensively studied since their discovery more than 35 years ago. It was hoped that the understanding these systems might be useful for development of antibiotics or the beneficial use of bacteria to improve agriculture or environment remediation. Indeed, it has been suggested that TCSs are excellent new targets for novel antibacterial drugs, as these systems are present in bacteria (and lower eukaryotes such as yeasts, and in higher plants) but are absent in mammals.

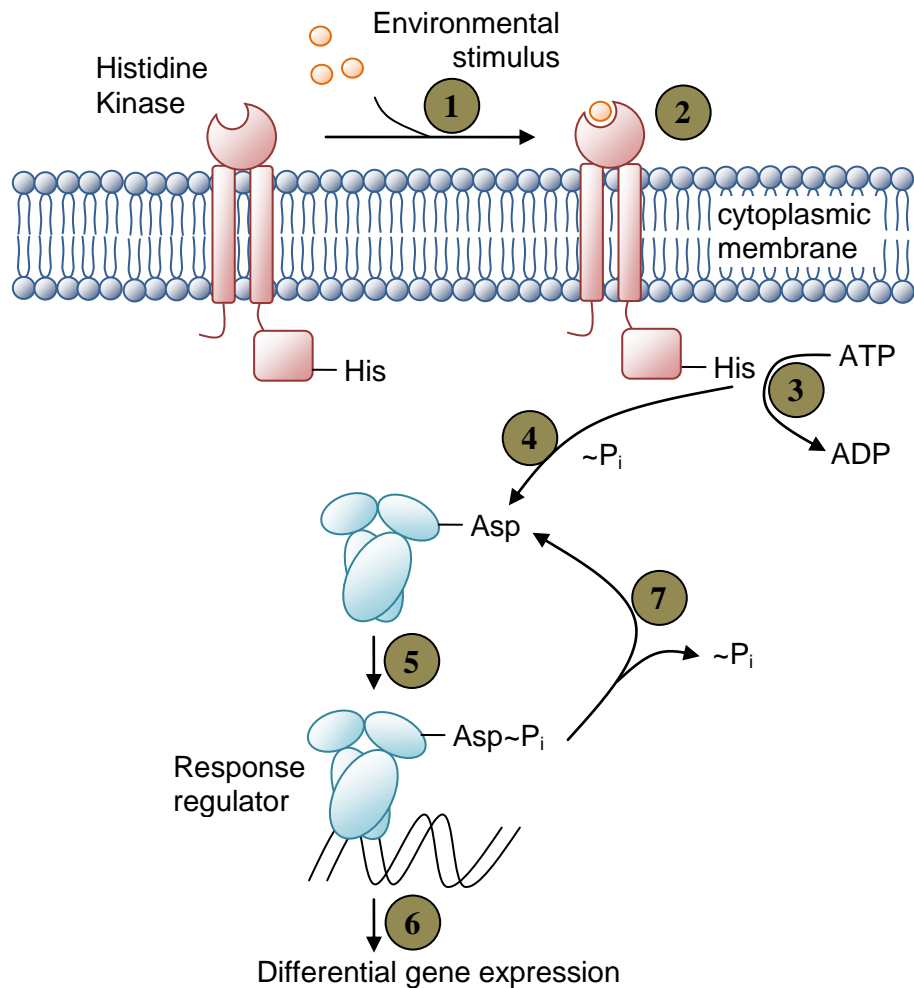


Figure 1.15 Two-component signal transduction system. The system is comprised of two proteins, a membrane-bound sensor histidine kinase and a cytoplasmic response regulator. The histidine kinase detects a specific environment stimulus (1) and undergoes a conformational change (2) resulting in the phosphorylation/activation of the kinase domain (transmitter) (3). The phosphate of the activated kinase domain is transferred to the receiver domain of its cognate response regulator, resulting in response regulator phosphorylation and activation (4). The activated response regulator mediates an appropriate cellular response (output) (5). Usually, the response regulator has a DNA-binding motif, and binding to its target promoter DNA sites leads to differential gene expression (6). The response regulator returns to a pre-stimulus state by dephosphorylation (7), either by the intrinsic phosphatase activity exhibited by the transmitter domain or by additional phosphatases.

1.4.3. Sensor histidine kinase

1.4.3.1. Protein phosphorylation in bacteria. The discovery of protein phosphorylation dates back to the middle of the 19th century when milk and egg proteins were found to covalently attached to inorganic phosphate. Lipmann and Levene (1932) detected phosphoserine in egg yolk, and over two decades later, phosphothreonine was discovered in casein, meanwhile phosphotyrosine as a mode of protein modification

was unknown until 1979 (Eckhart *et al.*, 1979). Phosphorylation of histidine was first reported by Boyer *et al.* (1962) who identified this type of phosphorylation in mitochondrial succinyl-CoA synthase. Phosphorylation was first recognised as important in the regulation of enzymatic activities by Krebs & Fischer (1956). Subsequently, numerous examples of regulation through phosphorylation were identified in eukaryotes, but lack of protein phosphorylation in prokaryotes led to the view that this type of modification was exclusive to eukaryotes. However, this view was changed upon publication in 1969 of a paper describing the isolation of a cyclic AMP-dependent kinase from *E. coli* which transferred phosphate from ATP to histones of calf thymus (Kuo & Greengard, 1969). A few other serine- and threonine-phosphoryl groups from an acidic hydrolysate of bacterial proteins were also identified (Wang & Koshland, 1978). During studies at that time, acidic conditions were frequently employed for purification procedures, for example in acidic media or TCA precipitation. Whilst, serine-, threonine-, and tyrosine phosphates are generally tolerant of acidic conditions, the phosphoramidate bond of phosphohistidines is prone to acid hydrolysis, and therefore the presence of histidine phosphates was often missed. Phosphohistidines also have a higher free energy than ATP and tend to function substoichiometrically as phosphotransfer intermediates rather than as regulatory modifications *per se*. Therefore these factors delayed the discovery of histidine protein kinases in bacteria, the first being purified and studied by Hess *et al.* (1988) and Ninfa *et al.* (1988).

1.4.3.2. Histidine kinase. There are currently ~22000 histidine kinases identified in the P2CS database (Barakat *et al.*, 2009). P2CS is a database for prokaryotic TCSs, which contains compilations of TCS genes from 755 completely sequenced genomes and 39 metagenomes. The first histidine kinases to be sequenced included EnvZ (Mizuno *et al.*, 1982), PhoR (Tommassen *et al.*, 1982), CpxA (Albin *et al.*, 1986), NtrB (Nixon *et al.*, 1986), DctB (Ronson *et al.*, 1987), VirA (Leroux *et al.*, 1987), and CheA (Stock *et al.*, 1988). The first model for two-component signal transduction was proposed by Ninfa *et al.* (1988) where NtrB functions as a kinase/phosphatase to control the cellular level of NtrC-phosphate in response to nitrogen availability, and NtrC acts to regulate transcription at nitrogen-regulated promoters. Sensor histidine kinases exhibit concentrated regions of homology and conservation termed homology boxes and these are described in Section 1.4.3.3. The most conserved region is the kinase core (Figure 1.16).

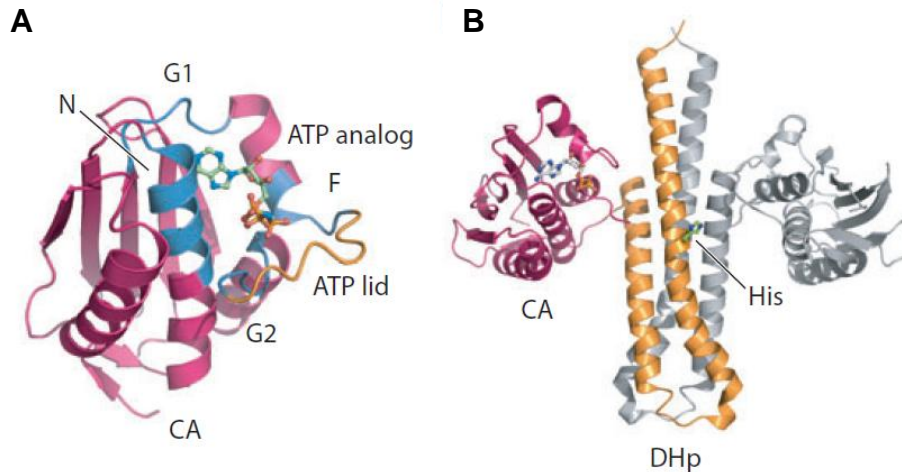


Figure 1.16 The kinase cores of histidine kinases PhoQ and HK853. (A) Structure of C-terminal catalytic and ATP binding (CA) domain of *E. coli* PhoQ (PDB: 1ID0) (Cheung *et al.*, 2008). Homology boxes important for ATP binding are shown in blue. (B) Structure of the entire kinase core of *Thermotoga maritima* HK853 (PDB: 2C2A) (Marina *et al.*, 2005). Dimeric HK853 is shown with one monomer in orange and pink, and the other in grey. An ATP analogue is bound to the CA domain which is covered by the ATP lid (orange). Diagram taken from Gao & Stock (2009).

1.4.3.3. Histidine kinase domains and conserved regions (homology boxes).

Membrane-bound histidine kinases possess several identifiable domains and conserved regions that are involved in signal perception (the sensory domain), signal transduction across the membrane (the transmembrane domain), phosphorylation and dimerisation domain for ATP binding and catalysis. Homology boxes have been identified in several of these domains, including the H, N, G1, F, and G2 boxes (Parkinson & Kofoid, 1992; Stock *et al.*, 1995; Grebe & Stock, 1999). The conserved H-box, which is located in the histidine kinase and dimerisation domains contains the site of histidine phosphorylation, whereas the N, G1, F, and G2 boxes define the nucleotide binding cleft (Figure 1.16).

1.4.4. Histidine kinases of *E. faecalis*

1.4.4.1. Membrane-bound sensor histidine kinases of *E. faecalis*. There are seventeen TCS pairs along with one orphan response regulator identified in the *E. faecalis* strain V583 (Hancock & Perego, 2002). The seventeen sensor histidine kinases of *E. faecalis* (Figure 1.17) are subgrouped according to the conservation of amino acids surrounding the phosphorylatable histidine residue in the H-box, a scheme proposed by Fabret *et al.* (1999).

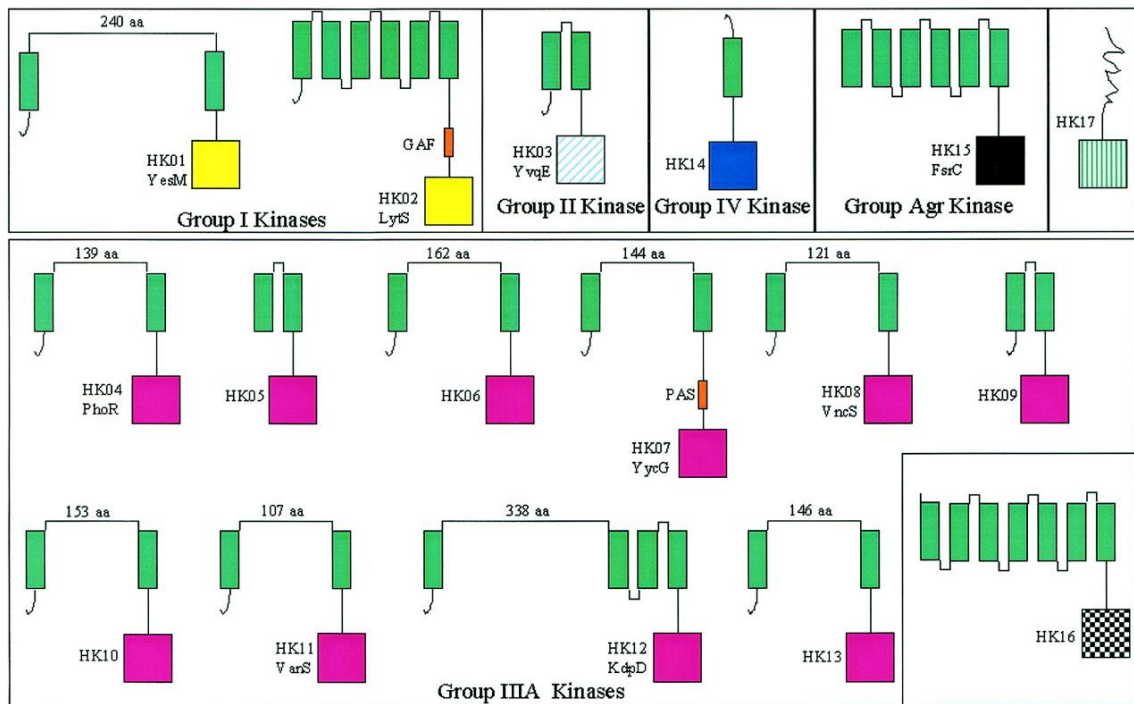


Figure 1.17 Schematic diagram of the sensor histidine kinases of *E. faecalis*. Diagram taken from Hancock & Perego (2002). Sensor histidine kinases of *E. faecalis* are grouped according to the scheme of Fabret *et al.* (1999). Transmembrane domain (green rectangles), histidine phosphorylatable domain (square), PAS or GAF domain (orange rectangle).

All the sensor histidine kinases of *E. faecalis* V583, apart from one (HK17), are predicted to be membrane-bound (Figure 1.17) (Hancock & Perego, 2002), consistent with the observation that localisation of the kinase in the membrane of the bacterial cell is a general feature of most TCSs (Kim & Forst, 2001). Two of the kinases belong to group I and one to group II. Ten kinases were assigned to group IIIA, which is the most abundant class of kinases in Gram-positive bacteria (Ferretti *et al.*, 2001; Kuroda *et al.*, 2001; Tettelin *et al.*, 2001). One kinase was classified into group IV. The remaining kinases were not classified as they do not appear to fit into the current classification scheme of Fabret *et al.* (1999); these include FsrC the quorum sensing kinase, and HK16 and HK17 of unknown function (Table 1.2).

Group ¹	Kinase Accession no.	Phosphorylatable histidine region ²	Similarity ³ (possible functions or systems it control)
I	EF2219 (HK01) Q832K7	LQSQINPHFLYNTLEYIR	YesM of <i>B. subtilis</i> (sugar utilisation and cell wall sensitivity to bacitracin)
	EF3197 (HK02) Q82Z75	LQAQVNPHEFFNAINTIS	LytS of <i>S aureus</i> (cell wall metabolism and murein hydrolase activity)
II	EF2912 (HK03) Q82ZY5	HRLAREIHDSVSQQLFAM	YvqE of <i>B. subtilis</i> ; VraS of <i>S. aureus</i> (increase antibiotic resistance such as vancomycin and bacitracin)
IIIA	EF1704 (HK04) Q834F1	DFVSNVSHLKTPTVTSLLG	PhoR of <i>B. subtilis</i> (growth of cell, heat sensitivity and virulence)
	EF3290 (HK05) Q82YZ0	ELITNVSHDIRTPLTSIIG	CroS (growth, cell morphology and virulence)
	EF1261 (HK06) Q835W1	QFMADASHERMTPLTTING	YcIK of <i>B. subtilis</i> (growth, oxidative stress, virulence, membrane function and structure)
	EF1194 (HK07) Q836C1	EFVSNVSHLRTPLTSMRS	YycG of <i>B. subtilis</i> (cell wall biosynthesis and division, membrane composition, genetic competence and virulence)
	EF1863 (HK08) Q833S3	DDFKGASHELKTPPLASLKI	VncS of <i>S. pneumonia</i> (vancomycin resistant, heat sensitivity, and genetic mobile elements)
	EF0927 (HK09) Q837B6	DYIDSWVHEIKVPLAAITL	Similar to lactobacilli protein kinases of unknown function
	EF1051 (HK10) Q836Q7	QFVEDVSHLKTPIAAVSV	LisK of <i>L. monocytogenes</i> ; CsrS of <i>S. pyrogenes</i> (virulence, acid and heat sensitivity)
	EF2298 (HK11) Q47745	YFFAAASHLKTPIAAVSV	VanS _B of <i>E. faecalis</i> (vancomycin resistant)
	EF0570 (HK12) Q838F4	NLLRAVSHDLRTPLTVISG	KdpD of <i>Clostridium actobutylicum</i> (potassium concentrations, osmotic shock, and turgor pressor)
	EF0373 (HK13) Q838R2	ELIANISHDLKTPITSIIG	WakK (yycG or Vick) of <i>S aureus</i> (autolysis, cell wall metabolism and biofilm formation)
IV	EF1209 (HK14) Q836A8	SALQSQSHEFMNKMHVIIYG	CitA of <i>E. coli</i> ; YufL of <i>B. subtilis</i> (Citrate and C4 carboxylate metabolism, malate utilisation)
Not grouped	EF1820 (HK15) Q833V6	EELAMFRHDYKNLLYSLYS	FsrC of <i>E. faecalis</i> (quorum sensing, biofilm, gelatinase and serine protease production)
	EF1335 (HK16) Q835P1	TNVLKCLKHDLKNQYLITLG	VirS of <i>C. perfringes</i> (toxin production, regulation of antibiotic uptake)
	EF1632 (HK17) Q834L6	VAIREIHHRVKNNLQSVVS	Unknown function

Table 1.2 Histidine kinases of *Enterococcus faecalis* V583. Table after Hancock & Perego (2002). ¹Grouped based on the homology of the phosphorylatable histidine region (Fabret *et al.*, 1999). ²Amino acid sequence of the phosphorylatable histidine region. ³Similarity search by FASTA3.

1.4.4.2. Predicted biological functions. Information concerning three-dimensional structures and molecular mechanisms and functions of intact membrane TCSs, including those of *E. faecalis*, is generally lacking. Concerning the *E. faecalis* V583 systems, only three systems, FsrCA and VanSR and VicKR have been functionally characterised (Evers & Courvalin, 1996; Nakayama *et al.*, 2001; Ma *et al.*, 2008). Similarity searches, comparisons with other known bacterial TCSs (Hancock & Perego, 2002), systematic inactivation of the response regulator components (Teng *et al.*, 2002; Le Brenton *et al.*, 2003; Hancock & Perego, 2004; Muller *et al.*, 2008), experiments using mouse peritonitis and peritonitis macrophage models (Teng *et al.*, 2002; Muller *et al.*, 2008) have all provided valuable insights into the potential roles of each TCS of *E. faecalis*. Some were shown to be associated with virulence or responses to antimicrobial agents and others in stress responses, but their precise molecular environmental signals have not been identified, even in the case of the functionally-characterised VanSR. Below is a summary of what is currently known about each system.

The TCS EF2219-EF2218 (HK01-RR01) function is unknown, but it is homologous to YesM of *B. subtilis* that is located next to pectin/rhamnogalacturonan utilisation genes (Poncet *et al.*, 2009). Inactivation of the response regulator RR01 resulted in increased sensitivity to cell wall active agent bacitracin (Hancock & Perego, 2004b).

EF3197-EF3196 (HK02-RR02) is highly homologous to LytSR of *S. aureus*. The sensor histidine kinase EF3197 possesses 55% identity with LytS (Brunskill & Bayles, 1996; Hancock & Perego, 2002), suggesting a possible role for this system in cell wall metabolism and lysis through intrinsic murein hydrolase activity (Rice & Bayle, 2008). However, no changes in rate of autolysis were observed in an RR02 mutant, only an altered uptake of antibiotics such as erythromycin was observed in this mutant, and possibly the system was involved in cell wall permeability functions (Hancock & Perego, 2004b). The EF3197 kinase was investigated in this study and is more fully described in Section 5.2.

EF2912-EF2911 (HK03-RR03) is homologous to YvqE of *B. subtilis* that is involved in resistances to bacitracin and nisin (Mascher *et al.*, 2003; Hansen *et al.*, 2009), and is also homologous to VraS of *S. aureus* that contributes to methicillin-resistance (Kuroda *et al.*, 2003; Boyle-Vavra *et al.*, 2006; Yin *et al.*, 2006); overexpression of response regulator VraR leads to an increased resistance to

vancomycin (Kuroda *et al.*, 2000). Interestingly, the system is highly conserved in Gram-positive bacteria including *S. aureus*, *S. pyogenes*, and *S. pneumonia* (identity of 60 – 70%). Mutations in RR03 resemble that of RR01, exhibiting increased sensitivity to bacitracin (Hancock & Perego, 2004b).

EF1704-EF1703 (HK04-RR04) is similar to the PhoPR system of *B. subtilis* (Hancock & Perego, 2002; Teng *et al.*, 2002; Howell *et al.*, 2003). HK04 is 30% identical to PhoR and RR04 is 54% identical to PhoP (Hulett, 1996). Disruption of the response regulator gene RR04 resulted in sensitivity towards heat in strain JH2-2 (Le Breton *et al.*, 2003) and in a growth defect in strain V583 (Hancock & Perego, 2004b). The system was shown to be induced under phosphate deprivation conditions, and required for full survival in macrophages, confirming its role in virulence (Muller *et al.*, 2008).

EF3290-EF3289 (HK05-RR05) is named the CroSR system and is required for intrinsic β -lactam resistance (Hancock & Perego, 2004b; Comenge *et al.*, 2003). A RR05 mutant caused growth defects and alteration in cell morphology in strain JH2-2 (Le Breton *et al.*, 2003). The system is also involved in virulence since the mutation also impaired survival in macrophages (Muller *et al.*, 2008).

EF1261-EF1260 (HK06-RR06) is also associated with virulence and is most similar to YclK of *B. subtilis* (Kobayashi *et al.*, 2001). A RR06 mutant was shown to exhibit impaired survival and susceptibility to oxidative stress (Muller *et al.*, 2008). RR06 mutant was found to be more susceptible than wild type to heat and SDS, suggesting the system could be linked with membrane functions and structure (Hancock & Perego, 2004b).

EF1194-EF1193 (HK07-RR07), also named VicKR is reported an essential system for *E. faecalis*. The system is highly similar to essential YycFG of *B. subtilis*, which is also found highly conserved in Gram-positive bacteria (Bolotin *et al.*, 2000; Ferretti *et al.*, 2001; Glaser *et al.*, 2001; Kunst *et al.*, 1997; Kuroda *et al.*, 2001; Nolling *et al.*, 2001; Tettlelin *et al.*, 2001). The EF1194 kinase is 70% similar and 45% identical to the YycG kinase (Yamamoto *et al.*, 2001). Studies have shown that the system is involved in regulation of genes involved in cell wall biosynthesis or cell membrane composition (Howell *et al.*, 2003; Ng *et al.*, 2003; Ng *et al.*, 2004; Dubrac & Msadek, 2004; Mohedano *et al.*, 2005; Ng *et al.*, 2005), cell division (*ftsAZ* operon) (Fukuchi *et al.* 2000), virulence and genetic competence (Ng *et al.*, 2005; Wagner *et al.*, 2002; Senadheera *et al.*, 2005) and biofilm formation (Senadheera *et al.*, 2005).

Hancock & Perego (2004) provided suggestive evidence for the essential nature of this system, since knockouts in the response regulator gene *vicR* appeared to result in a lethal phenotype.

EF1863-EF1864 (HK08-RR08) is most similar (56% similarity and 35% identity) to the VncSR system of *S. pneumoniae* implicated in vancomycin tolerance (Novak *et al.*, 1999), and linked to mobile genetic elements (Hancock & Perego, 2004b). In common with the EF1704-EF1703 and EF3329 mutations, disruption of RR08 led to heat sensitivity (Le Breton *et al.*, 2003).

Little is known about the EF0927-EF0926 (HK09-RR09) system. Similarity searches showed that the HK09 is most similar to some lactobacilli proteins, but of unknown function.

EF1051-EF1050 (HK10-RR10) is also named EtaSR, and has been shown to contribute to virulence in a mouse peritonitis model and peritoneal macrophages (Teng *et al.*, 2002; Muller *et al.*, 2008). Mutants show increased acid sensitivity (Teng *et al.*, 2002) and heat resistance correlated with DnaK and GroEL levels (Le Breton *et al.*, 2003). This system is most similar to the LisKR and CsrSR systems of other Gram-positive bacteria. The EF1051 kinase was investigated in this study and is described in Section 5.3.

EF2298-EF2299 (HK11-RR11) also called VanS_BR_B regulates VanB-type vancomycin resistance. Inactivation of the genes leads to significantly more sensitivity towards vancomycin (Evers & Courvalin, 1996; Hancock & Perego, 2004b).

EF0570-EF0571 (HK12-RR12) shares similarity with KdpDE from several Gram-positive and Gram-negative organisms (Treuner-Lange *et al.*, 1997). The genes are often located on a pathogenicity island, though not always present in all *E. faecalis* strains and therefore suggesting that it may not be essential for a commensal life style (Teng *et al.*, 2002; Hancock & Perego, 2004b; Shankar *et al.*, 2002).

EF0373-EF0372 (HK13-RR13) is homologous to Walk (synonyms: VicK, yycG) of *S. aureus*, having 34.2% identity and 73.8% similarity. Walk is shown to regulate genes involved in autolysis, biofilm formation and cell wall metabolism (Wagner *et al.*, 2002; Clausen *et al.*, 2003; Jansen *et al.*, 2007). Expression of EF0373-EF0372 was only detectable at high temperature (55°C), suggesting a possible role in stress response (Le Breton *et al.*, 2003).

EF1209-EF1210 (HK14-RR14) is most similar to CitBA and DcuSR that are involved in regulation of citrate and/or C4 dicarboxylate transport (Scheu *et al.*, 2008;

Yamamoto *et al.*, 2009). EF1209 is 30% identical and 65% similar to DcuS of *E. coli*. Mutant RR14 lacks the response regulator component which resulted in a growth defect (Hancock & Perego, 2004b) and increased sensitivity at high temperature (50°C), suggesting that the system is involved in stress responses (Le Breton *et al.*, 2003). The system is also similar to YufLM of *B. subtilis* (the kinase being 34% identical and 65% similar to YufL) that is involved in regulation of malate utilisation (Tanaka *et al.*, 2003; Doan *et al.*, 2003).

EF1820-EF1821 (HK15-RR15) is the Fsr quorum sensing system of *E. faecalis* (Qin *et al.*, 2000; Nakayama *et al.*, 2001; Qin *et al.*, 2001; Nakayama *et al.*, 2006). In a systemic inactivation of all two-component systems present in *E. faecalis*, it was demonstrated that the *fsr* system is the only system affected in biofilm formation (Hancock & Perego, 2004b). The *fsr* regulatory locus is comprised of four genes; designated *fsrA*, *fsrB*, *fsrC* and *fsrD*, and its genes organisation are highly similar to the *agr* system in *Staphylococcus* and other Gram-positive bacteria (Autret *et al.*, 2003; Nakayama *et al.*, 2006). The Fsr kinase (EF1820) was investigated in this study and is described in Section 5.4.

EF1335-EF1336 (HK16-RR16) is homologous to the VirRS system of *Clostridium perfringens* (28% identity and 61% similarity) that is involved in toxin production (Cheung *et al.*, 2009). A knockout of RR16 in *E. faecalis* V583 displayed faster doubling times during growth in the presence of subinhibitory concentrations of erythromycin, suggesting that the system could affect uptake of this class of drug (Hancock & Perego, 2004b).

EF1632-EF1633 (HK17-RR17) is the only system in *E. faecalis* with a cytoplasmic sensor histidine kinase (Hancock & Perego, 2002). The system showed no similarity to known TCSs and the function is unknown.

1.4.5. Mechanism of stimulus perception

Sensor histidine kinases of the TCSs play important roles in bacterial signal transduction by sensing external signals and directly affecting the phosphorylation state of the cytoplasmic kinase domain through stimulating its kinase or phosphatase activity. A classic sensor histidine kinase is homodimeric and typically has two transmembrane segments with an extracellular loop through which signal is sensed. With the increasing number of kinase domain structures it was shown that depending on the subcellular localisation of the specific system and properties of the detected signal, input domains

could be extracellular, membrane-embedded, or intracellular (Figure 1.18) (Cheung & Hendrickson, 2010). Current structures of some portions of the histidine kinases provide insights into the architecture relating to the overall mechanisms of the kinases (Figure 1.19). The precise mechanisms by which signals are perceived by the kinase and transduced across the membrane remains elusive, particularly in the absence of intact proteins possessing their transmembrane and sensory domains.

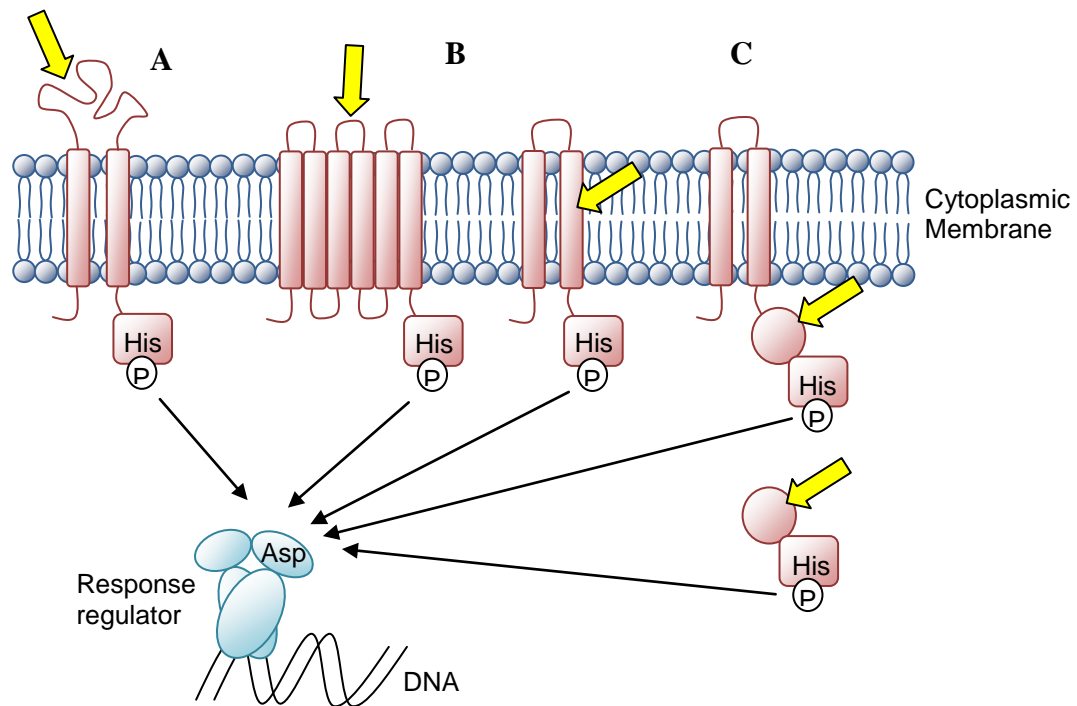


Figure 1.18 Mechanisms of stimulus perception by sensor kinases. (A) Periplasmic-sensing kinase. (B) Transmembrane region-sensing kinases. Stimulus perception either occurs with the membrane-spanning helices alone or with combination of the transmembrane regions and short cytoplasmic loops. (C) Cytoplasmic-sensing kinase (could be soluble or membrane bound). The stimulus is represented by a yellow arrow.

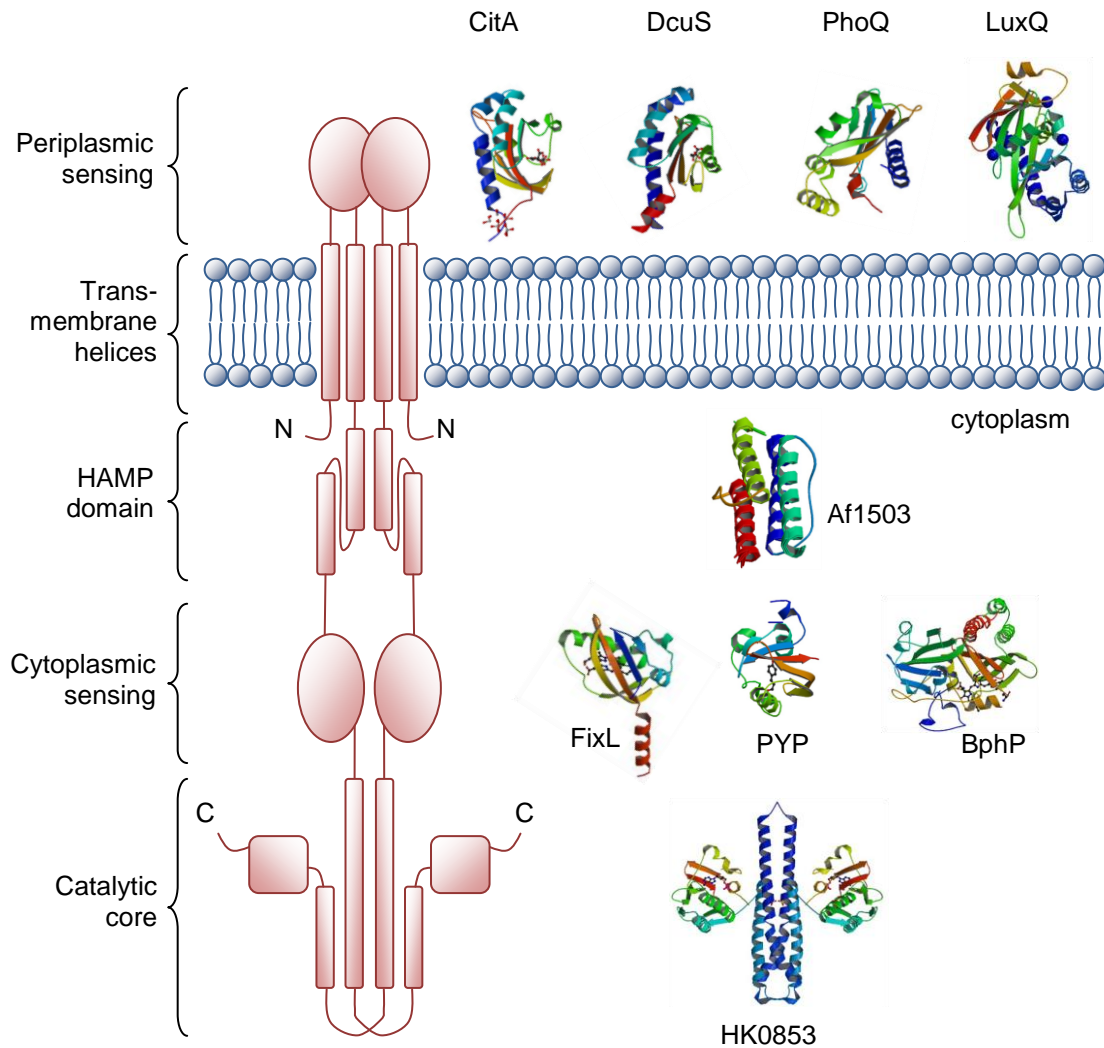


Figure 1.19 Structures of histidine kinase domains. Diagram after Szurmant *et al.* (2007). Sensor histidine kinase is a large modular protein comprised of the conserved C-terminal catalytic core and several optional N-terminal elements involved in signal sensing or propagation of signal to the catalytic core. Example structures for periplasmic domains include: CitA of *Klebsiella pneumoniae* (PDB: 1P0Z) (Reinelt *et al.*, 2003), DcuS of *E. coli* (PDB: 3BY8) (Cheung & Hendrickson, 2008), PhoQ of *E. coli* (PDB: 3BQA) (Cheung *et al.*, 2008), LuxQ of *Vibrio harveyi* (PDB: 2HJE) (Neiditch *et al.*, 2006). An example of HAMP (Histidine kinases, adenylyl cyclises, methyl-accepting chemotaxis proteins, and phosphatases) domain: Af1503 of *Archaeoglobus fulgidus* (PDB: 2ASX) (Hulko *et al.*, 2006). Example structures of cytoplasmic sensing domain: FixL of *Bradyrhizobium japonicum* (PDB: 1DRM) (Gong *et al.*, 1998), PYP of *Rhodospirillum centenum* (PDB: 1MZU) (Rajagopal & Moffat, 2003), BphP of *Deinococcus radiodurans* (PDB: 1ZTU) (Wagner *et al.*, 2005). Example structures of catalytic domain: HK0853 of *Thermotoga maritima* (PDB: 2C2A) (Marina *et al.*, 2005).

1.4.5.1. Periplasmic-sensing. The three classes of periplasmic sensor domain include: mixed alpha-beta folds, all-alpha folds, and periplasmic binding protein-like folds. Recent sequence analysis revealed a putative additional class of all-beta sensor domains,

though currently there is no structural evidence for this. The first three periplasmic sensing domain structures solved include PhoQ (Cheung *et al.*, 2008; Cho *et al.*, 2006), DcuS (Cheung & Hendrickson, 2008; Pappalardo *et al.*, 2003) and CitA (Reinelt *et al.*, 2003; Sevvana *et al.*, 2008) (Figure 1.19), all with alpha-beta folds, and were termed PDC (PhoQ-DcuS-CitA) domains. These periplasmic domains sense divalent ions, C₄-dicarboxylates and citrate respectively. PDC sensors are the most prevalent, distinguishable by a long N-terminal helix and often a short C-terminal helix, and a central five-stranded anti-parallel β -sheet scaffold. The central β -sheet scaffold has the same topology as Per-Arnt-Sim (PAS) domains (Moglich *et al.*, 2009), but also contains other distinctive structural features (Cheung *et al.*, 2008). The most recent structures of PDC sensor domain include the oligosaccharide sensor AbfS (Emami *et al.*, 2009) and the PhoR periplasmic domain (PDB code: 3cwf). Often observed in extracellular sensor domains are domain insertions in which a second membrane-distal PDC domain is inserted between the first and second helices of the membrane-proximal PDC domain, forming a double-PDC domain (Cheung & Hendrickson, 2010). Double-PDC domains were first reported in quorum sensor LuxQ (Neiditch *et al.*, 2005; Neiditch *et al.*, 2006) that forms a complex with the periplasmic binding protein LuxP in order to sense its autoinducer. Comparable structures were also found in a C₄-dicarboxylate sensor (DctB) (Cheung & Hendrickson, 2008; Zhou *et al.*, 2008b) that has a ligand binding pocket in the membrane-distal domain (Figure 1.20), as well as recent structures including: sensors HK1_s-Z2 (PDB code: 3LI9), HK1_s-Z3 (PDB code: 3LIB), HK1_s-Z6 (PDB code: 3LIC), HK1_s-Z8 (PDB code: 3LID), and HK1_s-Z16 (PDB code: 3LIF (Zhang and Hendrickson, to be published)).

All-alpha folds periplasmic sensor domains are represented by NarX (Cheung & Hendrickson, 2009), TorS (Moore & Hendrickson, 2009) (Figure 1.20), and a structure in PDB (3KKB) (Tan *et al.*, to be published). A membrane-distal left-handed four-helix bundle is stacked against a membrane-proximal right-handed four-helix bundle (Figure 1.20). NarX senses nitrite and nitrate through direct binding, whereas TorS forms a complex with the accessory periplasmic binding protein TorT to bind trimethylamine-N-oxide (Baraquet *et al.*, 2006).

The third class of periplasmic sensor domain has a similar fold to periplasmic binding proteins, and is represented by HK29_s (Cheung *et al.*, 2009b) where the signal for this protein is not yet determined.

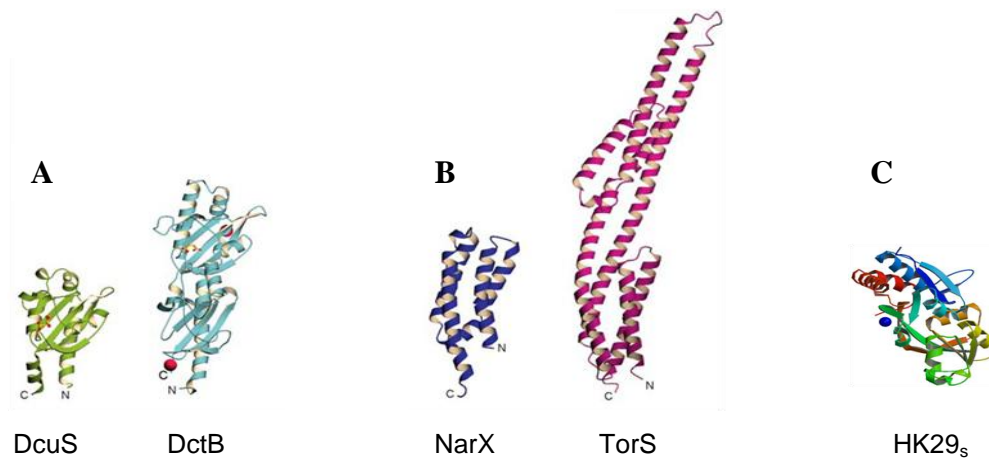


Figure 1.20 Periplasmic sensing domains. Structures are shown for: (A) PDC sensors DcuS and DctB; (B) all-alpha periplasmic sensor NarX and TorS; and (C) sensor with similar fold to periplasmic binding proteins HK29_s (PDB code: 3H7M). Diagram modified from Cheung & Hendrickson (2010).

1.4.5.2. Transmembrane-sensing. Sensor histidine kinases that possess many transmembrane helices often lack periplasmic sensing domains. In these cases the sensing domain usually lies within the transmembrane region. There are currently no structures of transmembrane-sensing domains of sensor histidine kinase, but biochemical studies have shown that the membrane regions for some histidine kinases are important for stimulus sensing. Examples of transmembrane sensing kinases include the DesK kinase that responds to temperature changes (Albanesi *et al.*, 2009; Martin *et al.*, 2009), SenS redox sensing which requires binding of secreted octameric haem-binding protein HbpS to its N-terminal transmembrane region (Bogel *et al.*, 2009), Etr1, a plant histidine kinase that detects ethylene within the hydrophobic N-terminal transmembrane region (Voet-van Vormizeele & Groth, 2008), RegB and ArcB that sense redox potential through the redox state of membrane-located quinines (Geogellis *et al.*, 2001; Swem *et al.*, 2003; Mapica *et al.*, 2004; Bekker *et al.*, 2010), PrrB that senses redox potential through interactions with the cbb3-type cytochrome c oxidase complex of the respiratory electron transport chain (Oh & Kaplan, 2000), and AgrC quorum sensor that detects its autoinducing peptide through two short extracellular loops in proximity to the transmembrane region (Geisinger *et al.*, 2008; Jensen *et al.*, 2008). As there are currently no structures of transmembrane sensing domains of histidine kinases, the structure of a phototaxis sensory rhodopsin II-transducer complex (HtrII-SrII) (Gordeilij *et al.*, 2002), which forms a four-helix bundle in the membrane, provides some insights into the helical arrangement of histidine kinases in the membrane.

1.4.5.3. HAMP domains. Many sensory proteins, including some histidine kinases, have a domain consisted of ~50 residues located just inside the membrane and within the cytoplasm, called the HAMP (Histidine kinases, adenylyl cyclises, methyl-accepting chemotaxis proteins, and phosphatases) domain. Hence, these are found in histidine kinases, adenylyl cyclases, methyl-accepting chemotaxis proteins, and phosphatases (Aravind & Ponting, 1999). HAMP domains are thought to link the transmembrane domain of the kinase to its catalytic domain, important for transducing signals from the membrane to the kinase core (Hulko *et al.*, 2006). An example structure of HAMP domain is shown by Af1503 in Figure 1.19.

1.4.5.4. Cytoplasmic-sensing. Current data indicate that many cytoplasmic sensing domains of sensor kinases have a PAS-fold, distinctive from that of PDC-fold in extracellular sensing domains. A putative PAS domain has been identified in EF1194 VicK sensor of *E. faecalis* (Hancock & Perego, 2002; Ma *et al.*, 2008). PAS domains have a five-stranded anti-parallel β -sheet core flanked by α -helices, and it may bind with a cofactor and involve in protein-protein interaction. The first two structures of histidine kinase cytoplasmic sensing domain include FixL of *Bradyrhizobium japonicum* (Figure 1.19) and *Rhizobium meliloti* (Figure 1.21) (Gong *et al.*, 1998; Gong *et al.*, 2000; Miyatake *et al.*, 2000; Hao *et al.*, 2002; Dunham *et al.*, 2003; Key & Moffat, 2005; Gilles-Gonzalez *et al.*, 2006; Key *et al.*, 2007; Moore & Hendrickson, 2009), which sense oxygen through haem. Other examples of cytoplasmic sensing in histidine kinases include NreB that detects oxygen through an iron-sulfur ($[4Fe-4S]^{2+}$) group (Mullner *et al.*, 2008), the LovK photosensor which forms a flavin adduct upon absorption of blue light (Purcell *et al.*, 2007), the MmoS redox-sensor (Figure 1.21) that contains two tandem PAS domains (PAS-A and PAS-B) with a FAD cofactor bound to an N-terminal PAS domain (Ukaegbu & Rosenzweig, 2009).

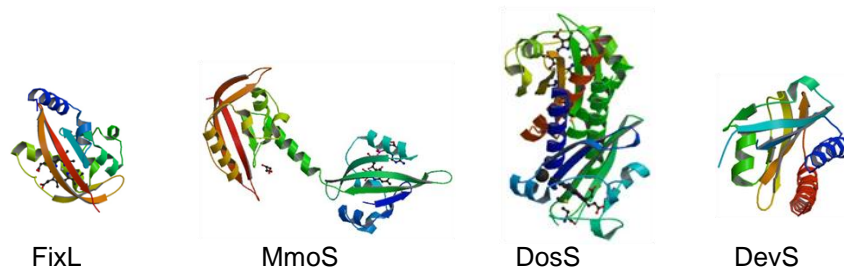


Figure 1.21 Cytoplasmic sensing domains. Structures from left to right are: PAS domain of FixL (PDB code: 1D06); PAS domains of MmoS (PDB code: 3EWK); GAF domain of DosS (PDB code: 2W3D); GAF domain of DevS (PDB code: 2VKS).

Some cytoplasmic sensor domains have a GAF fold (Ho *et al.*, 2000), which consists of a six-stranded anti-parallel β -sheet core related to the PAS domain. The first reported structures of GAF domains were those of histidine kinases DosS (Cho *et al.*, 2009) and DosT (Podust *et al.*, 2008) of *Mycobacterium tuberculosis*. Both of these sensors are haem-containing redox sensors that detect changes in the redox state of bound iron, or in binding of oxygen, respectively. Another example of haem-bound redox sensor include the DevS (Lee *et al.*, 2008) of *Mycobacterium smegmatis*, which also has a GAF domain (GAF-A) that is controlled through binding of oxygen. Additionally, two other structures of GAF domains from sensor histidine kinases were deposited in the PDB (PDB codes: 3HCY and 3CIT) but of unknown function.

Apart from PAS and GAF domains, there is a third family of cytoplasmic sensor domains found in phytochromes, soluble histidine kinase photoreceptors of bacteria, fungi, and plants. One example is the *Pseudomonas aeruginosa* bacteriophytochrome photosensory core domain (PaBphP-PCD) (Figure 1.21); which has a tripartite PCD with a PAS, GAF, and phytochrome (PHY) domain arranged along an extended α -helix (Yang *et al.*, 2008; Yang *et al.*, 2009). The PHY domain is very similar to the PAS and GAF domains, having a five-stranded anti-parallel β -sheet scaffold, and the PAS and GAF domain together forms the chromophore binding domain (CBD).

The aims and objectives of this work is detailed in Section 1.5.2.

1.5. Aims and objectives of this work

The determination of membrane protein structures using crystallographic methods has been hampered by several major bottlenecks including protein production, purification and crystallisation. The ultimate aim is to facilitate determination of the structures of membrane proteins, not just of isolated domains, but of the full-length protein, which will allow a complete understanding of structure relating to function.

1.5.1. The NCS1 family of transporters

1.5.1.1. Importance of the NCS1 proteins. Obtaining crystals of membrane transport proteins has been difficult due to their low expression level, hydrophobicity and them being dynamic, which makes them hard to extract from the membrane and to crystallise.

Three structures of Mhp1, each in a different state were elucidated providing insights into the mechanism of alternating access (Weyand *et al.*, 2008; Shimamura *et al.*, 2010). NCS1 transporters are particularly interesting to study as they are

homologous to Mhp1, the structures which are reminiscent of LeuT and vSGLT, the bacterial homologues of the human neurotransmitter and sugar transporters respectively (Yamashita *et al.*, 2005; Faham *et al.*, 2008; Weyand *et al.*, 2008), and hence these proteins are structurally-related. Elucidating the mechanism of substrate transport by these proteins will therefore be of wide significance.

NCS1 transporters also play important roles in the salvage pathways of bacteria, fungi and plants; in the capture of nutrients such as nucleobases, hydantoins, nucleosides and other related compounds from the environment. In many organisms, the transport of nucleobases into the cell serves as nitrogen sources. At least one protein, Mhp1, has commercial significance; others could putatively have commercial significance too.

1.5.1.2. Aims and research objectives. At the beginning of this research, Mhp1 was undergoing crystallisation in a collaborative project between two principal investigators Henderson and Iwata, and no structures of the NCS1 proteins, including Mhp1, were solved. The aim of my work was to clone and overexpress many NCS1 proteins for characterisation and crystallisation studies in order to elucidate their structure-function relationships. However, during the course of this research three structures of Mhp1 were elucidated and this further highlights the importance of the NCS1 transporters.

The ultimate aim of this research was to define the structure-activity relationships and establish the molecular mechanism of the NCS1 family of bacterial membrane transporters.

Research objectives

- 1) Selecting protein targets – this will involve a search for Mhp1 homologues (the targets), followed by *in silico* analysis of the targets including identifying the conserved residues and topology predictions.
- 2) Cloning and heterologous expression of bacterial NCS1 transporters (Mhp1 homologues) – the genes encoding NCS1 transporters will be amplified from the organism of interest by PCR, followed by insertion into plasmid pTTQ18-His₆ which showed successful amplification of many membrane proteins (Ward *et al.*, 2000) as well as introducing a hexahistidine tag. The resultant recombinant construct will be transformed into an *E. coli* host for expression trials.

- 3) Investigation of substrates transported by the NCS1 transporters – intact cells expressing the NCS1 transporter will be screened for the transported substrate. The substrate will be in the radiolabelled form for transport assays.
- 4) Investigation of transport specificity and kinetics – once substrate(s) for the target protein are identified, intact cells expressing the NCS1 transporter will be tested for transport kinetics and substrate specificity.
- 5) Purification of the NCS1 transporters and testing their integrity – solubilisation will be performed to extract the target protein from membranes expressing it, followed by purification using Ni-NTA chromatography. The purified protein will be subjected to circular dichroism to test its secondary structure integrity.

1.5.2. The sensor histidine kinases of *Enterococcus faecalis*

1.5.2.1. Importance of sensor histidine kinases of *E. faecalis*. Currently, one of the most serious concerns encountered in the clinic is the increasing proportion of enterococci that are resistant to the last-line antibiotic vancomycin. *E. faecalis* V583 is of course is one of those strains. Evidence has shown that bacterial resistance to antibiotics can be regulated by two-component systems (Evers & Courvalin, 1996; Matsushia & Janda, 2002). Since *E. faecalis* is an important pathogenic organism that causes up to 20% of hospital-acquired infections and since it exhibits survival under harsh environmental conditions and exhibits increasing resistance to current therapies including vancomycin, this study included studies of the membrane histidine kinases of the bacterium. The study was timely, given the recent advances in methods for expression and purifying intact sensory kinases that possess their sensory domains (Potter *et al.*, 2002; Potter *et al.*, 2006). The aims included production of intact sensor kinases for signal identification, and initiation of structural studies for intact kinases using crystallisation and other methods.

1.5.2.2. Aims and research objectives. The TCSs work is a collaborative project between two principal investigators Henderson and Phillips-Jones. The ultimate aim is to define the structure-activity relationships and establish the molecular mechanism of the sensor histidine kinases of *E. faecalis*.

Research objectives

- 1) *In silico* analysis of the membrane-bound sensor histidine kinases of *E. faecalis*
– to identify conserved residues and to undertake topology predictions.

- 2) Cloning and heterologous expression of the genes encoding the sensor histidine kinases of *E. faecalis* – The sensor histidine kinase genes will be amplified from *E. faecalis* by PCR, followed by cloning into plasmid pTTQ18-His₆. The resultant plasmids will be transformed into expression host for expression trials.
- 3) Purification of sensor histidine kinases and confirmation of their intact state – the sensor kinases will undergo solubilisation trials followed by purification using Ni-NTA chromatography. The purified proteins will be confirmed as intact by Western blotting together with N-terminal sequencing and/or mass spectrometry.
- 4) Comparison of activities of sensor histidine kinases when in the membrane-bound and purified forms – the sensor histidine kinases will be subjected to autophosphorylation assays using radiolabelled ATP.
- 5) Identification of ‘signalling’ ligands and/or inhibitors that modulate the activities of the histidine kinases of *E. faecalis*.
- 6) To initiate structural studies in order to increase understanding of how signals are transduced across the membrane by intact membrane histidine kinases. This will include crystallisation trials that may ultimately lead to the elucidation of a three-dimensional structure at atomic resolution, and the use of fluorescence and circular dichroism experiments in the presence and absence of signalling ligands.

1.6. Outline of research

Chapter 1: Introduction – This chapter highlights the importance of bacterial cell membrane and provides a literature review on two families of membrane proteins: the NCS1 transporters; and the sensor histidine kinases of *E. faecalis*. The aims and objectives of the research are specified at the end of the chapter.

Chapter 2: Methods – All methods used to carry out the research are detailed in this chapter; including recombinant DNA technologies, protein preparation and manipulation, and biochemical and biophysical techniques.

Chapter 3: Overexpression, purification and characterisation of the NCS1 family of transport proteins – This chapter describes the investigation of NCS1 transporters, outlines how the genes of interest are isolated from the bacterium and expressed in *E. coli* cells allowing their extraction from the membrane (Section 3.1). The substrates for some transporters were also explored and identified (Section 3.1). Three NCS1

transporters for which the substrates were identified were further investigated and are described in Section 3.2 to 3.4.

Chapter 4: Purification and substrate binding studies of *Microbacterium liquefaciens* hydantoin permease (Mhp1) – Binding of substrate and ligand to Mhp1 were investigated and described in this chapter.

Chapter 5: Overexpression and purification of the intact membrane histidine kinases of *Enterococcus faecalis* V583 for activity, structural and ligand binding studies – This chapter outlines the expression, purification and activities of sensor histidine kinases of *E. faecalis* (Section 5.1). Further sections investigate the sensor kinases EF3197, EF1051 and FsrC in detail (Section 5.2 to 5.4).

Chapter 6: Conclusions and future perspectives – A detailed conclusion for the research are discussed in this chapter, highlighting the impact of the research, followed by concluding remarks.

Chapter 2

Material and Methods

2.1. Sources of materials

2.1.1. Antibiotics

Carbenicillin (Melford laboratories, U.K.); siamycin I (synthesised by Kenzo Nishiguchi, Kyushu University, Japan); tetracycline (Sigma-Aldrich, U.K.).

2.1.2. Bacterial growth media

CaCl₂, glycerol, KH₂PO₄, MES, MgSO₄, NaCl, Na₂HPO₄, NH₄Cl, (Sigma-Aldrich, U.K.); casamino acids (Q.BIOgene, U.K.); isopropyl β-D-1-thiogalactopyranoside (IPTG) (Melford laboratories, U.K.); microagar, tryptone and yeast extract (Oxoid Ltd., U.K.).

2.1.3. Filters, membranes and chromatography paper

0.45 μm cellulose nitrate/acetate membrane filter, type HA (Millipore Ltd., U.K.); 0.22 μm nitrocellulose membrane filters, type GTSF (Millipore Ltd., U.K.); 0.1, 0.2 and 0.4 μm Nucleopore[®] polycarbonate (Whatman International Ltd. U.K.).

2.1.4. Protein assay, membrane preparation, purification, phosphorylation assay and fluorimetry

Acetic acid, bovine serum albumin (BSA), chicken lysozyme, choline Cl, dimethyl sulfoxide (DMSO), dithiothreitol (DTT), ethylenediaminetetraacetate (EDTA), HgCl₂, imidazole, KCl, mercaptoethanol, MgCl₂, MgSO₄, NaCl, NaOH, naphthol blue black, sodium dodecyl sulphate (SDS), sucrose, trichloroacetate (TCA), tris base ultrapure (TRIS) (Sigma Chemicals Co., U.K.); detergents list – see Section 2.5.2 (Melford laboratories, U.K. and Sigma-Aldrich); ethanol, methanol, Na₂HPO₄, NaH₂PO₄ (Fisher Scientific UK Ltd., U.K.); nickel-nitriloacetic acid (Ni-NTA) resin (QIAGEN Ltd., U.K.); gelatinase biosynthesis-activating pheromone (GBAP).

2.1.5. Protein standards

High-range rainbow coloured protein molecular weight marker (GE Healthcare Ltd., U.K.); precision plus protein[™] standards (BioRad laboratories, U.K.); Sigma7 protein molecular weight markers (Sigma-Aldrich, U.K.).

2.1.6. Radiolabelled compounds

Radiolabeled ³²P or ³³P-ATP (GE Healthcare Ltd., U.K.); radiolabeled nucleobases (Perkin Elmer., UK; Sigma-Aldrich Co., U.K.); ¹⁴C-allantoin was synthesised by Dr

Simon Patching; ³H-L-5-benzylhydantoin and ³H-indolymethylhydantoin were synthesised by Dr Simon Patching and Dr Shun'ichi Suzuki.

2.1.7. Recombinant DNA technology

1 kb Plus DNA ladder and 1 kb DNA ladder (Invitrogen Ltd, U.K.); *E. coli* BL21(DE3) cells (EMD Biosciences, Inc. Novagen Brand, U.S.A.); dNTPs, Pfu Turbo[®] DNA polymerase, *Taq* thermal polymerase and thermal polymerase buffer, XL10-Gold ultracompetent cells (Stratagene, U.S.A); *EcoRI*, *HindIII*, *NdeI*, *PstI*, restriction enzyme buffers, T4 DNA ligase and T4 DNA ligase buffer (New England Biolabs[®] Inc., U.K.); Lambda DNA *EcoRI* and *HindIII* markers (Promega Ltd., U.K.); oligonucleotide primers (MWG-Biotech Ltd., U.K.); QIAquick gel extraction kit & QIAprep spin miniprep kit using a microcentrifuge (QIAGEN Ltd., U.K.); SYBR Safe[™] DNA gel stain (Invitrogen, U.S.A.).

2.1.8. Scintillation fluid and vials

Scintillation fluid (Perkin Elmer, U.K.); scintillation vials (Sarstedt Ltd., U.K.).

2.1.9. Substrates

Hydantoin (Fluka chemika); L-5-benzylhydantoin and L-5-indolymethylhydantoin were synthesised by Dr Simon Patching and Dr Shun'ichi Suzuki; nucleobases, nucleosides, allantoin, allantoic acid and uric acid (Sigma-Aldrich, U.K.).

2.1.10. SDS-PAGE and Western blotting

2 % bis-acrylamide, 40 % acrylamide (BioRad laboratories, U.K.); acetic acid, ammonium persulphate, BSA standards, brilliant blue R, bromophenol blue, glycine, KCl, mercaptoethanol, methanol, TEMED, TRIS, Triton X-100, Tween 20 (Sigma-Aldrich, U.K.); Fluorotrans[™] membrane (Pall BioSupport, U.K.); supersignal[®] west pico chemiluminescent substrate, HisProbe[™] HRP (Pierce Science UK Ltd., U.K.).

2.2. General microbiology

2.2.1. Bacterial strains used in this study

The *Escherichia coli* strains used throughout this study are detailed in Table 2.1.

Strain	Genotype	Source
XL10-Gold	Tet ^r Δ(<i>mcrA</i>)183Δ(<i>mcrCB-hsdSMR-mrr</i>)173 <i>endA1 supE44 thi-1 recA1gyrA96 relA1 lac</i> Hte [F ⁺ <i>proAB lacI^qZΔM15Tn10</i> (Tet ^r)Amy Cam ^r]	Stratagene™
BL21(DE3)	(F ⁻ <i>ompT hsdS_B (r_Bm_B) gal dcm</i> (DE3))	Novagen™

Table 2.1 *E. coli* strains used for cloning and expression. Genotypes after the nomenclature taken from (Sambrook *et al.*, 1989).

2.2.2. Growth and storage media

All media used for growth and storage of bacterial cultures are detailed in Table 2.2.

Medium	Ingredients
Luria Bertani (LB)	10 g/L tryptone; 10 g/L NaCl; 5 g/L yeast extract; pH to 7.5
Minimal	0.2% (w/v) casamino acids; 20mM glycerol; 2mM MgSO ₄ ; 0.2mM CaCl ₂ ; 1x M9 salts (stock of 5x M9 salts was made: 64 g/L Na ₂ HPO ₄ ; 15 g/L KH ₂ PO ₄ ; 2.5 g/L NaCl; 5 g/L NH ₄ Cl)
2x Tryptone / Yeast extract (2TY)	10 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl; pH to 7.5
Freezing mixture	12.6 g/L K ₂ HPO ₄ ; 0.9 g/L Na ₃ C ₆ H ₅ O ₇ ·2H ₂ O; 0.18 g/L MgSO ₄ ; 1.8 g/L (NH ₄) ₂ SO ₄ ; 3.6 g/L KH ₂ PO ₄ ; 96.0 g/L glycerol
TSS	LB-medium containing 10% (w/v) PEG 8000; 5% (w/v) DMSO; 25mM MgCl ₂ ; pH6.5
SOB	20 g/L Tryptone; 5 g/L yeast extract; 0.5 g/L NaCl; 20mM MgSO ₄

Table 2.2 Growth media for bacterial cultures.

Solutions were prepared using deionised or MilliQ™ water, and sterilised by autoclaving at 15 psi for 20 min at 121°C. Heat-labile solutions were filter-sterilised using 0.2 µm cellulose acetate filters. Carbenicillin (100 mg/ml) and tetracycline stocks (15 mg/ml) were prepared in water and ethanol respectively, filter sterilised and stored at -20°C.

A frozen stock of each clone was made by inoculating a single colony of BL21(DE3) cells harbouring the clone into 5 ml LB medium with 100 µg/ml carbenicillin and growing for approximately 4 hours. The culture, 0.5 ml was taken and added to an equal volume of freezing mix, quick frozen in liquid nitrogen and stored at -70°C.

2.2.3. Plasmid pTTQ18

Plasmid pTTQ18 (Stark, 1987) is based on the pUC high expression series of plasmids and contains a polylinker/*lacZ* alpha flanked by the strong hybrid *trp-lac* (*tac*) promoter (de Boer *et al.*, 1983) and the *rrnB* transcription terminator. A hexa-histidine site was engineered into plasmid pTTQ18 by Liang (2005) that introduces hexa-histidine tag (His-tag) into the C-terminal end of the protein for purification purposes, yielding pTTQ18-His₆ (Figure 2.1). This plasmid subsequently showed successful overexpression of many prokaryotic membrane proteins (Ward *et al.*, 2000; Saidijam *et al.*, 2003; Liang *et al.*, 2005; Potter *et al.*, 2002; Saidijam *et al.*, 2005; Surade *et al.*, 2006; Suzuki & Henderson, 2006; Betanney, 2008) and was used in this project.

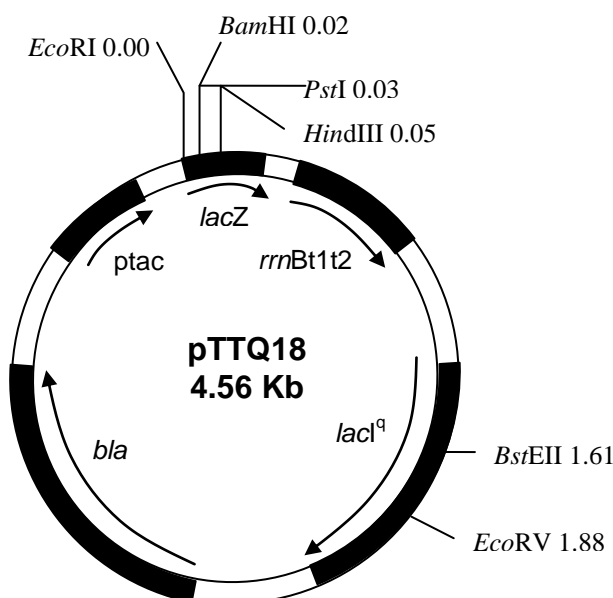


Figure 2.1 The plasmid pTTQ18. The multiple cloning site consists of several restriction enzymes sites illustrated by *EcoRI*, *BamHI*, *PstI* and *HindIII*. The actual multiple cloning site consists of (in the 5'-3' order) *EcoRI*, *EclI36*, *SacI*, *Asp718*, *KpnI*, *AvaI*, *SmaI*, *XmaI*, *BamHI*, *XbaI*, *AccI*, *SalI*, *Sse8387*, *PstI*, *SphI*, and *HindIII*. The plasmid comprises the synthetic *tac* promoter (*ptac*), over-expressed LacI repressor gene (*lacI^q*), and regions encoding β-lactamase (*bla*), and β-galactosidase alpha fragment gene (*lacZα*), and *E. coli rrnB* operon transcriptional terminator (*rrnBt12*) and pUC18 origin of replication (not shown). Diagram adapted from Ward *et al.* (2000).

2.3. DNA preparation and manipulation

Standard methods such as polymerase chain reaction (PCR) amplification, gel electrophoretic DNA separation, DNA extraction from agarose gel, plasmid preparation and restriction digestion were performed as described in Sambrook *et al.* (1989).

2.3.1. Cloning strategy – overview

Each target gene was inserted into the multiple cloning site (MCS) downstream of the *tac* promoter in the plasmid pTTQ18-*His*₆ in order to amplify protein expression (Figure 2.2). Two different restriction enzymes, *Eco*RI and *Pst*I were used to ensure correct orientation of the gene on ligation into plasmid pTTQ18-*His*₆, as well as preventing re-ligation of the plasmid. Firstly the membrane protein gene was amplified by PCR using bacterial genomic DNA as template, followed by digestion with *Eco*RI and *Pst*I and ligation with *Eco*RI-*Pst*I digested pTTQ18-*His*₆. The plasmid construct with the gene inserted was then transformed into *E. coli* XL10-Gold cells, followed by colony PCR to identify positive clones.

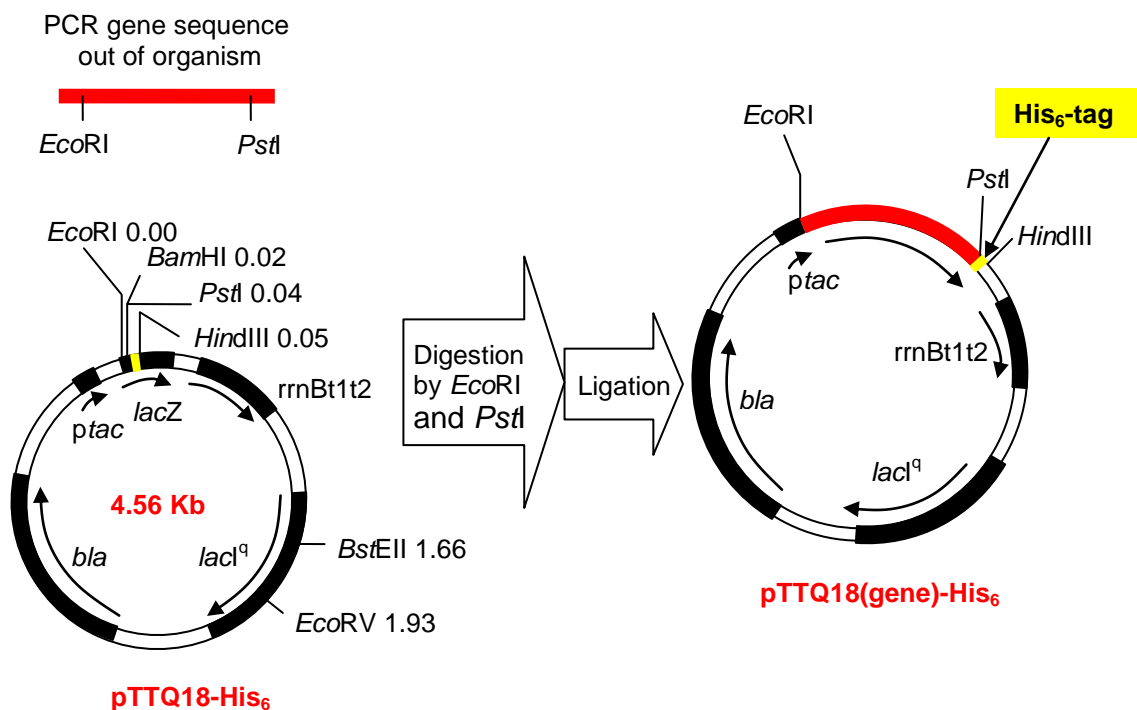


Figure 2.2 Cloning strategy for membrane proteins using plasmid pTTQ18-*His*₆. This cloning strategy was previously used by Peter J.F. Henderson lab members, successfully overexpressing membrane proteins (Ward *et al.*, 2000; Saidijam, 2004; Saidijam *et al.*, 2005; Bettaney, 2008).

2.3.2. Primer design and amplification of membrane protein genes

The DNA sequence of the gene of interest was obtained from TransportDB (Paulsen *et al.*, 1998; Ren *et al.*, 2007) or UniProt (The UniProt Consortium, 2008). Webcutter 2.0 (Heiman, 2000) was used to map the restriction sites in the gene to ensure absence of *EcoRI* and *PstI* sites, as these will be the restriction enzymes used for cloning. If an *EcoRI* or *PstI* site was found in the gene, *NdeI* or *HindIII* were used instead respectively. Since the *HindIII* site is located downstream of the hexahistidine site in the plasmid, the primer introducing the *HindIII* site must therefore additionally contain the His₆-tag sequence. The primers are designed to introduce an *EcoRI* site (GAATTC) or *NdeI* site (CATATG) in the 5' end, and a *PstI* site (CTGCAG) or *HindIII* (AAGCTT) in the 3' end of the gene to promote subsequent ligation with plasmid pTTQ18-His₆. The gene sequence was also submitted for topology prediction using TMHMM (Krogh *et al.*, 2001) to test that the C-terminal end of the protein is in the cytoplasm for incorporation of His₆-tag.

Some considerations for designing primers include: primers should be between 25 and 45 bases in length; melting temperature $\geq 70^{\circ}\text{C}$; a minimum GC content of 40% and usually terminating in G or C bases; should not form primer dimers or other secondary structures; primability $> 90\%$; stability 60-70%.

AmplifX 1.5.4 software was used to design and predict the quality of primers. Each primer pair was used in a PCR reaction to isolate a specific membrane protein gene from the bacterium of interest. Primers were prepared in sterilised water (15 pmol/ μl) and stored at -20°C . The primer pairs used in this project for the amplification of NCS1 genes are listed in Table 2.3, and for histidine kinase genes in Table 2.4.

For a 50 μl PCR mix, 2.5U PFU Turbo, 50 ng bacterium genomic DNA, 0.2 mM dNTPs, 0.3 pmole forward and reverse primers, 1x PFU Turbo buffer, 1 mM MgCl_2 were used. The reaction was carried out as follows:

Steps	Time
1) Warm-up at 94°C	3 min
2) Denaturation at 94°C	1 min
3) Annealing at 56°C	1 min
4) Extension at 72°C	4 min
5) Final extension at 68°C	15 min

NCS1 proteins	Primer sequence 5'→3' Forward primer (top) Reverse primer (bottom)
b0336	GCAG GAATTC GCATATGTCGCAAGATAACAACCTTTAGCC CCC AAGCTT TTA ATGATGATGATGATGATGATG TCGACACTGTTAGCCTCCACA
b0511	CCG GAATTC GCATATGGAACATCAGAGAAAACTATTCCAG AAAA CTGCAG CACCTATGGTTTTTTTGGCTCTCCTGTTTTTTTC
Bsu36470	CCG GAATTC GCATATGAAATTAAAAAGAGAGTCAGCAGCAATCCA AAAA CTGCAG CTTCAGCCTGGCGGACCTGCGCATGTT
lp_3374	CCG GAATTC GCATATGCAGCAAGAGTCAAAGTATCAAATAGAA AAAA CTGCAG CGACACGTTTGGCAGGTAACGGGTAGAAAAATTCG
PA0438	GCAG GAATTC GCATATGTCCGACAGCAGCAACTACAG CCC AAGCTT TTA ATGATGATGATGATGATGATG ATCAGCATGGAGGACCTCTTC
PA0443	CCG GAATTC GCATATGCAACAGAGCAGATCGGAAGTGACCGAGCGCGA AAAA CTGCAG CGGGCAGCGGTTTGGCCAATGCCGCCGGGCTCCG
PA0476	CCG GAATTC GCATATGTCCAGCACCCCGAAGCCGCCGCTCG AAAA CTGCAG CGTGCTGCACGCTGTCCACCGCATCGCTTCG
PA2073	CCG GAATTC GCATATGAGCAATAACAACGAGGCGTCACGCCCC AAAA CTGCAG CGGCGCCCCCTCTTCCAGCAGGCGCTCGC
PPA0619	CCG GAATTC GCATATGACGCAGACCGCCACCAAAAACG AAAA CTGCAG CTCGTGACTCCTGACGACGCACGGT
PPA2211	CCG GAATTC GCATATGGTGTCCGCGCCAAAATAGTGCTC AAAA CTGCAG CCCGAGAAACAATGCTTGGATCCAGC
SCO0572	GCAG GAATTC GCATATGGCCATGCACAAGAACCCCGCCG CCC AAGCTT TTA ATGATGATGATGATGATGATG CGGTGCTCCTACGGTGGCCGTG
SCO5579	CCG GAATTC GCATATGAGCAAGACCGCCGAGAACGAAGGC AAAA CTGCAG CGACGGTCACGGTCTCGGCCCTCGGGCTGCT
SCO6417	CCG GAATTC GCATATGATTGGCCTGCCCATGACCGACACCGCGCC AAAA CTGCAG CGGCGGAACGGTTGCGGCTTCCGGCCCGCTCGGT
SCO7500	CCG GAATTC GCATATGACTGCCCTGGAAAGACCGCCGCTCGTC AAAA CTGCAG CTTCCGCGGCGACGGCGATGGACTCGCCGTCCAC

Table 2.3 Primer sequences for PCR amplification of NCS1 family genes. The *EcoRI*, *PstI* and *HindIII* sites are in blue, red and green respectively. The hexahistidine tag is coloured purple for the primers containing *HindIII* restriction site.

For the amplification of genes from a bacterium with high GC content, such as for *Streptomyces coelicolor* and *Pseudomonas aeruginosa*, the annealing temperature was lowered to 40°C with addition of 5-10% DMSO in the PCR mix as detailed below.

Steps	Time
1) Warm-up at 94°C	3 min
2) Denaturation at 94°C	1 min 30 secs
3) Annealing at 40°C	1 min
4) Extension at 72°C	4 min
5) Cycle to step 2 for 2 more times	
6) Denaturation at 94°C	30 secs
7) Annealing at 40°C	1 min
8) Extension at 72°C	4 min
9) Cycle to step 6 for 25 more times	

HK no.	Histidine kinases	Primer sequence 5' → 3' Forward primer (top) Reverse primer (bottom)
01	EF2219	CCG GAATTC GCATATGAAGCGAATCTTCAAAAACTGT AAAA CTGCAG CTGCTACAAACGAAAGACGAATTG
02	EF3197	CCG GAATTC GCATATGGTTCGAATTATTTATTTTAAATGATGGAA AAAA CTGCAG CACGATTAATACGTGCATCTAATTCCTCCTGT
03	EF2912	CCG GAATTC GCATATGACCGATCGGATTTCAAGACGCAT AAAA CTGCAG CTCCACTAACATTACTTTGATCACTTGC
04	EF1704	CCG GAATTC GCATATGAAAAAGAGACTGCGGATTGAATATTT AAAA CTGCAG CGGCTTCTTTTTTTGTTAAAAGCAAAGTG
05	EF3290	CCG GAATTC GCATATAGCTCGTTAAACCTAAAAAGATTGAAAA AAAA CTGCAG CACTCTCTGAATTTCTTGTGGTACGC
06	EF1261	CCG GAATTC GCATATGAAGTATTTGTATCAACAATTA AAAA CTGCAG CGTTTTTCAGTAATTTCAACATCAGGAAAT
07	EF1194	CGATCAT GAATTC CAAGAAAAAAGTTCACTTT CGATC CTGCAG CTTCCCACCAATCCTCTTCATA
08	EF1863	CGATCAT GAATTC CAAAAAATTAAAGATATTT CATC CTGCAG CAGAAAATATTTTATTTTAAAGAT
09	EF0927	CCG GAATTC GCATATGACCATTCTAAAGTATTTAAAAGAT AAAA CTGCAG CTTTCACCTCATTATAATAACTGAGGGA
10	EF1051	CCG GAATTC GCATATGAAAAGAACGATTAATAAAGAACTG AAAA CTGCAG CTTGTACGTTGCTCTCTCTTTTATTCCAGCGA
11	EF2298	CCG GAATTC GCATATGGAAAGAAAAGGGATTTTCATTAAG AAAA CTGCAG CTAGTGTGATGTGGGCGGTAAATC
12	EF0570	CCG GAATTC GCATATGCAATCTGGCGATGGACGCC AAAA CTGCAG CCTTCAAATAAATTCGAACCAATGTTT
13	EF0373	CCGGAATTC GCATATG ACCATTAAACGCCGCTTTTTTTA AAAA CTGCAG CTTTTTTCATCCTCCAACAACGGCAGC
14	EF1209	CCG GAATTC GCATATGAAGCGTGGCGGAAATTATGGTGT AAAA CTGCAG CATAATCAATAGGTTTCAATTTTCTCTGCTC
15	EF1820	CCG GAATTC GCATATGATTTTGTGCTTATTAGCTACTAACGTTT AAAA CTGCAG CTTCGTTAACAACTTTTTTACTGTTATGATT
16	EF1335	CCG GAATTC GCATGCTTGATATAACATTACGAAGCCGAAGCC AAAA CTGCAG CGTTATCACCGCATTTGTTTCTTTGTA

Table 2.4 Primer sequences for PCR amplification of *E. faecalis* sensor kinase genes. The *EcoRI*, *PstI* and *NdeI* sites are in blue, red and brown respectively. Primers for EF3197, EF2912, EF1051, EF0373, EF1209 and EF1820 were designed by myself. Primers for EF1194 were designed by Dr Victor Blessie, and the remaining primers were designed by Dr Hayley M. Yuille.

2.3.3. Preparation of plasmid pTTQ18-His₆

To clone each membrane protein gene into pTTQ18-His₆, the plasmid pNorAH6 (pTTQ18-His₆ containing the *norA* gene) (Hoyle, 2000) was isolated from *E. coli* strain BLR and digested with restriction endonucleases *EcoRI* or *NdeI* and *PstI* or *HindIII*. This yields two fragments, a small fragment of ~1200 bp which is the *norA* gene and a larger fragment of pTTQ18-His₆ which was isolated from the agarose gel for cloning.

2.3.4. Agarose gel electrophoresis and restriction digestion of PCR product

The PCR products were separated and analysed on 1% (w/v) agarose gel electrophoresis. The 1kb DNA ladder was used for size determination of the PCR product (Table 2.5), the key product confirmed was excised and extracted from the agarose gel using QIAquick gel extraction kit (QIAquick Spin Handbook 07/2002) and subjected to digestion with *EcoRI* / *NdeI* and *PstI* / *HindIII* or for 3 hours or overnight at 37°C, this yields the gene with the overhangs for ligation into pTTQ18-His₆.

Fragments	1 kb DNA extension ladder (Invitrogen™) (1 µl = 50 ng)		Lambda DNA/ <i>EcoRI</i> + <i>HindIII</i> marker (Promega™) (5 µl = 25 ng)	
	Base pairs	Amount (ng)	Base pairs	Amount (ng)
1	12,216	10	21,226	109.0
2	11,198	7	5,148	26.5
3	10,180	7	4,973	25.4
4	9,162	7	4,268	21.9
5	8,144	7	3,530	19.0
6	7,126	7	2,027	10.4
7	6,108	6	1,904	9.8
8	5,090	6	1,584	8.15
9	4,072	6	1,375	7.0
10	3,054	6	947	5.0
11	2,036	6	831	4.2
12	1,636	10	564	2.9
13	1,018	5	125	0.6
14	506	3	-	-

Table 2.5 DNA molecular weight markers used for determination of DNA concentration and size. Table modified from Saidijam (2004).

Restriction digestion of the PCR products is as follows: 30 µl PCR product; 29 µl sterile MQ™ water; 40 Units *EcoRI* / *NdeI* restriction enzyme; 40 Units *PstI* / *HindIII* restriction enzyme; 7 µl 10x restriction buffer.

The following buffers or stocks were for agarose gel electrophoresis: 50x TAE buffer (2 M Tris, 1 M Acetate, 50 mM EDTA); 1% (w/v) agarose gel (70 ml 1x TAE, 0.7 g agarose, 1 µl 10 mg/ml SYBR Safe™ DNA gel stain); 10x DNA loading buffer (0.005% bromophenol blue, 50% glycerol).

2.3.5. Quantification of DNA

2.3.5.1. The Lambda DNA marker method. To determine the concentration of a DNA sample, 1 μ l of the DNA was run on a 1% agarose gel electrophoresis alongside the Lambda DNA *EcoRI+HindIII* Markers of known DNA concentrations. The intensity of the DNA band was compared to the markers to determine the concentration. This is an approximate estimation of the DNA concentration.

2.3.5.2. The absorbance method. A dilution of 1:100 of pTTQ18-His₆ in MQTM water was made and absorbance at 260 nm and 280 nm measured using the WPA Biowave S2100 Diode Array Spectrometer. The measured absorbance at 260 nm was used to calculate the DNA concentration using the standard rule where 1 mg/ml of DNA has an $A_{260\text{nm}}$ of 20. A ratio of $A_{260\text{nm}} / A_{280\text{nm}}$ gives the purity of the DNA sample. For a pure DNA sample, $A_{260\text{nm}} / A_{280\text{nm}} = 1.8$. Above 1.8 is RNA contamination, and below 1.8 is protein contamination.

2.3.6. Ligation of gene into plasmid pTTQ18-His₆

Ligation reactions were performed using 1:1, 1:3 and 1:5 plasmid to gene molar ratios (Table 2.6). Ligation mixture 4 and 5 are controls with no ligase or no gene.

	Ligation mixture				
	1	2	3	4	5
pTTQ18-His ₆	5 ng/ μ l	5 ng/ μ l	5 ng/ μ l	5 ng/ μ l	5 ng/ μ l
gene	5 ng/ μ l	15 ng/ μ l	25 ng/ μ l	15 ng/ μ l	No gene
DNA T4 ligase	1 μ l	1 μ l	1 μ l	No ligase	1 μ l
10x T4 ligase buffer	1.5 μ l	1.5 μ l	1.5 μ l	1.5 μ l	1.5 μ l
sterile H ₂ O	10.5 μ l	8.5 μ l	6.5 μ l	9.5 μ l	11.5 μ l
Total volume	15 μ l	15 μ l	15 μ l	15 μ l	15 μ l

Table 2.6 Ligation reactions using various plasmid:insert ratios.

Firstly, water, plasmid and gene were added into a 1.5 ml eppendorf tube and were mixed by gentle pipetting. The sample was centrifuged at 1400 g for 2 secs and placed in a water bath at 45°C for 5 min, then it was placed on ice for 2 min and the T4 ligase buffer and T4 ligase added. The ligation mix was centrifuged at 1400 g for 2 secs and left at room temperature for 10 min before incubation for 1 hour at 37°C. Prior to transformation into *E. coli* XL10-Gold cells (Section 2.3.8), the ligation mix was cooled on ice for 5 min.

2.3.7. Preparation of *E. coli* competent cells and transformation

2.3.7.1. Chung et al. (1989) method. *E. coli* Competent cells were prepared by inoculating a single colony into 5 ml of LB medium and incubation overnight at 37°C and 220 rpm. This stationary-phase overnight culture (0.5 ml) was transferred to 50ml LB medium and incubated at 37°C and 220 rpm until the Absorbance at 600 nm (A_{600nm}) reached 0.3 – 0.4 (after approximately 2 hours). The cells were chilled on ice for 10 min, followed by centrifugation at 2500 g for 10 min at 4°C and resuspended in 1/10 the original volume of ice-cold TSS (Table 2.2) solution (5 ml ice-cold TSS); then it was ready to be transformed. Cells prepared by this method were stored at 4°C for a maximum of 6 hours without significant loss of competency, as well quick frozen in liquid nitrogen for long-term storage at -70°C.

For transformation, 1 µl of DNA (1-200 ng) was added to 100 µl ice cold competent cells and incubated on ice for 30-60 min, followed by addition of 900 µl of TSS or LB medium plus 20 mM glucose and the cells were incubated at 37°C for 60 min. The cells were then centrifuged at 2500 g for 2 min, plated on 1.5 % agar-LB, and incubated at 37°C overnight. For transformation of the ligation mix, 15 µl of ligation mix was added to 85 µl competent cells and the method followed as above.

2.3.7.2. Inoue et al. (1990) method. Approximately 10-12 large colonies were inoculated into 250 ml SOB medium in a 2 litre flask, incubated at 18°C with 220 rpm until cell density reached A_{600nm} of 0.6 (approximately 36-40 hours). Subsequently, the cells were placed on ice for 10 min prior to centrifugation at 2500 g for 10 min at 4°C. The cell pellet was gently resuspended in 80 ml ice-cold transformation buffer (10 mM PIPES, 55 mM MnCl₂, 15 mM CaCl₂·2H₂O, 250 mM KCl) with 7% DMSO and were left on ice for a further 10 min prior to aliquoting into 500 µl volumes. The cells were frozen in aliquots in liquid nitrogen and stored at -70°C.

For transformation, a frozen stock of competent cells was thawed on ice for approximately 20 min, 85 µl was added per ligation reaction and mixed gently (for plasmid transformation, 3 µl plasmid in 50 µl competent cells). The ligation was placed on ice for 30 min, followed by heat shock at 42°C for 1 min and on ice again for 2 min. The reaction was made up to 1 ml with LB medium, incubated in an orbital incubator at 37°C for 1 hour, then pelleted at 2500 g. Some supernatant was removed (800 µl) and the cells were resuspended in the remaining 200 µl and plated out onto 1.5% agar-LB containing 100 µg/ml carbenicillin and grown overnight at 37°C.

2.3.8. Transformation of XL-10 Gold cells

The methods of Chung *et al.* (1989) or Inoue *et al.* (1990) were used, described in Section 2.3.7.

2.3.9. PCR screening for the presence of the correct gene in a recombinant clone

As the efficiency of restriction digestion may not be 100%, it is possible the prepared pTTQ18-His₆ used for ligation may contain undigested plasmids, resulting in carbenicillin-resistant colonies of the original parent plasmid with the *norA* gene (Section 2.3.3.). The colonies grown are therefore subjected to PCR screening to confirm the presence of the correct gene. To PCR screen, a colony of XL-10 Gold cells was suspended in 20 µl sterile MQTM water, and 10 µl taken and added to an equal volume of 2x PCR mix (0.2U Taq DNA polymerase, 1x Taq buffer, 0.4 mM dNTPs and 1pmol forward and reverse primers of the relevant gene), using the following conditions.

Steps	Time
1) Warm-up at 95°C	5 min
2) Denaturation at 95°C	1 min
3) Annealing at 40, 56 or 62°C	30 secs
4) Extension at 72°C	1 min
5) Cycle to step 2 for 25 more times	
6) Final extension at 72°C	10 min

The PCR product (3 µl PCR product, 6 µl MQTM water, and 1 µl 10x DNA loading buffer) was subjected to 1% agarose gel electrophoresis to identify the amplified positive clones. The remaining 10 µl of the positive colony was grown in 5ml LB medium containing 100 µg/ml carbenicillin at 37°C and 220 rpm overnight, followed by plasmid extraction (Section 2.3.10.)

2.3.10. Extraction of plasmid

Plasmids were extracted using the ‘QIAprep Spin Miniprep Kit Using a Microcentrifuge’ method (QIAprep Miniprep Handbook 05/2004).

2.3.11. Restriction digestion analysis

The extracted plasmids were subjected to restriction digestion analysis with *EcoRI* / *NdeI* and *PstI* / *HindIII* (Table 2.7) to confirm the correct size of the gene in plasmid pTTQ18-His₆.

	Double digest	Single digest	Control
Plasmid	2 µl	2 µl	2 µl
<i>EcoRI</i> or <i>NdeI</i> (20,000 U/ml)	1 µl	1 µl	-
<i>PstI</i> or <i>HindIII</i> (20,000 U/ml)	1 µl	-	-
10x <i>EcoRI</i> buffer	1 µl	1 µl	1 µl
MQ™ water	5 µl	6 µl	7 µl

Table 2.7 Restriction digestion reactions to verify the present of target gene. Plasmids were single and double digested. Control is undigested.

2.3.12. Automated DNA sequencing

The plasmid constructs were submitted for automated DNA sequencing. This was performed using equipment and protocols by Amersham Biosciences DYEnamic ET™ Terminator Cycle Sequencing Protocol in conjunction with the ABI 373A automated DNA sequencer in the sequencing facility at the University of Leeds (performed by Denise Ashworth). The following thermal cycling conditions were used:

Steps	Time
1) Rapid thermal ramp to 95°C	20 secs
2) Rapid thermal ramp down to 50°C	15 secs
3) Rapid thermal ramp up to 60°C	60 secs
4) Cycle to step 1 for 25 more times	

To sequence the plasmid constructs 10 µl 50 ng/µl plasmid and 4 µl 1.6 pmol/µl primers were prepared and submitted for sequencing. Table 2.8 shows the primers used for DNA sequencing. These are forward and reverse pTTQ18-His₆ primers that bind upstream and downstream of the inserted gene.

	Primer sequence 5'→3'
pTTQ18-His ₆ forward primer	ATGAATTCGCAT
pTTQ18-His ₆ reverse primer	GCTGCAGGCGGTCGTGGCAGCCACCATCACC ATCACCATTAATAG

Table 2.8 Primers used for sequencing of plasmid constructs.

2.3.13. Transformation of plasmid into *E. coli* BL21(DE3) cells

The plasmid constructs were transformed into *E. coli* BL21(DE3), for expression trials, using the methods of Chung *et al.* (1989) or Inoue *et al.* (1990) (Section 2.3.7).

2.4. Protein preparation and manipulation

Methods for protein preparation and manipulation are found in Ward *et al.* (2000).

2.4.1. Testing for the expression of cloned protein

Cultures for expression tests were prepared by inoculating 5 ml of LB media with 100 µg/ml carbenicillin with a single colony of BL21(DE3) cells harbouring pTTQ18(gene)-His₆ and incubating overnight at 30°C or 37°C and 220 rpm. 1.5 ml of the overnight culture was transferred to 50 ml LB medium with 100 µg/ml carbenicillin and 20 mM glycerol in a conical flask and grown at 37°C and 220 rpm. The cells were induced with 0.5 mM IPTG when A_{680nm} reached 0.4-0.6 and were harvested 3 hours post-induction by centrifugation at 2500 g for 10 min. Cells were stored at -20°C if a membrane preparation was not immediately carried out.

2.4.2. Preparation of total *E. coli* membranes

The total membrane fraction was prepared by the water lysis procedure (Ward *et al.*, 2000).

2.4.3. Determination of protein concentration

The Schaffner and Weissman (1973) protein assay was used to determine the concentration of proteins in the membrane preparations. A calibration curve was constructed from a range of standards of 0.1 mg/ml Bovine Serum Albumin (BSA) between 0-20 µg (Figure 2.3). All procedures were performed at room temperature. Each sample was made up to 270 µl with deionised water. 30 µl 1M TRIS pH7.5, 10% w/v SDS, and 60 µl 60% w/v TCA were added, the mixture vortexed and left for 20 min. The precipitated proteins were transferred drop wise onto a pre-soaked 0.45 µm HAWP Millipore membrane filter under a vacuum manifold. Remnant proteins in the tube were washed with 0.3 ml of 6 % TCA and transferred to the filter. The filter was washed two times with 2 ml 6 % TCA and transferred to a Petri dish containing 30 ml naphthalene black stain (0.1 % naphthalene black in methanol, acetic acid and water 45:10:45 v/v) for 3 min, followed by washing in deionised water for 30 secs and subsequent washing two times in destain (methanol, acetic acid, water 90:2:8 v/v) for 1 min. The filters were then washed once in deionised water and blotted onto paper towel and the stained spots cut out.

Each spot was transferred to a 1.5 ml eppendorf tube and the stained protein eluted by addition of 1 ml 25 mM NaOH, 0.05 mM EDTA, 50% aqueous ethanol and left for 1 hour, vortexing at intervals. Absorbance at 630 nm for the protein samples was recorded and the linear relationship between absorbance and BSA concentration (Figure 2.3) was used to determine unknown concentration. The standard curve was drawn using the GraphPad Prism computer software.

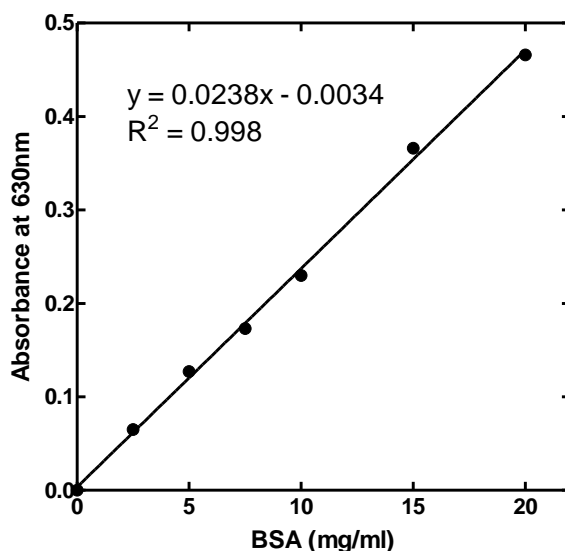


Figure 2.3 A typical calibration curve using BSA standards (Section 2.4.3.).

2.4.4. Analysis of proteins by SDS-PAGE

The prepared membranes were analysed by SDS-PAGE composed of a 15% (v/v) resolving gel and a 4 % (v/v) stacking gel (Table 2.9). Proteins (1-30 μ g) were solubilised in 1x sample loading buffer and incubated at 37 °C for 20 min. The molecular weight standards used are detailed in Tables 2.10 and 2.11. The samples were loaded onto the gel and electrophoresis performed for approximately 2 hours at 30 mA in running buffer. The gel was stained overnight in Coomassie brilliant blue and destained in SDS gel destain solution.

The following stock solutions were made: *4x sample loading buffer* (60 mM TrisHCL pH 7.2; 10% v/v glycerol; 2% w/v SDS; 0.005% bromophenol blue; 3% β -mercaptoethanol); *5x SDS running buffer* (6 g/L Tris; 28.8 g/L glycine; 1 g/L SDS); *Coomassie brilliant blue* (90 ml/L glacial acetic acid; 450 ml/L methanol; 1 g/L Coomassie brilliant); *SDS gel destain* (50 ml/L methanol; 70 ml/L glacial acetic acid).

Stock solutions	Resolving gel (15%)	Stacking gel (4%)
40% acrylamide	4.22 ml	0.77 ml
2% bis-acrylamide	0.49 ml	0.39 ml
1.5 M TrisHCl (pH 8.7)	2.81 ml	-
1.5 M TrisHCl (pH 6.8)	-	0.75 ml
10% SDS	0.10 ml	0.05 ml
deionised H ₂ O	3.48 ml	3.20 ml
10% ammonium persulphate (APS)	37 μ l	30 μ l
TEMED	12 μ l	9 μ l

Table 2.9 Composition of SDS-PAGE gel with 15% resolving and 4% stacking gels.

Protein	Mwt (kDa)
Bovine serum albumin	66.0
Ovalbumin	45.0
G-3-P dehydrogenase	36.0
Carbonic anhydrase	29.0
Trypsinogen	24.0
Trypsin inhibitor	20.1
Alpha-Lactalbumin	14.2

Table 2.10 SDS-PAGE molecular weight markers (Sigma). This marker was used for protein molecular weight (MWt) determination in SDS-PAGE.

Protein	Mwt (kDa)
Myosin (blue)	220.0
Phosphorylase b (brown)	97.0
Bovine serum albumin (red)	66.0
Ovalbumin (yellow)	45.0
Carbonic anhydrase (orange)	30.0
Trypsin inhibitor (blue)	20.1
Lysozyme (magenta)	14.3

Table 2.11 High-range rainbow molecular weight markers (Amersham Biosciences) used for protein molecular weight determination in Western blots and chemiluminescent detection.

A linear relationship between log molecular weight (Mwt) of standard protein markers and its distance migrated on the SDS-PAGE gel was generated to determine the size of a particular protein on the SDS-PAGE gel (Figure 2.4). The standard curve was drawn using the GraphPad Prism computer program.

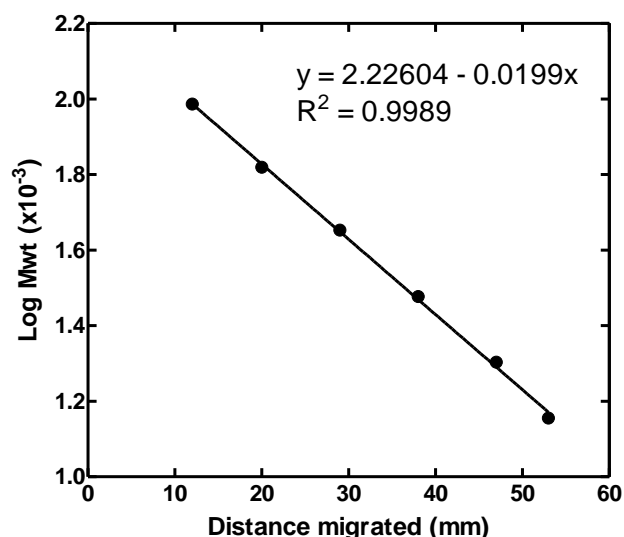


Figure 2.4 A typical log Mwt-distance curve to determine size of protein on a SDS-PAGE.

2.4.5. Drying of SDS-PAGE gels onto cellophane

SDS-PAGE gels were photographed using FluoroS MultiImager (Biorad™) and gels were dried onto cellophane. Prior to drying the gel, the destained gel and the cellophane were pre-soaked in deionised H₂O for 5 min. The gels were enclosed between two layers of cellophane and held firmly using a gel frame. All air bubbles were removed and the gel was left to dry at room temperature overnight.

2.4.6. Detection of His₆-tagged proteins by immunoblotting & chemiluminescent

The proteins separated by SDS-PAGE gel were transferred to Fluorotrans™ membrane by construction of a blot sandwich between two layers of filter papers (Towbin *et al.*, 1979). The filter papers were pre-soaked in 1x SDS running buffer and the membranes soaked in methanol. The transfer of proteins was achieved by application of 50 mA for 2 hours to the blot sandwich in a TE77 semi-dry transfer unit. After the proteins were transferred, the membrane was washed in 1x Tris buffer saline containing Tween 20 and Triton X100 (TBSTT) (20 mM TRIS pH 7.5, 500 mM NaCl, 0.05% v/v Tween 20, 0.2% Triton X100) for 10 min and blocked overnight in 3 % BSA TBSTT at 4°C. The membrane was then washed twice for 10 min in 1x TBSTT, then incubated in 10 ml HisProbe™-HRP (1 in 5000 dilution) for 1 hour at room temperature. A further four washes in 1x TBSTT were carried out, followed by incubation for 5 min in 10 ml SupersignalWest Pico Chemiluminescent (5 ml West Pico stable peroxide solution and 5 ml West Pico luminol/enhancer solution)(Supersignal®).

The membrane was wrapped in cling film and exposed for 1 min to X-ray film in X-ray 20. Rainbow markers were used for subsequent identification of the position of the signal produced by superimposition onto the PVDF membrane.

2.4.7. Optimisation of protein expression

The effects of growth medium (LB / 2TY / Minimal) and IPTG concentration on protein expression levels were investigated. Previous expression tests involved 0.5 mM IPTG; therefore a range of 0.1 mM, 0.25 mM, 0.5 mM, 0.75 mM and 1 mM IPTG was investigated using the same method described in Section 2.4.1 to 2.4.4. Also the time of harvesting after IPTG induction was investigated. The levels of overexpressed protein were analysed using scanning densitometry.

2.5. Biochemical techniques

2.5.1. Preparation of inner membranes using sucrose gradients

A total of 4-6 L culture of *E. coli* BL21(DE3) expressing his-tagged proteins were grown aerobically in 500-750 ml batches in 2 L flasks and induced with IPTG as described in Section 2.4.1. and harvested by centrifugation at 5000 g for 15 minutes. Pelleted cells were resuspended in 20 mM Tris, 0.5 mM EDTA and 10 % glycerol, and lysed by explosive decompression using a cell disruptor to produce predominantly inside-out membrane vesicles that were purified by separation on sucrose density gradients (Ward *et al.*, 2000).

2.5.2. Solubilisation trials of membrane proteins using detergents

This method is described previously (Gutmann *et al.*, 2007). Inner membranes (100 µg) were incubated with 100 µl solubilisation buffer containing 20 mM Tris pH 8.0, 20 mM imidazole pH 8.0, 300 mM NaCl, 20% v/v glycerol and 1% w/v detergents at 4°C for 1 hour prior to centrifugation at 100,000 g for 30 min to pellet any insoluble proteins. The detergents trialed included: n-hexyl-β-D-maltoside; n-octyl-β-D-maltoside; n-nonyl-β-D-maltoside; n-decyl-β-D-maltoside; n-undecyl-β-D-maltoside; n-dodecyl-β-D-maltoside; n-hexyl-β-D-glucopyranoside; n-heptyl-β-D-glucopyranoside; n-octyl-β-D-glucopyranoside; n-nonyl-β-D-glucopyranoside; n-decyl-β-D-glucopyranoside; cyclohexyl-n-hexyl-β-D-maltoside; Triton X-100; ethyleneglycol monoethylether (C₈E₁); octaethyleneglycol monododecylether (C₁₂E₈);

nonaethyleneglycol monodocecylether (C₁₂E₉); N,N-dimethyl dodecylamine oxide (LDAO); N,N-dimethyl decylamine oxide (DDAO); N,N-dimethyl octylamine oxide (ODAO); N-octyltetraoxyethylene; N-octylpentaoxyethylene; methyl-6-O-(N-heptylcarbamoyl)- α -D-glucopyranoside (HECAMEG); n-octyl- β -D-glucoside (β OG); and n-octyl- β -D-maltopyranoside. The total inner membrane and solubilised membrane fractions were analysed using SDS-PAGE and compared, to determine the location of each protein and therefore its solubilisation in detergent (Figure 2.5).

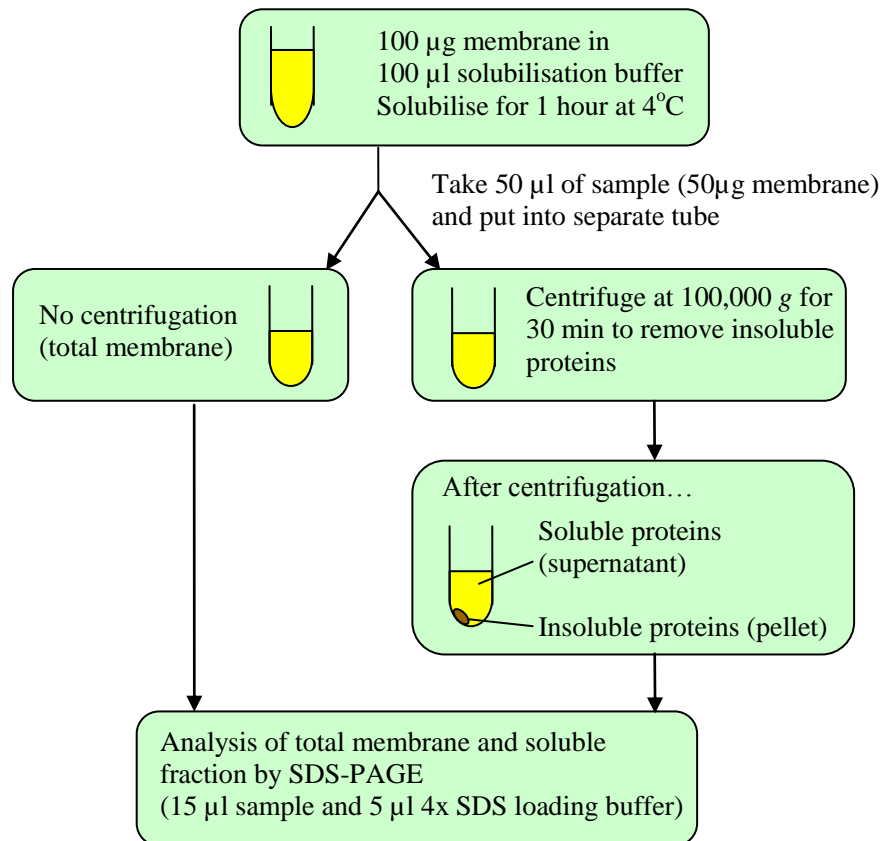


Figure 2.5 Procedures for solubilisation trial of membrane proteins. Membranes were solubilised in 1% w/v detergent, and one half of the sample was subjected to centrifugation to remove insoluble proteins. Soluble protein fraction obtained post-centrifugation was analysed together with the total membrane fraction by SDS-PAGE.

2.5.3. Purification of hexahistidine-tagged proteins

2.5.3.1. Purification of NCS1 family proteins. 2-4 mg/ml of inner membranes were incubated in solubilisation buffer (20 mM Tris pH8.0, 20 mM imidazole pH 8.0, 300 mM NaCl, 20% glycerol v/v and 1% w/v DDM) for 1 hour at 4°C (Mhp1 of *M. liquefaciens* and PuI of *B. subtilis* were solubilised for 2 hours or overnight). The insoluble fractions were removed by centrifugation at 100,000 g for 30 min and the solubilised fractions incubated with Ni-NTA for 30 min at 4 °C (Mhp1 membranes were

bound to Ni-NTA overnight), followed by transfer to a Bio-rad Inc. column to remove any unbound proteins. The resins were washed with 20 mM Tris pH 8.0, 10 or 20 mM imidazole pH 8.0, 1M NaCl (no NaCl for Mhp1), 5% glycerol and 0.05% DDM. The bound (hexahistidine-tagged) proteins were then eluted with 10 mM Tris HCl pH 8.0, 300 mM imidazole pH 8.0, 2.5% glycerol and 0.05% DDM (No NaCl was used in the elution for Mhp1). The protein was concentrated using a Centricon™ 30 kDa MWCO at 3000 g.

2.5.3.2. Purification of sensor kinases. Inner membranes (2.5-5 mg/ml) were incubated for 1-2 hours at 4°C in solubilisation buffer (20 mM Tris HCl pH 7.9, 20 mM imidazole pH 7.9, 2 mM β-mercaptoethanol, 20% glycerol, 250-300 mM NaCl and 1 % w/v detergent). Insolubilised proteins were removed by centrifugation at 100,000 g for 30 min, and the soluble proteins incubated between 0.5-15 hours with Ni-NTA at 4°C with gentle rotation, followed by transfer to a Bio-rad Inc. column to remove the unbound proteins. The bound proteins in the resin were washed with 20 mM Tris HCl pH 7.9, 20 mM imidazole pH 7.9, 2 mM β-mercaptoethanol, 10-20% v/v glycerol and 0-0.05% DDM, and in most cases 250 mM NaCl. The His₆-tagged protein was trialed with increasing concentrations of imidazole, 0.1 M, 0.2 M, 0.3 M and 0.5 M in wash buffer to elute the protein, and samples were analysed by SDS-PAGE. The eluted fractions containing the His₆-tagged protein were concentrated using Centricon™ 30 kDa MWCO at 3000 g, and exchanged into the required buffers for further analysis.

2.5.3.3. Size exclusion chromatography. Purified proteins were buffer exchanged into 25 mM Tris-HCl pH 8.0, 100 mM NaCl, 5% glycerol and 0.05% DDM (FPLC buffer) using BioRad Econo-Pac® 10DG column and concentrated using Centricon™ 50 kDa MWCO at 3000 g, followed by centrifugation at 13,000 rpm, 4°C, for 5 minutes on a bench-top centrifuge to remove particulate material. The concentrated protein sample (<1ml) at 10 – 15 mg/ml was loaded onto a S-200 column equilibrated with FPLC buffer and eluted at a flow rate of 1 ml/min. All buffers used were degassed and a constant temperature (4°C) was maintained throughout the size exclusion chromatography. Proteins were eluted in 10 ml fractions (first 40 ml) then in 1ml fractions, detected by UV absorbance at 280 nm, and subsequently analysed by chromatograph and SDS-PAGE to determine the monodispersity and purity. A sample (15 µl) was taken from each 1 ml fraction and 5 µl 4x SDS loading buffer was added prior to loading onto SDS-PAGE for analysis.

2.5.4. Uptake of radiolabeled substrates by the NCS1 family of transporters

Transport assays using intact bacteria were carried out as described Henderson & Macpherson (1986). A single colony was inoculated into 5 ml LB medium containing 100 µg/ml carbenicillin and was cultured aerobically overnight at 30°C and 220 rpm. For uptake assays, 1.5 ml of overnight culture was inoculated into 50 ml minimal medium (Section 2.2.2.) and induced with 0.5 mM IPTG when cell density at $A_{680\text{nm}}$ reached 0.4-0.6 and harvested 1 hour post-induction by centrifugation at 2800 g for 10 min. Subsequently, the cells were washed three times with 30ml KCl.MES buffer (150 mM KCl, 5 mM MES pH6.6) and resuspended in the same buffer to a cell density of 2 at $A_{680\text{nm}}$. For each uptake, 482.5 µl of cell suspension was energised with 20 mM glycerol in a bijoux bottle at 25°C and aerated for 3 min. At time = 0, 12.5 µl of radiolabelled substrate was added to the cells and vortexed; at time = 15 secs and time = 2 min, 200 µl of the sample was transferred onto pre-soaked 0.45 µm Whatman cellulose nitrate membrane filters under vacuum, followed by immediate washing with 4 ml ice-cold KCl.MES. The filter was transferred to a 20 ml scintillation vial. Standards were prepared by adding 4 µl of radiolabelled substrate to a pre-wetted filter. Background was a pre-wetted filter on the manifold washed with KCl.MES buffer. All filters were then submerged in 10ml scintillation fluid (Emulsifier-Safe™) and counted using a TRI-CARB PACKARD liquid scintillation analyzer 1500. The number of milligrams of cells and the rate of radiolabelled substrate uptake were determined using the following equations:

$$A_{680\text{nm}} \text{ of } 1 = 0.68 \text{ mg of dry cell mass}$$

$$\text{mg of cells} = \text{total sample volume} \times A_{680\text{nm}} \times 0.68$$

$$\text{Uptake (nmol/mg)} = (\text{dpm-background}) \times \frac{\text{total volume}}{\text{sample volume}} \times \frac{1}{\text{mg cells}} \times \frac{\text{nmol standard}}{\text{dpm standard}}$$

2.5.5. Autophosphorylation assays for sensor kinases and signal testing

2.5.5.1. Assays using inner membranes. These assays were performed as described previously (Potter *et al.*, 2002; Potter *et al.*, 2006; Ma *et al.*, 2008). Briefly, inner membrane proteins (200 µg) were incubated at 24°C in assay buffer containing 50 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 50 mM KCl, 0.005% DDM and 1 mM DTT. Total reaction volumes were 200 µl and reactions were initiated by the addition of 50 µM

ATP containing 50 μCi [γ - ^{32}P] or [γ - ^{33}P]. Samples (20 μl) were removed at time intervals between 0 and 60 min, and reactions stopped by the addition of 6 μl 4x SDS-PAGE sample loading buffer. Samples were incubated at 37°C for 30 min prior to SDS-PAGE analysis containing 4% and 12% stacking and resolving gels respectively (Table 2.12). The gels were washed briefly in deionised water and dried under vacuum at 70°C. Labelled protein bands were visualized by autoradiography using BioMax film (Kodak) exposed for 1-7 days, and quantified using a Fuji BAS 1000 phospho-imaging analyzer system (Fujifilm Co., Japan) and AIDA 2D analytical software (Raytest).

2.5.5.2. Assays using purified sensor kinases. Purified histidine kinases were buffer-exchanged in storage buffer containing 10 mM Tris-HCl pH 8.0, 25% glycerol and 2 mM β -mercaptoethanol. This protein (800 pmoles) was incubated in reaction buffer containing 50 mM Tris-HCl pH 7.6, 10 mM MgCl_2 and 50 mM KCl. Total reaction volumes of 150 μl were prepared and the reaction initiated by addition of 50 μM ATP containing 37.5 μCi [γ - ^{32}P] or [γ - ^{33}P]. Samples (15 μl) were removed at time intervals between 0 and 60 min and reactions stopped by the addition of 5 μl 4x SDS-PAGE sample loading buffer. All samples were analysed as described in Section 2.5.5.1. (Note: HEPES-Na was used instead of Tris-HCl for sensor kinase EF1051).

2.5.5.3. Signal testing. For signal screening, 80 pmoles of proteins were tested in the presence of individual candidate signal molecules per 15 μl reaction buffer as described (Section 2.5.5.2.). Protein and signal were incubated at 24 °C for 20 minutes prior to the addition of radiolabelled ATP to initiate the autophosphorylation reactions which proceeded for 40 or 60 minutes. Stock solutions of synthetic GBAP, which is identical in structure to natural GBAP, were prepared in 20% (w/v) acetonitrile, stored at -20°C as described previously (Nakayama *et al.*, 2001).

Stock solutions	Resolving gel (12%)	Stacking gel (4%)
30% acrylamide	3 ml	0.66 ml
1.5 M TrisHCl (pH 8.8)	1.9 ml	-
1.5 M TrisHCl (pH 6.8)	-	0.36 ml
10 % SDS	0.075 ml	0.05 ml
deionised H ₂ O	2.5 ml	3.61 ml
10% ammonium persulphate (APS)	0.05 ml	0.05 ml
TEMED	0.003 ml	0.005 ml

Table 2.12 Composition of mini-gels (12% resolving and 4% stacking gels).

2.6 Biophysical techniques

2.6.1. Circular dichroism studies of secondary structure of purified membrane proteins

Purified proteins were buffer exchanged into 20 mM sodium or potassium phosphate buffer pH 7.6 with 0.05% DDM prior to analysis of secondary structure using Jasco J-715 spectropolarimeter. Purified protein (30-40 μ g) or 0.05 mg/ml suspended in 300 μ l Na / K phosphate buffer was placed into a Far UV cuvette. CD spectral analysis was performed at 18°C with the following parameters: 1nm step resolution; scan rate at 10 nm/min; response time at 1 sec; sensitivity at 50 mdeg; bandwidth at 1.0 nm; and accumulation of ten scans. A sample containing the buffer was measured and subtracted from the protein sample. An alternative parameter was also used for some proteins and indicated otherwise.

2.6.2. Electrospray ionisation mass spectrometry

Purified membrane proteins (50 μ l at ~5mg/ml) were prepared for electrospray ionisation (ESI) mass spectrometry by mixing with 150 μ l of analytical grade methanol, followed by addition of 50 μ l of chloroform and mixed to form a single phase, then 150 μ l of water was added and mixed vigorously to initiate phase separation. After centrifugation at 10,000 g for 2 min to precipitate the protein, the bulk of the aqueous methanol phase was removed and 100 μ l methanol was added. Following a second identical centrifugation step and removal of the solution the pellet of precipitated protein was dried under a flow of nitrogen. The protein was dissolved in 50 μ l 100% formic acid and immediately subjected to size-exclusion chromatography using SephadexTM LH-20 to complete the preparation of membrane proteins, as described previously (Venter *et al.*, 2002). Ionisation of the sample was by nano-ESI using a nanomate (Advion Bioscience Inc., Ithaca, N.Y.), and mass spectrometry by LCT Premier (Waters, U.K. Ltd., Manchester, U.K.) with a single time-of-flight analyser.

2.6.3. Fluorimetry studies of purified proteins

Steady-state fluorescence measurements were performed by Photon Technology International QM-1 spectrophotometer (PTI, U.K.) with LPS 200B lamp. Reaction was placed in Hellma[®] precision cells made of Quartz Suprasil[®] with 10 mm path-length for measurements. Acquisition parameters prior to scan: excitation at 285 nm; emission at 300-400 nm; length at 100 nm; step size at 1 nm; integration at 1 sec; and average at 1.

2.6.3.1. Titration by substrate. Purified proteins were buffer-exchanged into 10 mM Tris-HCl pH 7.6 and 0.05% DDM. A stock of 100 mM substrate was made in 100% DMSO for titration. Benzylhydantoin (0-20 μ l) was titrated into a reaction volume of 980 μ l containing 140 μ g protein, 10 mM Tris-HCl pH 7.6, 135 mM choline Cl, 0.05% DDM and plus / minus 15 mM NaCl, shown in Table 2.13. Note: potassium phosphate buffer was sometimes used instead of Tris-HCl and is indicated otherwise.

Volume additions of 100 mM substrate (μ l)	Final [substrate] (mM)	Final DMSO (%)
0.0	0.00	0.00
1.0	0.10	0.10
1.0	0.20	0.20
3.0	0.50	0.50
2.5	0.75	0.75
2.5	1.00	1.00
5.0	1.50	1.50
5.0	2.00	2.00

Table 2.13 Titration by substrate. Proteins were incubated at 18°C throughout the experiment, and were left stirring for 4 minutes prior to each volume addition. The final substrate and DMSO concentrations, with each volume of substrate added, are indicated in the table.

2.6.3.2. Titration by salt. Purified proteins were buffer-exchanged into 10 mM potassium phosphate pH 7.6 and 0.05% DDM. Stocks of 100 mM salt was made in deionised water. Salt (0-50 μ l) was titrated into a reaction volume of 950 μ l containing 140 μ g protein, 10 mM potassium phosphate pH 7.6 and 0.05% DDM. The final concentration of salt for each volume addition is shown in Table 2.14.

Volume additions (μ l)	Final [Salt] (mM)
0.0	0.00
0.25	0.025
0.25	0.05
0.5	0.10
1.5	0.25
2.5	0.50
5.0	1.00
15.0	2.50
25.0	5.00

Table 2.14 Titration by salts. Proteins were incubated at 18°C throughout the experiment, and were left stirring for 4 minutes prior to each volume addition. The final salt concentrations, with each volume of salt added, are indicated in the table.

2.6.3.3. Titration by HgCl₂. Purified proteins were buffer-exchanged into 10 mM Tris-HCl pH 7.6 and 0.05% DDM. Stocks of 10 mM and 100 mM HgCl₂ were made in deionised water. HgCl₂ (0-19 µl) was titrated into a reaction volume of 980 µl containing 140 µg protein, 10 mM Tris-HCl pH 7.6, 140 mM choline Cl (or 125 mM choline Cl with 15 mM NaCl and 2mM substrate) and 0.05% DDM. The volume additions of HgCl₂ are shown in Table 2.15.

Volume additions (µl)	Final [HgCl ₂] (mM)
1.0 µl of 10 mM	0.01
2.0 µl of 10 mM	0.03
2.0 µl of 10 mM	0.05
5.0 µl of 10 mM	0.10
1.5 µl of 100 mM	0.25
2.5 µl of 100 mM	0.50
5.0 µl of 100 mM	1.00

Table 2.15 Titration by HgCl₂. Proteins were incubated at 18°C throughout the experiment, and were left stirring for 4 minutes prior to each volume addition. The final HgCl₂ concentrations, with each volume added, are indicated in the table.

2.6.3.4. Titration by ATP. Purified proteins were buffer-exchanged into 10 mM Tris-HCl pH 7.6. A stock of 2 mM ATP was made in deionised water. ATP (0-15 µl) was titrated into a reaction volume of 975 µl containing 140 µg protein and 10 mM Tris-HCl pH 7.6. The volume additions of ATP are shown in Table 2.16.

Volume additions of 2 mM ATP (µl)	Final [ATP] (mM)
0.0 µl	0.0
0.5 µl	1.0
0.5 µl	2.0
1.5 µl	5.0
2.5 µl	10.0
5.0 µl	20.0
15.0 µl	50.0

Table 2.16 Titration by ATP. Proteins were incubated at 18°C throughout the experiment, and were left stirring for 4 minutes prior to each volume addition. The final ATP concentrations, with each volume added, are indicated in the table.

2.6.3.5 Titration by GBAP. Purified proteins were buffer-exchanged into 10 mM Tris-HCl pH 7.6. A stock of 1 mM GBAP was made in 3.4% acetonitrile. GBAP (0-5.4 µl) was titrated into a reaction volume of 989.3 µl containing 140 µg protein and 10 mM Tris-HCl pH 7.6. The volume additions of GBAP are shown in Table 2.17.

Volume additions of 1 mM GBAP (μ l)	Final [GBAP] (mM)
0.0 μ l	0.0
0.7 μ l	0.7
0.6 μ l	1.3
1.4 μ l	2.7
2.6 μ l	5.3
5.4 μ l	10.7

Table 2.17 Titration by GBAP. Proteins were incubated at 18°C throughout the experiment, and were left stirring for 4 minutes prior to each volume addition. The final GBAP concentrations, with each volume added, are indicated in the table.

2.6.4. N-terminal sequencing

Proteins were separated by SDS-PAGE (Section 2.4.4) and the gel was rinsed briefly in blotting buffer (10 mM CAPS pH 11, 10% v/v methanol HPLC grade, in Milli-Q water) to remove excess Tris/glycine/SDS which may interfere with sequencing. A stock 100 mM CAPS buffer (pH using NaOH) was made and stored at 4°C, and methanol and water was added shortly before use. FluorotransTM PVDF-membrane was soaked in methanol and equilibrated in the blotting buffer for several minutes. The proteins were transferred from the gel to the membrane by electroblotting at 50 mA for 2 hours in a TE77 semi-dry transfer unit. The membrane was then stained in Coomassie Blue for 10 minutes and destained overnight in 50% (v/v) methanol in H₂O. The stained protein band was excised from the membrane and N-terminal sequencing by Edman degradation (Edman, 1950) was performed by Dr Jeff Keen (University of Leeds).

2.6.5. Removal of N-terminal formyl group by acid-methanol treatment

The N-terminal formyl group of the membrane protein was removed by incubating the protein (0.1 mg) in at least ten volumes of 2M HCl in anhydrous methanol for 48 hours at room temperature with occasional vortexing (Venter, 2001).

2.6.6. Crystallisation trials

Proteins (10 – 15 mg) were purified by Ni-NTA and subjected to size-exclusion chromatography as described in Section 2.5.3.3 to produce homogenous proteins. The pure homogeneous protein fraction was concentrated using Vivaspin 50,000 MWCO to ~15 mg/ml for crystallization trials using the Hampton screening kits (Crystal Screen I and II, MemFracTM and MemStartTM). The vapour diffusion drops were set up in a Corning® Plate 3552 (96 wells) by Douglas Instruments Oryx6 crystallisation robot.

Chapter 3

Overexpression, purification and
characterisation of the NCS1 family of
transport proteins

Chapter 3 – Introduction

This chapter is divided into four sections:

Section 3.1 – Cloning, expression and activities of the NCS1 family proteins

Section 3.2 – Characterisation of an *Escherichia coli* cytosine transporter (b0336)

Section 3.3 – Characterisation of a *Bacillus subtilis* allantoin transporter (Bsu36470)

Section 3.4 – Characterisation of a *Pseudomonas aeruginosa* uracil transporter (PA0443)

Section 3.1 describes the pipeline for the study of the NCS1 family of transporters; it includes an *in silico* analysis of the target proteins, followed by cloning, expression and substrate identification. Subsequent sections, 3.2 to 3.4, describe the characterisation in more detail of three NCS1 proteins; these are the cytosine transporter of *E. coli* (b0336), the allantoin transporter of *B. subtilis* (Bsu36470), and the uracil transporter of *P. aeruginosa* (PA0443).

3.1. Cloning, expression and activities of the NCS1 family proteins

3.1.1. Introduction – A genomic approach

Recently genomics approaches have been developed to overcome the lack of knowledge of membrane protein structural biology (Dobrovetsky *et al.*, 2005; Surade *et al.*, 2006; Carpenter *et al.*, 2008). The idea is to study a number of targets that are homologues of the targeted protein, in the hope that at least a few will produce diffracting crystals, knowing that some will fall out during the following stages.

In this project a pipeline involving some innovative methods developed in Professor Henderson's laboratory was exploited to investigate the NCS1 family of proteins (Figure 3.1). Firstly, a pre-cloning stage involved a search for Mhp1 homologues that belong to the NCS1 family and an *in silico* analysis of the selected homologues. Secondly, the homologues were cloned into an expression vector and their expression in *E. coli* was examined. Successfully overexpressed candidates were then progressed to purification, characterisation and crystallisation trials.

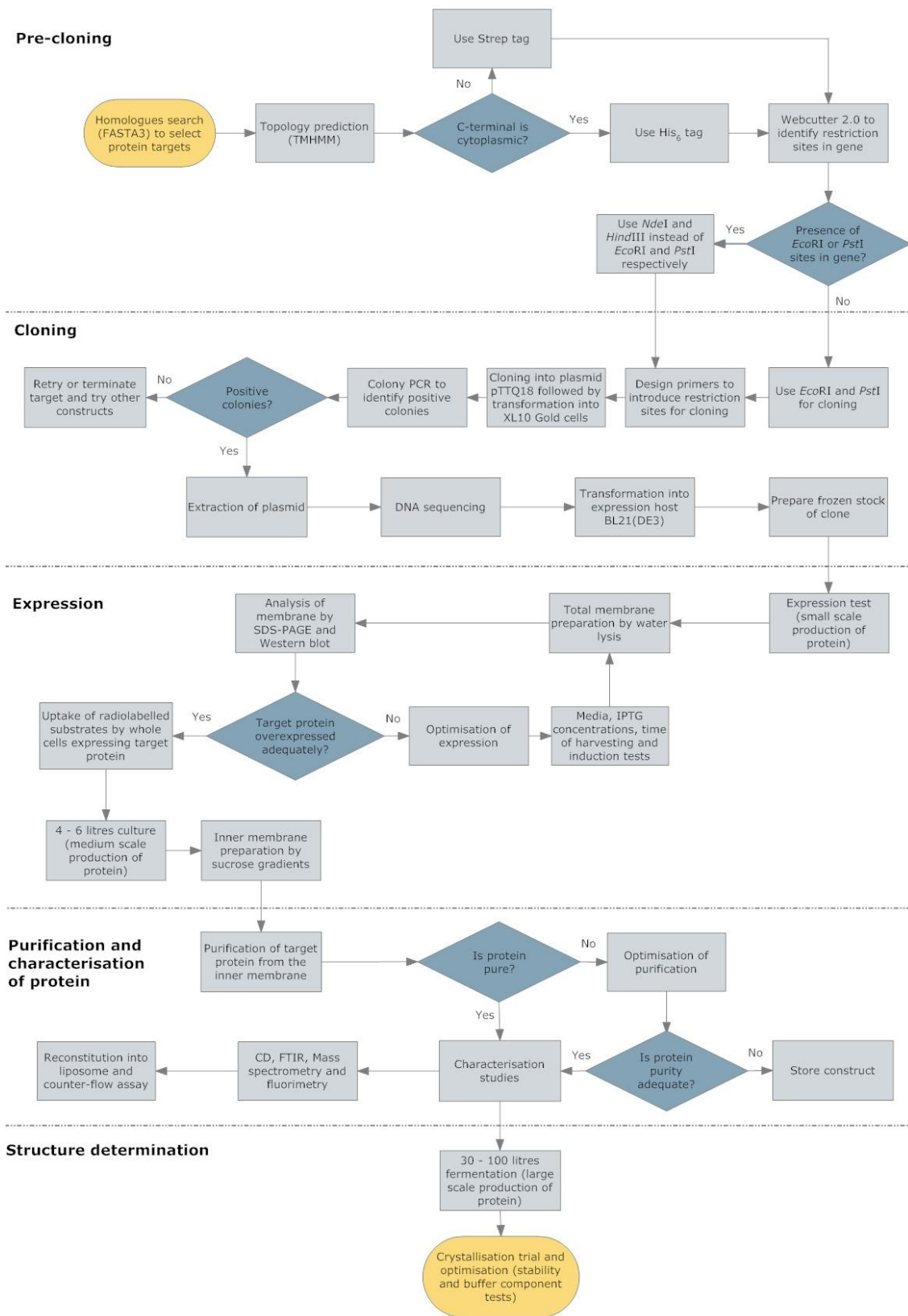





Figure 3.1 Pipeline of genomic approach for the study of NCS1 family proteins.
 Key: Start/end (), process () and decision ().

3.1.2. An homology search and in silico analysis of of Mhp1 homologues

A search for Mhp1 homologues belonging to the NCS1 family using FASTA3 (Pearson, 2000) was performed, and twenty proteins from various bacteria were chosen as targets based on the availability of the genomic DNA in the laboratory and their similarity to the Mhp1 protein. The selected Mhp1 homologues (Table 3.1) have 17-31% identity and 52-65% similarity to Mhp1 (Table 3.1).

An affinity tag, for identification and purification purposes, was introduced at the C-terminal end of the protein since alteration of the N-terminus may be inimical to expression. Depending on the position of the C-terminus of the protein, cytoplasmic or periplasmic, a His₆-tag or Strep-tag was used, respectively.

The topology of a membrane protein follows the 'positive inside' rule (von Heijne, 1986) and predictions using TMHMM (Krogh *et al.*, 2001) showed the selected proteins generally have a putative cytoplasmic C-terminus, apart from the PA0438 and Pden0678 proteins, which were predicted to be outside and have eleven transmembrane regions (Appendix 1). The TMHMM diagram clearly suggests there could be a twelfth transmembrane region in PA0438 and this will then make the C-terminus cytoplasmic (Appendix 1) and therefore a His₆-tag was introduced. The Pden0678 having a C-terminus that is periplasmic suggests a strep-tag should be used. The topology predictions by TMHMM are summarised in Table 3.2. The majority of the selected proteins have twelve putative transmembrane helices and their N-terminus and C-terminus are predicted to be located on the cytoplasmic side of the membrane (Table 3.2), which is a typical feature of the NCS1 as well as the MFS family transporters.

Comparisons of the selected NCS1 proteins against Mhp1 showed some conserved sequences in the family (Appendix 2). The phylogenetic tree of these selected proteins with the previously studied NCS1 proteins (Table 1.1) are shown in Figure 3.2, and showed HyuP of *Arthrobacter aureescens* clustered most closely to Mhp1. The *S. cerevisiae* Dal4, Fur4 and Fui1 proteins clustered closely together (Figure 3.2); they transport allantoin, uracil and uridine, respectively. The *S. cerevisiae* Thi10 and Nrt1 are on the same sub-branch and also clustered closely together, transporting thiamine and nicotinamide riboside, respectively (Figure 3.2). Other previously studied NCS1 proteins that clustered closely together include the *S. cerevisiae* Fcy2, *S. cerevisiae* Tpn1, *C. albicans* Fcy21 and *A. nidulans* FcyB proteins; which transport cytosine, vitamin B₆, purine and cytosine-purine respectively (Figure

3.2). Here, the *P. putida* CytX and *L. plantarum* lp3374 proteins do not cluster close to any of the NCS1 proteins in the phylogenetic tree (Figure 3.2).

	Protein name (bacteria)	Size of protein (amino acids) Accession no.	Putative/known substrate(s)	Identity to Mhp1 (%)	Similarity to Mhp1 (%)
0	Mhp1 (<i>M. liquefaeciens</i>)	489	hydantoins	100.0	100.0
1	b0336 (<i>E. coli</i>)	419 P0AA82	cytosine	22.8	58.1
2	b0511 (<i>E. coli</i>)	484 P75712	allantoin, uracil	26.1	58.7
3	Bsu36470 (<i>B. subtilis</i>)	490 P94575	allantoin, uracil	27.3	63.3
4	Bsu3867 (<i>B. subtilis</i>)	457 P94369	cytosine/purine	16.6	52.4
5	EF3000 (<i>E. faecalis</i>)	498 Q82ZQ0	allantoin nucleobase	25.8	62.2
6	lp3374 (<i>L. plantarum</i>)	457 Q88SN4	cytosine/purines, uracil, thiamine, allantoin	19.3	57.9
7	PA0438 (<i>P. aeruginosa</i>)	416 Q9I679	cytosine	24.1	57.5
8	PA0443 (<i>P. aeruginosa</i>)	496 Q9I674	cytosine/purines, uracil, thiamine, allantoin	29.3	59.8
9	PA0476 (<i>P. aeruginosa</i>)	575 Q9I642	cytosine/purines, uracil, thiamine, allantoin	26.4	61.8
10	PA2073 (<i>P. aeruginosa</i>)	476 Q9I242	cytosine/purines, uracil, thiamine, allantoin	23.1	54.4
11	PA5099 (<i>P. aeruginosa</i>)	480 Q9HU84	cytosine/purines, uracil, thiamine, allantoin	22.2	52.2
12	Pden0678 (<i>P. denitrificans</i>)	450 Q9I5P7	cytosine/purines, uracil, thiamine, allantoin	25.3	65.0
13	Pden4351 (<i>P. denitrificans</i>)	490 A1BA71	cytosine/purines, uracil, thiamine, allantoin	30.9	65.3
14	Pden1111 <i>P. denitrificans</i>	494 A1B125	nucleobase	28.6	60.6
15	PPA0619 (<i>P. acnes</i>)	487 A1AZN7	cytosine/purine	22.2	54.2
16	PPA2211 (<i>P. acnes</i>)	495 Q6A5P8	cytosine/purine	18.6	53.0
17	SCO0572 (<i>S. coelicolor</i>)	466 Q93RZ8	cytosine	23.3	57.1
18	SCO5579 (<i>S. coelicolor</i>)	485 Q9ZBQ0	cytosine/purine	24.6	57.6
19	SCO6417 (<i>S. coelicolor</i>)	522 O69811	allantoin	31.0	62.3
20	SCO7500 (<i>S. coelicolor</i>)	495 Q93J97	allantoin	27.9	61.6

Table 3.1 Proteins of the NCS1 family studied in this project. The size, putative/known substrates, identity and similarity to Mhp1, of the target proteins are shown. Similarity and identity values were calculated by FASTA3 (Pearson, 2000).

	Protein name	Predicted position of:		No. of predicted transmembrane regions
		N-terminus	C-terminus	
0	Mhp1	inside	inside	12
1	b0336 (CodB)	inside	inside	12
2	b0511	inside	inside	12
3	Bsu36470 (PucI)	inside	inside	12
4	Bsu3867	inside	inside	12
5	EF3000	inside	inside	12
6	Ip3374	inside	inside	12
7	PA0438	inside	outside/inside	11/12
8	PA0443	inside	inside	12
9	PA0476	outside	inside	11
10	PA2073	inside	inside	12
11	PA5099	outside	inside	13
12	Pden0678	inside	outside	11
13	Pden4351	inside	inside	12
14	Pden1111	inside	inside	12
15	PPA0619	inside	inside	12
16	PPA2211	inside	inside	12
17	SCO0572	inside	inside	12
18	SCO5579	inside	inside	12
19	SCO6417	inside	inside	12
20	SCO7500	inside	inside	12

Table 3.2 Summary of the topology predictions determined by TMHMM (Krogh *et al.*, 2001) for the selected NCS1 proteins.

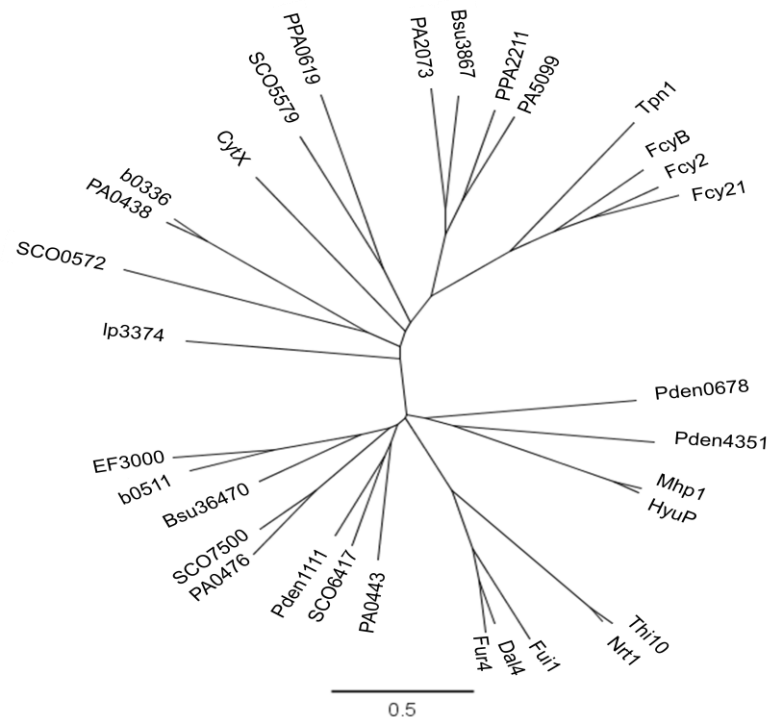


Figure 3.2 Phylogenetic tree of Mhp1, its homologues and previously studied NCS1 proteins. The phylogenetic tree was created by Geneious Pro 4.6.5 software using amino acid sequences of the NCS1 proteins.

3.1.3. Cloning of genes encoding NCS1 proteins

The genes of interest were amplified by PCR (Figure 3.3). The amplified products after restriction digestion were ligated into plasmid pTTQ18-His₆ (Section 2.3.3. to 2.3.6.) using the strategy described in Section 2.3.1.

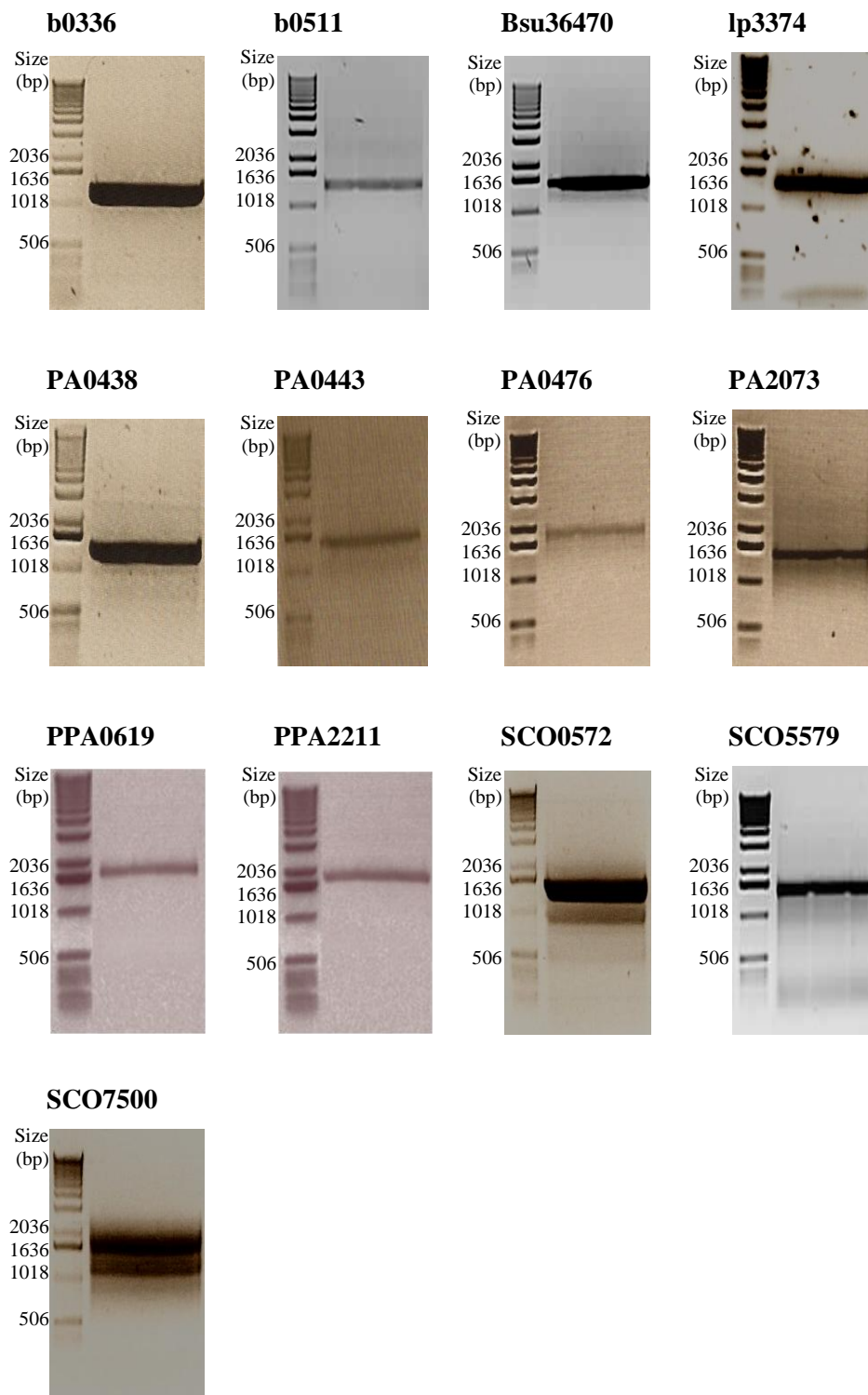


Figure 3.3 NCS1 family genes amplified by PCR. Samples after PCR were analysed on a 1% agarose gel. Lane: 1kb DNA ladder (left) and PCR products (right). Note: results for EF3000, Pden1111 and Pden4351 are not shown.

Out of the twenty NCS1 targets, sixteen are currently amplified by PCR (Figure 3.3). The EF3000 construct was made by Dr Kim Bettaney (Bettaney, 2008), and the Pden1111 and Pden4351 constructs were made by Dong Leng (unpublished). The cloning results for them are therefore not shown in this thesis. Genes from organisms with high GC content such as the *Pseudomonas* and *Streptomyces* were more difficult to amplify and addition of DMSO with lowered annealing temperatures were required to overcome this problem (Section 2.3.2).

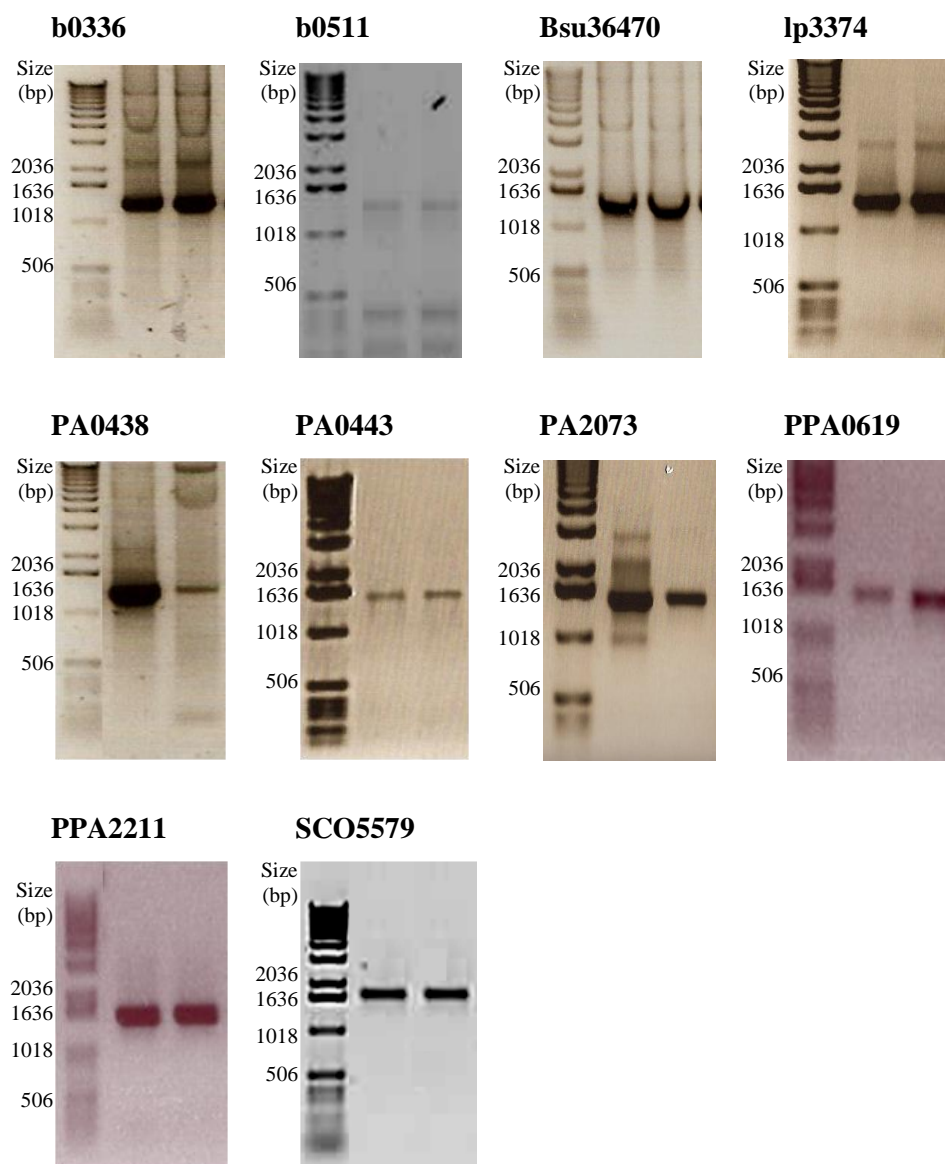


Figure 3.4 PCR screening for positive colonies. Colonies obtained from ligation and transformation were subjected to PCR screening to identify colonies containing the correct plasmids (Section 2.3.9). Samples after PCR were analysed on a 1% agarose gel. Lane: 1kb DNA ladder (left) and colony PCR products (middle and right) – only two out of six screened colonies are shown. Note: results for EF3000, Pden1111 and Pden4351 are not shown.

Out of the sixteen amplified genes, thirteen were successfully cloned into plasmid pTTQ18-His₆, and confirmed by PCR screening (Figure 3.4), restriction digestion analysis (Figure 3.5) and DNA sequencing (Appendix 3) as described in Sections 2.3.9 to 2.3.12. The recombinant clones were each transformed into *E. coli* BL21(DE3) cells (Section 2.3.13) for expression trials (Section 2.4.1).

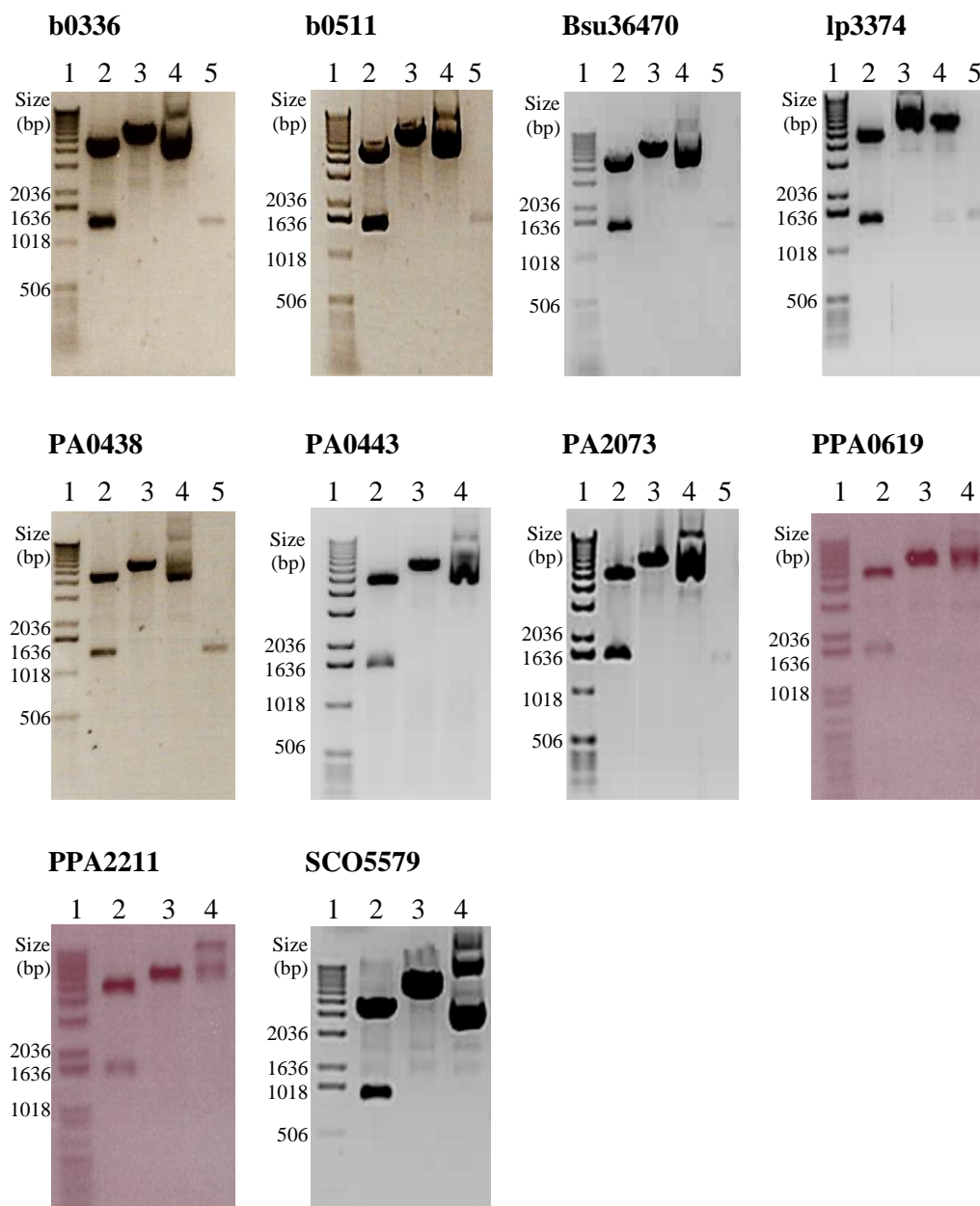


Figure 3.5 Restriction digestion analysis to confirm the correct size gene in plasmid pTTQ18-His₆. XL-10 Gold cells harbouring the putative clone were grown overnight as described in Section 2.3.9. The plasmids from the overnight culture were extracted and subjected to restriction digestion for 3 hours. Samples after digestion were analysed on a 1% agarose gel. Lane (1) 1kb DNA ladder. Plasmids harbouring the putative gene were digested as follows: Lane (2) *EcoRI* and *PstI*; (3) *EcoRI*; (4) undigested. Lane (5) *EcoRI* and *PstI* digested PCR product. Note: results for EF3000, Pden1111 and Pden4351 are not shown.

The genes for PA0476, SCO0572 and SCO7500 were successfully amplified by PCR (Figure 3.3), but their cloning into plasmid pTTQ18-His₆ is not yet completed. Therefore seven genes are still at the cloning stage, and thirteen of the twenty selected NCS1 proteins are currently successfully cloned.

3.1.4. Expression of the NCS1 proteins

The native expression level of any particular membrane protein is typically less than 0.1% of total cell protein and consequently amplified expression is required for production of milligram quantities of proteins. The selected NCS1 genes from various bacteria were therefore cloned into plasmid pTTQ18-His₆ (Section 2.2.3), under control of the *tac* promoter, which allows successful overexpression of many membrane proteins (Ward *et al.*, 2000; Ward *et al.*, 2001; Potter *et al.*, 2002; Saidijam *et al.*, 2003; Liang *et al.*, 2005; Saidijam *et al.*, 2005; Saidijam *et al.*, 2006; Suzuki & Henderson, 2006; Surade *et al.*, 2006). *E. coli* was chosen as the expression host since it is inexpensive, easy to handle and has proven successful for producing prokaryotic membrane proteins (Hockney, 1994; Grisshammer & Tate, 1995; Ward *et al.*, 2000; Wang *et al.*, 2003). Also, most of the structures of bacterial membrane proteins that are present in the Protein Data Bank (Berman *et al.*, 2000) were successfully expressed in *E. coli*.

Of the target list, the EF3000, Pden1111 and Pden4351 genes were previously shown to express in *E. coli* host cells (Bettaney, 2008; Leng, unpublished). Their expression in *E. coli* membranes is repeated in this project for comparison with the other NCS1 proteins (Figure 3.6). All the NCS1 proteins currently cloned in this project contained a His₆-tag, because their C-termini were predicted to be inside the inner membrane (Table 3.2 & Appendix 1). The amino acid sequences of the resulting expressed recombinant proteins are shown in Appendix 4.

Twelve out of thirteen cloned NCS1 proteins were successfully expressed in the *E. coli* membrane (Figure 3.6). Only PPA0619 was not detectable in the Western blotting using India-HisProbeTM (Figure 3.6). The selected NCS1 proteins were expressed at 3-24% of the total membrane, determined by scanning densitometry (Table 3.3). The Western blot to the His₆-tag (Figure 3.6) confirmed the insertion of the genes into plasmid pTTQ18-His₆, and that the C-termini of the cloned proteins were intact. No protein degradation was generally detected on the Western blot, though possibly a small amount of degraded or truncated SCO5579 protein was observed (Figure 3.6).

Proteins that were intensely overexpressed, such as the PA0443, PA2073 and SCO5579, tended to have an upper band on the Western blot which could indicate more than one form of the protein in the gel (Figure 3.6), perhaps because of the incomplete unfolding of the protein in SDS. The theoretical mass of each His₆-tagged protein was calculated using the Compute pI/Mw tool (Gasteiger *et al.*, 2005) and compared with the size of the protein on the SDS-PAGE gel determined from a log Mr-distance curve (Section 2.4.4). The NCS1 family proteins migrate at approximately 62-74% of their predicted sizes on SDS-PAGE (Table 3.5). It is typical for membrane proteins to migrate anomalously on a SDS-PAGE due to their hydrophobic nature, high binding of SDS or the retention of their secondary structure facilitating the migration through the gel (Ward *et al.*, 2000).

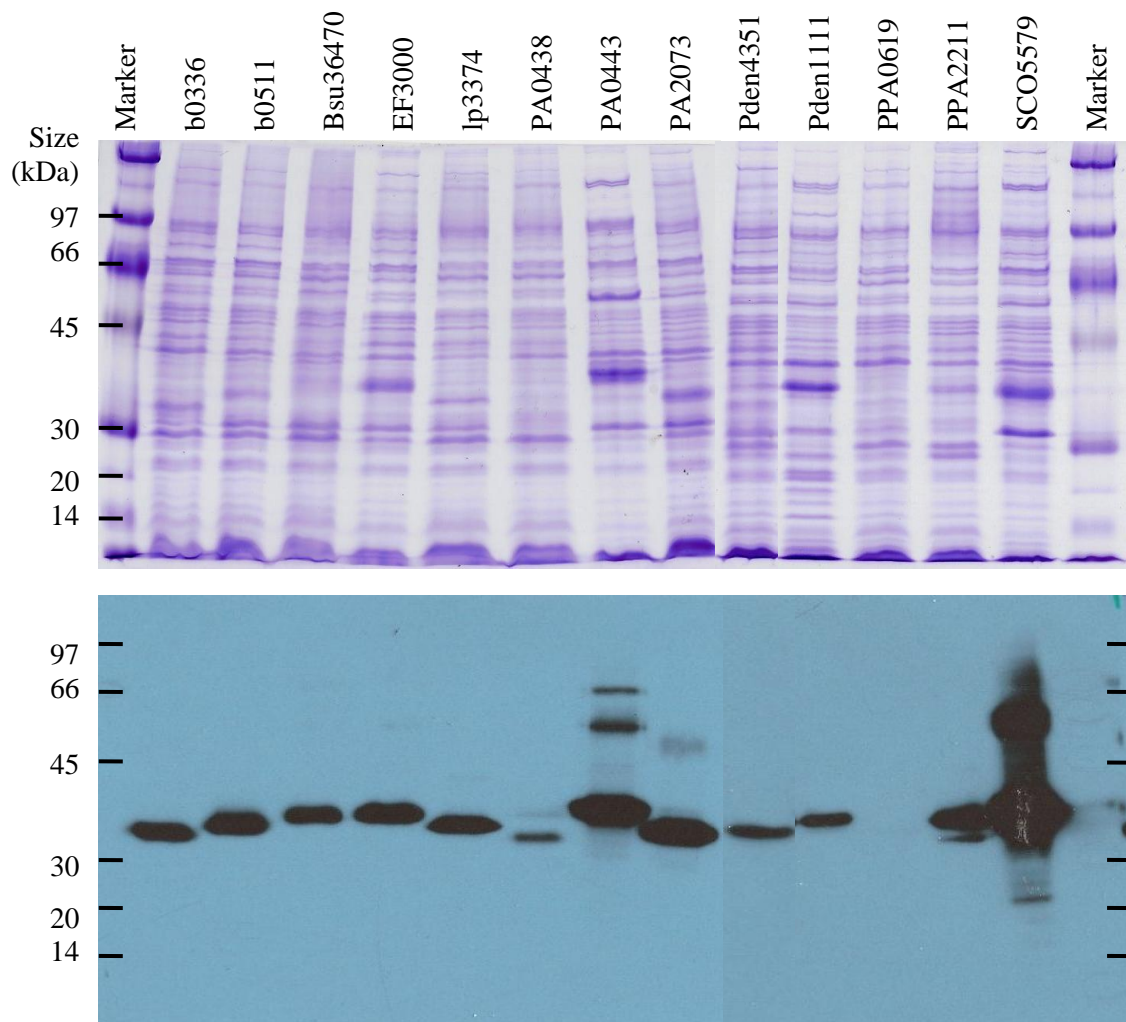


Figure 3.6 Expression of thirteen NCS1 genes in *E. coli*. The NCS1 genes from various bacteria were cloned into plasmid pTTQ18-His₆ and expressed in *E. coli* BL21(DE3) cells under the control of *tac* promoter. The cells were cultured in LB medium as described in Section 2.4.1. Cells were induced with 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes were prepared by water lysis and samples (15 µg) were analysed by SDS-PAGE (top) and Western blotting (bottom).

Protein name	Predicted size (kDa) ¹	Size on SDS-PAGE (kDa) ²	Difference in size (%)	Expression in total <i>E.coli</i> membrane (%) ³
b0336	44.9	33	73	9
b0511	54.3	35	64	7
Bsu36470	55.8	37	66	5
EF3000	56.0	39	70	16
lp3374	51.6	34	66	6
PA0438	44.3	31	70	3
PA0443	55.5	41	74	24
PA2073	53.2	35	66	17
Pden4351	54.9	34	62	7
Pden1111	55.0	40	73	20
PPA0619	53.7	-	-	-
PPA2211	54.7	40	73	8
SCO5579	52.6	39	74	23

Table 3.3 Expression level, predicted and observed size of the NCS1 transporters. ¹Predicted using Compute pI/MW tool (Gasteiger *et al.*, 2005); ²calculated using a log Mr-distance curve; ³determined by scanning densitometry of a Coomassie Blue stained SDS-PAGE in Figure 3.6.

3.1.5. Analysis of the codon usage in the NCS1 proteins

The fraction of usage of each codon in the protein was computed and plotted against the fraction of usage of the codon in the *E. coli* expression host (relative adaptiveness), and a mean difference value (an average of the relative adaptiveness) of codon usage calculated (Appendix 5). The analyses were computed and plotted by ‘Graphical codon usage analyser’ (Fuhrmann *et al.*, 2004).

A comparison of the codon usage of the NCS1 family proteins with the *E. coli* expression host showed a mean difference of 20.38 to 43.88% (Appendix 5). The codon in PA0438 has the highest mean difference to the *E. coli* expression host, and b0511 has the lowest mean difference (Appendix 5). These proteins were expressed at 3 and 7 %, respectively (Figure 3.6 and Table 3.3), which were low compared to the other NCS1 proteins (Table 3.3). The PPA0619 protein that did not express (Figure 3.6), has a calculated mean difference of 29.70% which was in the middle of the range for the NCS1 proteins (Appendix 5). These results suggest that differences in the codon usages of the NCS1 proteins from the *E.coli* expression host do not necessarily affect its expression level (Appendix 5).

An analysis of the number of *E. coli* rare codons in each of the NCS1 proteins showed they have between 11-56 rare codons (Appendix 6). PPA0619 has the highest number of rare codons, and the second-highest consecutive rare codons (Appendix 6). There are also three consecutive rare codons of the same bases (ACG,ACG,ACG) in PPA0619, which was not found in the other NCS1 proteins (Appendix 6), that could contribute to this protein not being expressed in the *E. coli* cells (Figure 3.6). The PA0443 protein has 28 rare codons (one of the least) and no consecutive rare codons, and was expressed at 24% in the *E. coli* membrane, which was the highest amongst all the NCS1 proteins tested (Appendix 6). The results indicate the number of rare codons and consecutive rare codons could affect the expression level of the NCS1 proteins.

3.1.6. Identification of substrates for the NCS1 proteins

The substrates transported by most of the selected NCS1 proteins are unknown but they are predicted to transport nucleobases, thiamine and allantoin (Saier *et al.*, 2006) (Figure 3.7). In order to identify the substrates for the selected transporters (Table 3.1), transport assays were performed using radiolabelled substrates and intact *E. coli* cells overexpressing each target protein (Section 2.5.6). Both uninduced and induced cells were tested and the uptake of a particular substrate was identified by higher uptake in the induced cells than in the uninduced; hence the uninduced cells served as a control.

Ten different substrates were chosen in an attempt to identify the substrates for nine of the selected NCS1 proteins (Table 3.4) that had moderate or high overexpression levels. It was previously reported that the CodB and PucI proteins are cytosine and allantoin transporters, respectively (Danielsen *et al.*, 1992; Schultz *et al.*, 2001). These proteins were tested and their transported substrates further confirmed (Table 3.4); allantoin was not the only substrate for the PucI protein, but uracil was also (Table 3.4). Further characterisation of these two proteins is described in Section 3.2 and 3.3.

Sequence alignment by ClustalW (Thompson *et al.*, 2002; Larkin *et al.*, 2007) (Appendix 2) showed that the b0511 and EF3000 proteins were highly similar to the PucI protein, 75.1 and 73.1 % respectively, but surprisingly, allantoin or uracil were not taken up by these proteins (Table 3.4).

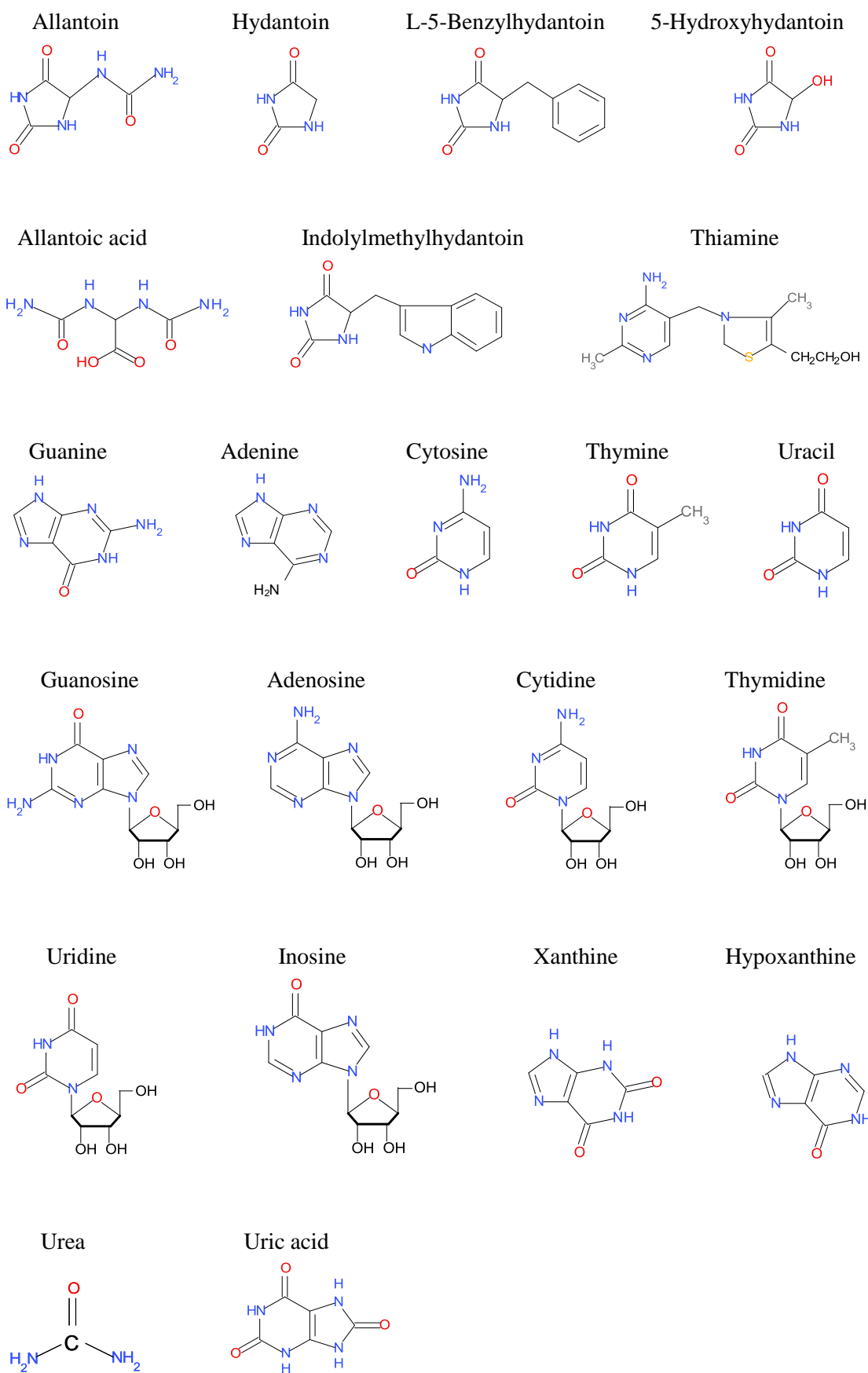


Figure 3.7 Chemical structures of possible substrates for the NCS1 family of transporters. The structures were drawn using Symyx Draw 3.2.

	Cell	adenine	guanine	cytosine	uracil	adenosine	thymidine	uridine	allantoin	thiamine	5-benzyl-hydantoin
b0336 (CodB)	U	N/T	N/T	0.130 0.138 0.144	N/T	N/T	N/T	N/T	N/T	N/T	N/T
	I			0.800 0.859 0.930							
b0511	U	0.626 0.671 0.698	N/T	0.086 0.091 0.113	0.335 0.324 -	N/T	N/T	2.882 3.886 4.369	0.107 0.213 -	N/T	3.349 3.622 3.641
	I	0.869 0.872 0.900		0.090 0.097 0.108	0.367 0.410 -			2.966 3.139 3.146	0.064 0.147 -		3.826 4.031 4.628
Bsu36470 (Pucl)	U	1.124 1.281 1.339	N/T	0.081 0.089 0.122	0.335 0.421 -	N/T	N/T	2.057 2.107 2.522	9.924 9.736 -	N/T	3.120 3.256 -
	I	0.890 0.968 1.056		0.097 0.103 0.104	0.608 0.720 -			1.571 1.807 1.841	43.91 44.16 -		2.632 2.684 -
EF3000	U	1.052 1.096	N/T	N/T	0.335 0.311	N/T	N/T	N/T	0.067 0.050	0.485 0.543	3.636 3.645
	I	1.312 1.394			0.286 0.306				0.063 0.066	0.279 0.493	3.817 4.035
PA0443	U	1.229 1.231 1.239	N/T	0.273 0.274 0.292	0.253 0.264 -	1.565 1.599 1.679	0.312 0.322 0.339	2.328 2.882 2.959	0.062 0.094 -	N/T	2.777 3.269 3.339
	I	1.064 1.126 1.157		0.170 0.236 0.236	0.841 0.886 -	1.638 1.706 1.718	0.318 0.357 0.358	1.967 2.114 2.903	0.041 0.072 -		3.191 3.227 3.285
PA2073	U	1.237 1.239 1.248	0.270 0.277 0.420	0.241 0.268 0.269	0.098 0.121 0.131	N/T	N/T	N/T	0.095 0.097 -	0.144 0.145 0.359	2.730 2.808 3.141
	I	0.807 0.820 0.835	0.198 0.198 0.204	0.136 0.140 0.143	0.900 0.104 0.114				0.098 0.103 -	0.259 0.271 0.362	2.788 3.204 3.206
Pden1111	U	0.762 0.908	N/T	0.110 0.116	0.304 0.309	1.372 1.382	0.264 0.272	1.067 1.150	0.026 0.039	0.394 0.414	2.462 2.606
	I	1.512 1.715		0.099 0.104	0.648 0.736	1.684 1.892	0.428 0.477	1.794 1.791	0.032 0.063	0.379 0.382	2.652 2.656
Pden4351	U	N/A	N/T	0.137 0.139	0.271 0.297	N/T	N/T	N/T	0.049 0.089	N/T	2.784 3.027
	I			0.201 0.212	0.741 0.764				0.070 0.110		3.037 3.449
SCO5579	U	1.114 1.289 1.340	N/T	0.141 0.162 0.177	0.285 0.321 -	N/T	0.262 0.268 0.269	1.757 1.902 2.290	0.055 0.076 -	N/T	2.320 2.330 3.063
	I	0.858 0.880 0.920		0.090 0.100 0.100	0.329 0.357 -		0.212 0.213 0.223	1.764 1.779 1.780	0.057 0.063 -		1.645 2.022 2.596

Table 3.4 Substrate screening of the NCS1 family transporters. Transport assays were performed as described in Section 2.5.6. Cells were cultured in minimal media supplemented with 100 µg/ml carbenicillin and 20 mM glycerol, and induced with 0.5 mM IPTG when A_{680nm} reached 0.4 – 0.6. Cells were harvested 1 hour post-induction. Uninduced (U) and induced (I) cells harbouring the target clone were tested. The uptake (nmol/mg of cells) two minutes after addition of radiolabelled substrate was examined. Uptake (in green), potential uptake (in amber) and no uptake (in red) by the proteins are indicated in the table. N/T, not yet tested. Criteria for green = $\geq 100\%$ increase I/U. Each value in the table represents a single transport experiment.

The *P. denitrificans* proteins, Pden1111 and Pden4351, took up uracil in the induced cells at 2.3 and 2.7 times higher than in the uninduced cells (Table 3.4). The Pden1111 protein also mediated uptake of adenosine and thymidine, though the extent is not as great when compared to uracil, but it suggested the ability to transport a range of substrates (Table 3.4). The PA0443 protein from *P. aeruginosa* transported uracil at 3.3 times higher in the induced cell and therefore is an uracil transporter (Table 3.4). This protein is further discussed in Section 3.4.

3.1.7. Discussion: cloning, expression and activities of NCS1 proteins

To date, the structures of one NCS1 family protein, Mhp1, has been solved (Weyand *et al.*, 2008; Shimamura *et al.*, 2010). In this genomic approach for studying the NCS1 family transporters, twenty proteins that were homologues of Mhp1 were selected (Table 3.1). *In silico* analysis of these proteins suggested that most of them have twelve transmembrane domains and possess a cytoplasmic N- and C- terminus (Table 3.2 and Appendix 1).

Thirteen of the selected proteins were cloned and a His₆-tag introduced at the C-termini, and another seven are undergoing cloning (Table 3.5). Of the thirteen cloned proteins, twelve were successfully expressed in *E. coli* and their presence was confirmed by Western blotting (Figure 3.6). The NCS1 proteins were expressed at 3-24% of the *E. coli* total membrane, and like other membrane proteins, they migrated anomalously on SDS-PAGE, at 62-74% of their actual mass (Table 3.3).

All but two of the NCS1 proteins are heterologously expressed in *E. coli*. There are many factors that could contribute to the low level of expression of some of the NCS1 proteins. For example, codon usage is important for genes to be expressed efficiently (Grosjean & Fier, 1982), and could be a problem in recombinant expression systems. Analysis of the NCS1 family proteins have shown that their differences in codon usage from that in the expression host does not affect their expression level (Appendix 5), but the number and consecutiveness of rare codons in the protein will affect their expression (Appendix 6).

To overcome this, host strains such as the RosettaTM (Novagen) and CodonPlus-RILTM (Stratagen), which are BL21 derivatives, could be used to provide the rare tRNA species to match the codon usage of the heterologous gene. The modest expression levels of two *E. coli* genes (b0336 and b0511) were 9 and 7 % respectively, lower than some of the heterologously expressed genes (Figure 3.6 and Table 3.3). There are 22

and 24 rare codons in these proteins, respectively; fewer in number than the other NCS1 proteins (Appendix 6). This could suggest that not only codon usage is an important factor for protein expression, but the overexpression of the gene itself might cause toxicity and limit the level of protein being expressed.

	Protein name	Cloned	Expression	Purified (Ni-NTA)	Substrate(s) transported	N-terminal sequence	Circular Dichroism
1	b0336	yes	moderate	yes	cytosine	not tried	yes
2	b0511	yes	moderate	not tried	not identified	not tried	not tried
3	Bsu36470	yes	moderate	yes	allantoin uracil	MNSHMKLKES	Yes
4	Bsu3867	UC	N/A	N/A	N/A	N/A	N/A
5	EF3000	yes	high	yes	not identified	not tried	not tried
6	lp3374	yes	low	not tried	not tried	not tried	not tried
7	PA0438	yes	low	not tried	not tried	not tried	not tried
8	PA0443	yes	high	yes	uracil	MNSHMQQSRS	yes
9	PA0476	UC	N/A	N/A	N/A	N/A	N/A
10	PA2073	yes	high	yes	not identified	MNSQMSNNNE	not tried
11	PA5099	UC	N/A	N/A	N/A	N/A	N/A
12	Pden0678	UC	N/A	N/A	N/A	N/A	N/A
13	Pden4351	yes	moderate	not tried	uracil	not tried	not tried
14	Pden1111	yes	high	yes	uracil	not tried	yes
15	PPA0619	yes	ND	N/A	N/A	N/A	N/A
16	PPA2211	yes	low	not tried	not tried	not tried	not tried
17	SCO0572	UC	N/A	N/A	N/A	N/A	N/A
18	SCO5579	yes	high	yes	not identified	MNSHMSKTAE	not tried
19	SCO6417	UC	N/A	N/A	N/A	N/A	N/A
20	SCO7500	UC	N/A	N/A	N/A	N/A	N/A

Table 3.5 Current status of the selected NCS1 family proteins. UC, under cloning; N/A, not available to test; ND, not detected in Western blotting. The EF3000 protein was cloned and purified by Dr Kim Bettaney. The Pden4351 and Pden1111 proteins were cloned by Mr Dong Leng, and Pden1111 was purified by Mr Shen Wei Lim.

Here, CodB (b0336) and PucI (Bsu36470) proteins were tested for the uptake of their reported substrates (Danielsen *et al.*, 1992; Schultz *et al.*, 2001) and the assays further confirmed that their substrates were cytosine and allantoin, respectively (Table 3.4).

Three uracil transporters were identified; Pden4351, Pden1111 and PA0443 (Table 3.4). PucI also transported uracil (Table 3.4). The substrates transported by some of the NCS1 proteins were not yet identified (Table 3.4). A wider range of substrates should now be tested and possibly the buffer components of the transport assay modified, e.g. by addition of NaCl. Since structure of Mhp1 revealed a sodium binding site (Weyand *et al.*, 2008), the selected NCS1 transporter may potentially rely on the presence of sodium for transport.

Currently, seven of the selected NCS1 proteins are purified from the inner membrane and this allows us to proceed to characterisation studies and crystallisation trials (Table 3.5). Some targets are still at the cloning stage (Table 3.5), and once cloned they will proceed to expression trials and onwards.

3.2. Characterisation of an *Escherichia coli* cytosine transporter – CodB (b0336)

3.2.1. Introduction – the *codBA* operon of *E. coli*

The *E. coli codBA* operon encodes two pyrimidine salvage proteins, the cytosine permease (*codB*) and the cytosine deaminase (*codA*), whose genes overlap, and are located in a cytosine-inducible operon (Danielsen *et al.*, 1992; Qi & Turnbough, 1995) (Figure 3.8). *E. coli* can utilise cytosine as its sole source of nitrogen by using CodB to mediate the uptake of exogenous cytosine, and CodA to catalyse the hydrolytic deamination of cytosine to uracil (a source of pyrimidines) and ammonia (a source of nitrogen) (Danielsen *et al.*, 1992).

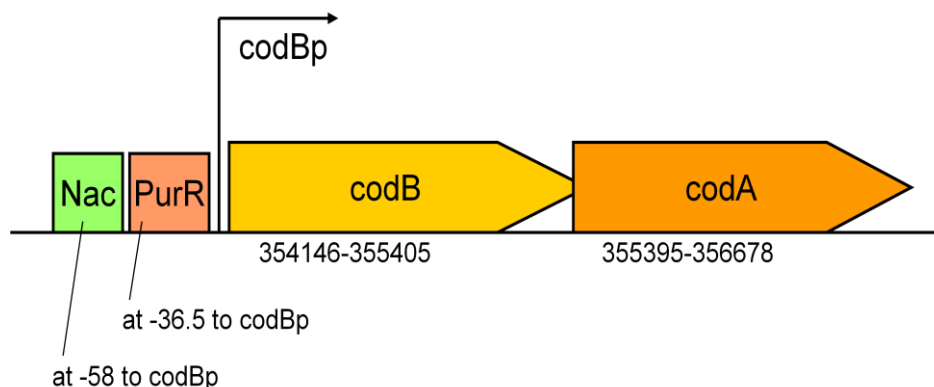


Figure 3.8 Map of the *E. coli codBA* region. The bent arrow denotes the *codB* promoter site. Diagram modified from the *codBA* operon in regulonDB (Gama-Castro *et al.*, 2008).

Transcription of the *codBA* operon of *E. coli* is regulated by nitrogen and purine, with approximately three times more *codBA* expression in cells grown in nitrogen-limiting medium (Muse *et al.*, 2003) and reduced *codBA* expression in cells grown on medium containing hypoxanthine (Danielsen *et al.*, 1992).

Expression of *codBA* is induced under nitrogen-limited growth conditions and repressed by purines or pyrimidines. Induction and repression are mediated by the nitrogen assimilation control (Nac) (Anderson *et al.*, 1989; Muse *et al.*, 1998; Muse *et al.*, 2003) and PurR proteins respectively (Anderson *et al.*, 1989; Kilstrup *et al.*, 1989), and their binding sites are located upstream of the *codBp* promoter (Figure 3.8).

Nitrogen-limited growth leads to a starvation for glutamine, causing a cascade of events, including activation of a transcription regulator, NtrC, a protein kinase that activates RNA polymerase carrying σ^{54} to transcribe a number of genes, one of which is *nac*, which codes for the Nac protein. Nac activates RNA polymerase carrying σ^{70} , which leads to transcription of several operons whose products can supply the cell with ammonia or glutamate from alternative organic sources (Bender, 1991; Muse & Bender, 1998), i.e. *codBA*, *nac*, glutamate dehydrogenase (*gdhA*) (McPherson & Wootton, 1983), glutamate synthase operon (*gltBDF*) (Castaño *et al.*, 1988; Castaño *et al.*, 1992), and *nupC* (Craig *et al.*, 1994).

PurR is a repressor that controls the synthesis of the enzymes of purine biosynthesis (Rolfes & Zalkin, 1988). This regulator binds two products of purine metabolism, hypoxanthine and guanine, to induce the conformational change that allows binding to DNA at the PurR box (Figure 3.8).

Section 3.2 describes the characterisation of the CodB protein. Firstly, the topology and the conserved residues in the protein are examined, followed by the kinetics and substrate specificity studies, and finally purification and circular dichroism.

3.2.2. The *E. coli* cytosine transporter (CodB): topology analysis and conserved residues

In a whole cell transport assay, the cloned *codB* gene was able to complement mutants with cytosine transport defects (Danielsen *et al.*, 1992) suggesting CodB is a cytosine permease. Previous topology analysis using alkaline phosphatase and β -galactosidase fusions suggested CodB has twelve transmembrane segments (Danielsen *et al.*, 1995). Here, a topology model of CodB in the cytoplasmic membrane, predicted using TMHMM (Krogh *et al.*, 2001), also showed twelve transmembrane helices

(Figure 3.9). The N- and C- termini of CodB comprise approximately twenty amino acids, and the lengths of the periplasmic and cytoplasmic loops between the helices is variable (Figure 3.9).

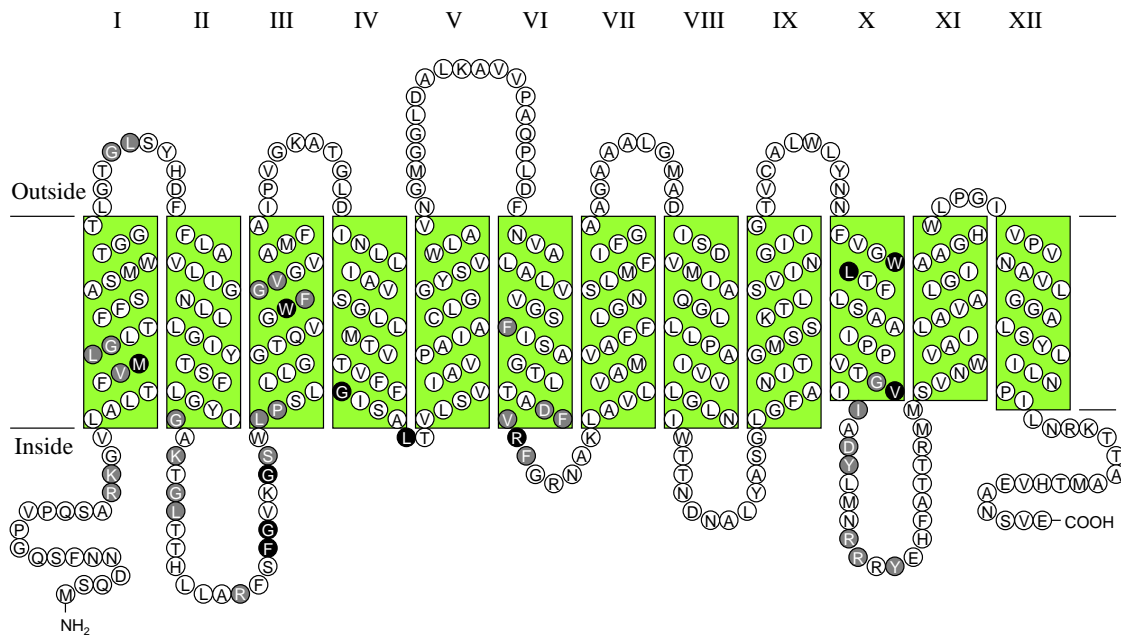


Figure 3.9 Topology model of *E. coli* cytosine permease (CodB). Topology prediction by TMHMM (Krogh *et al.*, 2001). The predicted helices are numbered from the N- to C-terminus and are shown in roman numerals. Conserved residues of the NCS1 family proteins are highlighted in grey and black according to the alignment in Appendix 2.

The conserved residues of CodB are highlighted in the topology model (Figure 3.9). Most of these are identical or highly similar within the NCS1 family proteins (Appendix 2). How some of these residues may contribute to its substrate recognition and binding are discussed in Section 3.2.9.

A search through the databases using FASTA3 showed the amino acid sequence of CodB is highly conserved throughout the *E. coli* strains, 99.3-100% identity; the *Salmonella* species, 79.9-93.2% identity (95.7-96.2% similarity); and the *Vibrio* species, 73.7-78.0% identical (93.0-94.8% similar). CodB is 76.5 % identical (92.6% similar) to PA0438, and 27.4% identical (65.5% similar) to SCO0572, and clustered very closely with these proteins, shown by the phylogenetic tree (Figure 3.2). This could suggest that these selected proteins are cytosine transporters.

3.2.3. Cloning and expression of CodB

To characterise the *E. coli* cytosine permease, the *codB* gene (b0336) was amplified from the genomic DNA of *E. coli* K-12 strain by PCR (Figure 3.3), and cloned into plasmid pTTQ18-His₆ with a His₆-tag added to the C-terminus of the protein. The clone was subjected to analysis by PCR screening and restriction digestion, which confirmed the correct sized gene in the plasmid construct (Figure 3.4 and 3.5). DNA sequencing also confirmed a single silent mutation in the cloned *codB* gene but confirmed the correct encoded amino acid sequence (Appendix 3 and 4). This was followed by transformation of the construct into *E. coli* BL21(DE3) cells for expression test, as described in Section 2.3 and 2.4. The CodB protein was overexpressed to 9% in the *E. coli* total membrane (Figure 3.6 and Table 3.3) and the growth curves of the uninduced and induced cells harbouring the pTTQ18(*codB*)-His₆ are shown in Figure 3.10. The induced culture continued to grow after addition of IPTG but was slower than the uninduced cells (Figure 3.10). The SDS-PAGE gel and Western blot confirmed the presence of the CodB protein in the induced cell, and not in the uninduced cell (Figure 3.10).

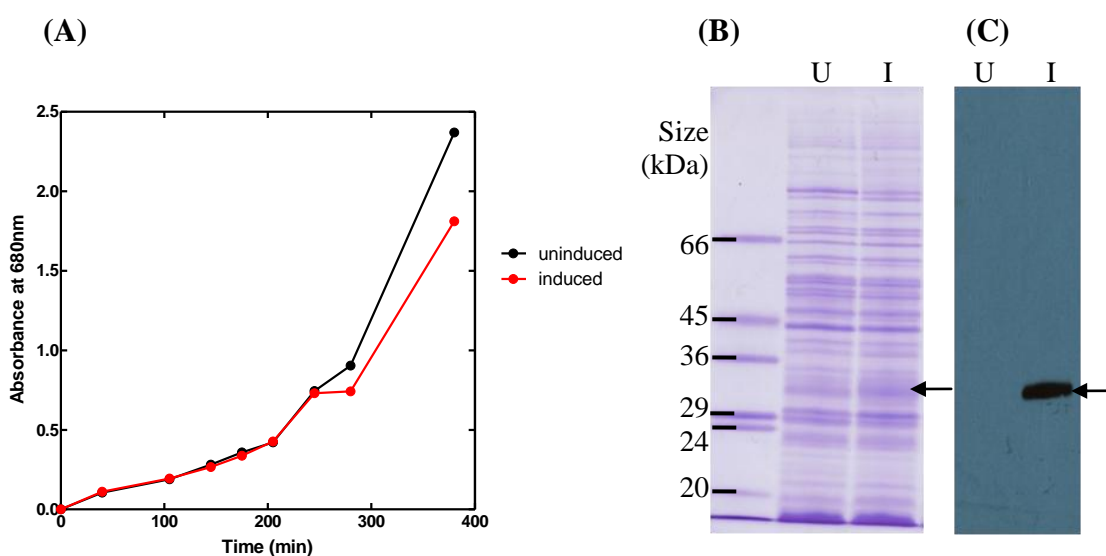


Figure 3.10 Growth and expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(*codB*)-His₆. (A) *E. coli* BL21(DE3) cells transformed with pTTQ18(*codB*)-His₆ were grown in LB medium supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction. Total *E. coli* membranes were prepared and analysed by (B) SDS-PAGE and (C) Western blotting. U (uninduced) and I (induced). Arrow indicates position of the CodB protein.

3.2.4. Uptake of radiolabelled cytosine by CodB

A time course of cytosine transport by CodB showed an increase in uptake of cytosine in the induced cells expressing CodB, up to a maximum of 1.8 nmol mg⁻¹ cells (Figure 3.11). The level of radioactivity decreased after the peak, which could indicate metabolism of cytosine inside the cells, to radiolabelled product(s) that leave the cell. In bacteria, cytosine taken up into the cell is either incorporated into DNA and RNA or hydrolytically deaminated to uracil and ammonia. The ammonia is excreted from the cell and the uracil is converted to barbituric acid, which is then degraded to urea and malonic acid (Figure 3.12). Oxidation of malonic acid then produces carbon dioxide and ³H-water (Figure 3.12).

The radiolabelled cytosine used in this project is ³H-5-cytosine. The decreased radioactivity in the induced cells could be caused by the excretion of radiolabelled water from the cell (Figure 3.12). Here, the proportion of uptake or excretion of these compounds could not be defined. The induced cells expressing CodB showed an increase in uptake of radiolabelled compound after 5 minutes of the time course, this could suggest the cells recommence taking up cytosine when levels get low (Figure 3.12).

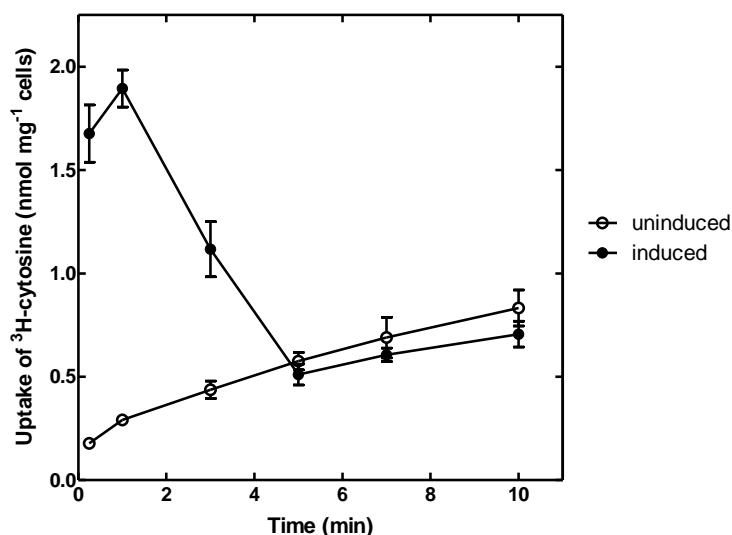


Figure 3.11 Time course of cytosine uptake by CodB. Uptake of ³H-cytosine in *E. coli* BL21(DE3) cells transformed with pTTQ18(*codB*)-His₆. Cells were cultured in minimal medium in the presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4-0.6, 0.5 mM of IPTG was added the induced sample and cells were harvested 1 hour post-induction, followed by washing of cells. Cells were energised with 20 mM glycerol, and incubated in 50 µM ³H-cytosine in 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Results represent means of triplicates (±SEM).

The rapid decrease in radioactivity might indicate a high level of cytosine is toxic to the *E. coli* cells, and so the cell tries to metabolise and excrete the toxic compound. At present there are no known cytosine exporters in bacteria, so it is most likely that the decrease in radioactivity is due to excretion of water as a waste product of cytosine metabolism.

In the uninduced cells, a gradual increase in radioactivity suggested cytosine is taken up into the cells (Figure 3.11). No expression of CodB is observed in the uninduced cells, shown by the SDS-PAGE and Western blot of total membrane of CodB expression (Figure 3.10); therefore the uptake observed could be endogenous cytosine transport system in the host cell.

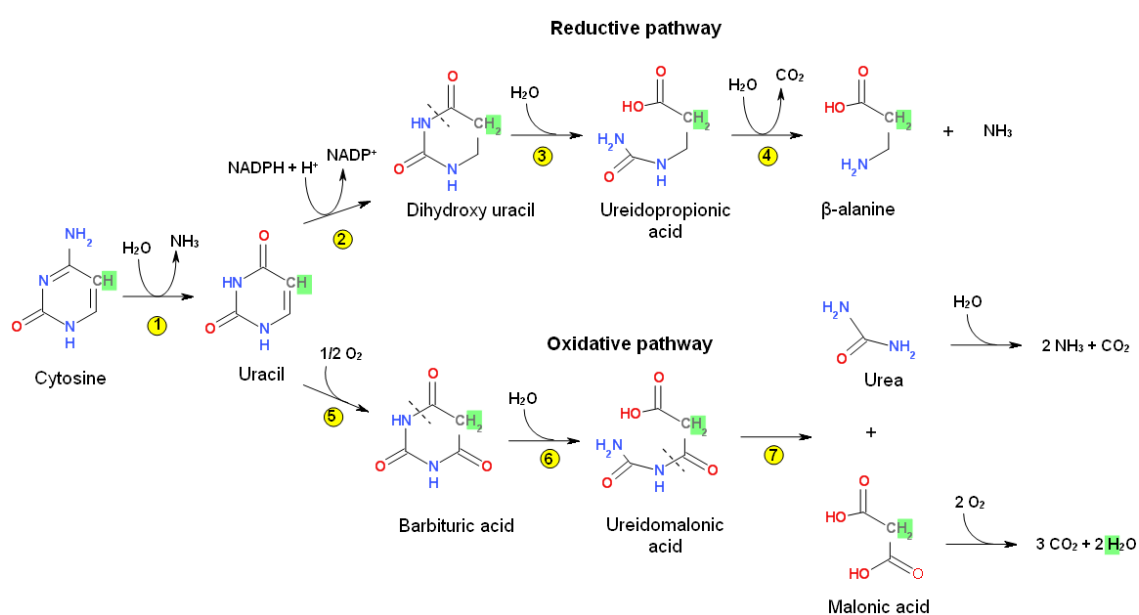


Figure 3.12 Pyrimidine catabolic pathways in bacteria. Bacterial reductive and oxidative pathways for catabolism of pyrimidine rings. Enzymes in the pathways are as follows: (1) cytosine deaminase; (2) dihydropyrimidine dehydrogenase; (3) dihydropyrimidinase; (4) ureidopropionase; (5) uracil dehydrogenase; (6) barbiturase; (7) ureidomalonase. The radiolabelled hydrogen in ³H-5-cytosine is highlighted in green, showing its catabolism through the two pathways. Diagram modified from Loh *et al.* (2006) and Hayaishi & Kornberg (1952).

3.2.5. Substrate specificity of CodB

To study the substrate specificity of CodB, ³H-cytosine uptake in *E. coli* cells expressing CodB was determined in the presence of ten-fold molar excess of potential substrates (Figure 3.7 and 3.12). This competition assay of CodB showed benzylhydantoin and uracil are able to inhibit transport of cytosine into the cells (Figure

3.13). Benzylhydantoin was able to inhibit uptake of cytosine at 15 seconds and 2 minutes post-addition of cytosine, whereas uracil only had an effect at 15 seconds (Figure 3.13). Possible competition of cytosine uptake in presence of ten-fold molar excess of potential competitors showed relative uptake of 80-120% (Figure 3.13). These results indicated CodB has a narrow substrate specificity, being specific for cytosine transport and no other nucleobases that were tested here apart from the effect of uracil at 15 seconds.

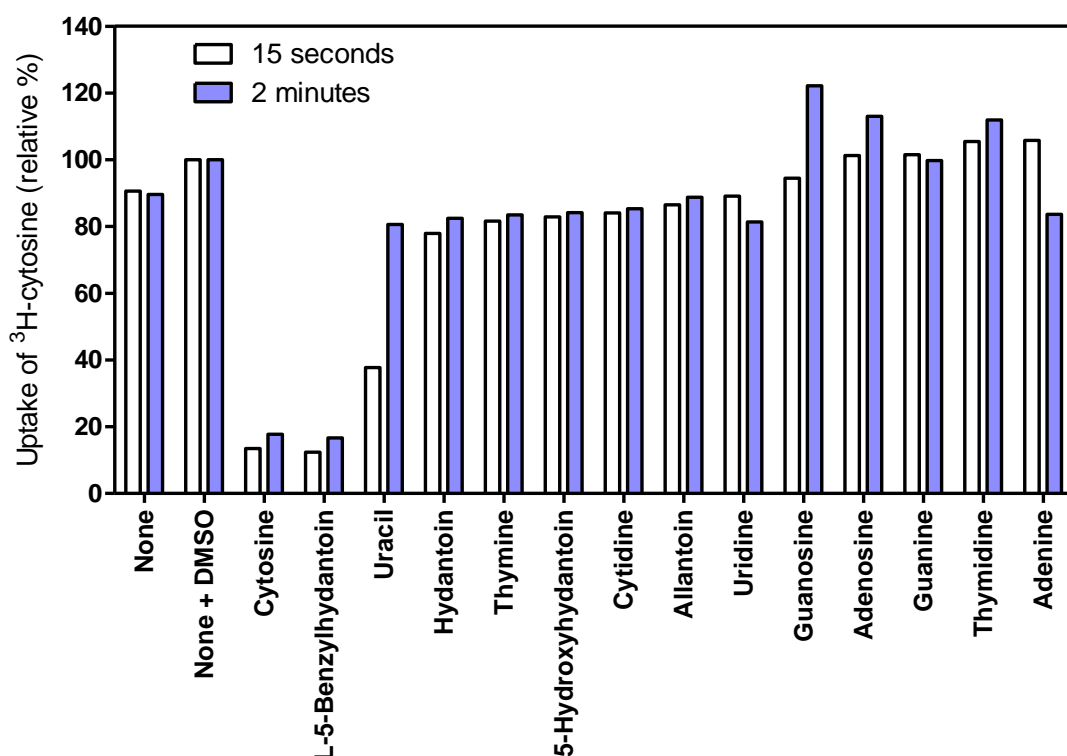


Figure 3.13 Substrate specificity of CodB. Competition of ^3H -cytosine ($50\ \mu\text{M}$) uptake into *E. coli* BL21(DE3) cells expressing CodB in the presence of a ten-fold molar excess of potential competitors. The noncompeted uptake rate was taken as 100% corresponding to 1.88 and 1.93 nmol cytosine per mg of cells for 15 seconds and 2 minutes post-addition of ^3H -cytosine, respectively. Buffers contain 2% DMSO. Results represent means of duplicated experiments. None = no competitors or DMSO (control).

Competition with an excess of unlabelled cytosine caused a reduction of the uptake at 15 seconds and 2 minutes, to 13 and 18% respectively (Figure 3.13), confirming that CodB mediates the transport of cytosine. L-Benzylhydantoin was also found to be effective at inhibiting ^3H -cytosine uptake when applied in excess, suggesting that CodB also binds benzylhydantoin (Figure 3.13).

The common features of cytosine, benzylhydantoin, uracil and hydantoin are highlighted in Figure 3.14. Hydantoin is a poor substrate for CodB, indicated by the small reduction in ^3H -cytosine uptake in presence of ten-fold molar excess of hydantoin (Figure 3.13). This suggested the six-membered ring, rather than the hydantoin group, of the benzylhydantoin might play a crucial role in binding to CodB (Figure 3.13).

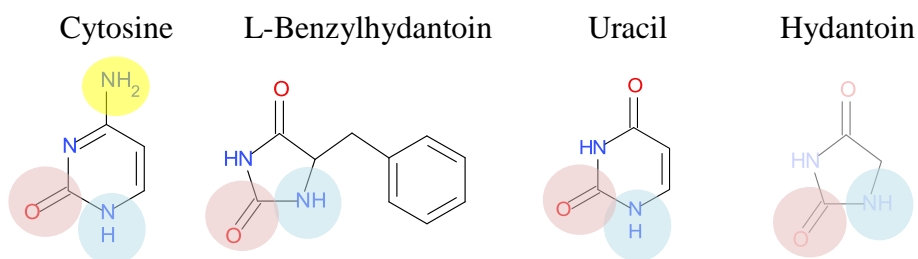


Figure 3.14 Common structural features of CodB substrates. The unique NH_2 group in cytosine (yellow); common structural groups in CodB substrates (pink and blue).

The structure of uracil is highly similar to cytosine. Uracil also has a six-membered ring, but the results suggested that the NH_2 group in cytosine is important for CodB to select cytosine over uracil (Figure 3.13 and 3.14). However, the highlighted pink and blue part of the selected substrates (Figure 3.14) may be important for initial recognition by CodB.

3.2.6. Kinetic analysis of cytosine transport by CodB

Transport studies in *E. coli* BL21(DE3) cells expressing CodB with ^3H -cytosine confirmed cytosine uptake is concentration dependent and showed a Michaelis-Menten constant (K_m) for cytosine of $30.96 \pm 11.06 \mu\text{M}$, and displayed saturation kinetics with a V_{max} of $1.495 \pm 0.1016 \text{ nmol mg}^{-1} \text{ cells}$ (Figure 3.15). Uninduced cells demonstrated a continuous uptake of cytosine and no saturation was observed with the concentrations tested (Figure 3.15). Since the SDS-PAGE and Western blot showed no expression of His₆-tagged CodB in the total *E. coli* membrane (Figure 3.10), the effect in the uninduced cells must be the host cells constitutively expressing proteins that transport cytosine, thus this could be the host cell's CodB. Therefore the kinetics of the uninduced cells was subtracted from the induced cells (Induced minus uninduced), giving rise to kinetics for the His₆-tagged CodB (Figure 3.10), which showed a K_m and V_{max} for cytosine of $7.748 \pm 4.894 \mu\text{M}$ and $1.006 \pm 0.0591 \text{ nmol mg}^{-1}$, respectively.

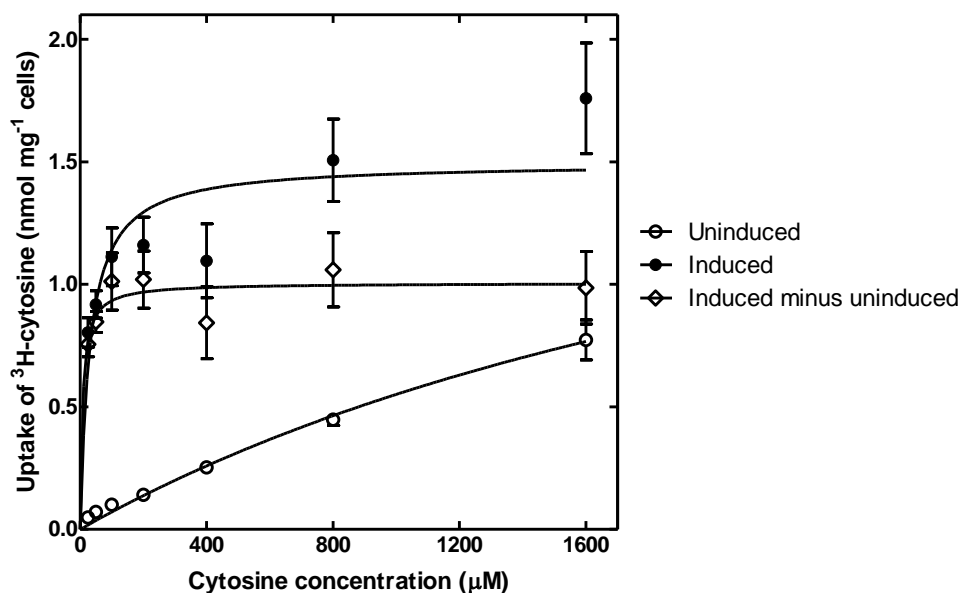


Figure 3.15 Michaelis-Menten kinetics of ³H-cytosine uptake by CodB. *E. coli* BL21(DE3) cells transformed with pTTQ18(*codB*)-His₆ were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4-0.6, 0.5 mM IPTG was added to the induced sample and cells were 20 mM glycerol harvested 1 hour post-induction, followed by washing of cells. Cells were energised with 20 mM glycerol, and incubated in various concentrations of ³H-cytosine (40 – 1600 µM) in 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Samples were taken 15 seconds post-addition of ³H-cytosine. Results represent means of triplicates (±SEM).

3.2.7. Purification of CodB from the inner membrane

A six litre culture of *E. coli* BL21(DE3) expressing the His₆-tagged CodB was grown aerobically as described in Section 2.4.1. The harvested cells were lysed by explosive decompression using a cell disruptor and the inner membrane was prepared by separation of the outer from the inner membrane on a sucrose gradient (Ward *et al.*, 2000) (Section 2.5.1). The membrane fractions obtained from the sucrose gradients were analysed by SDS-PAGE (Figure 3.16). The gel showed the presence of CodB in the prepared inner and outer membranes, suggesting poor separation by the sucrose gradients for this protein. CodB was present at 13% in the prepared inner membrane, and 9% and 8% in the mixed and outer membrane respectively (Figure 3.16), as determined by scanning densitometry. The sucrose gradients step appears to be an important step for making CodB more abundant, as indicated by densitometry, the protein was increased from 9% in the mixed membrane (Figure 3.6 and Figure 3.16) to 13% in the inner membrane after the sucrose gradients (Figure 3.16). The red arrow

shows the position of a abundant protein migrating at ~31 kDa in the mixed and outer membranes, that is not CodB (Figure 3.16). The abundance of this protein (red arrow) in the outer membrane and its absence or low abundance in the inner membrane after sucrose gradients, could suggest this protein is porin, which is found abundantly in the outer membranes of *E. coli*.

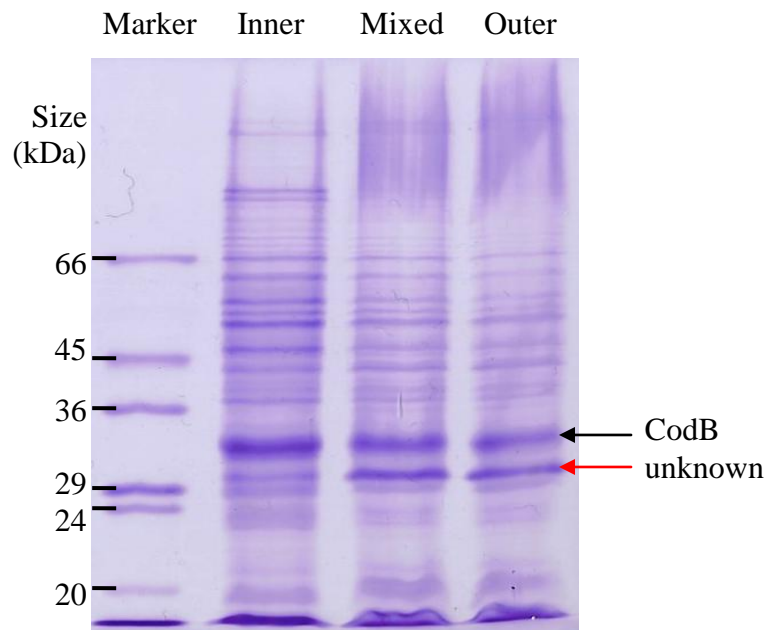


Figure 3.16 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) expressing CodB. Total *E. coli* membranes were prepared from *E. coli* BL21(DE) pTTQ18(*codB*)-His₆ grown on LB medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes were separated on 25-55% sucrose gradients and membrane fractions extracted from the gradients were analysed by SDS-PAGE. The black arrow indicates the position of the CodB protein, and the red arrow shows the position of the protein removed from the inner membrane.

The inner membranes were subjected to solubilisation and purification as described in Section 2.5.2 and Section 2.5.3. Fractions from the purification of CodB were analysed by SDS-PAGE and Western blotting (Figure 3.17). The majority of the CodB protein was solubilised in the 1% DDM buffer, indicated by very small amount of CodB in the insoluble fraction, shown by the SDS-PAGE and Western blot (Figure 3.17). Some CodB proteins that were not bound to the Ni-NTA resin was lost in the flow-through, unbound fraction (Figure 3.17), and indicates a longer binding time or increased amount of resin could reduce the loss.

Some contaminants were co-eluted with the CodB protein, migrating at ~45 kDa and >66 kDa on the SDS-PAGE, which were not detected by the Western blotting (Figure 3.17). The upper contaminant on the SDS-PAGE is often found in *E. coli* overexpression systems; this is the multidrug efflux protein (AcrB) of *E. coli*, which was previously confirmed by Mass spectrometry (Bettaney, 2008). A further purification step, such as size-exclusion chromatography, could be used to remove the contaminating proteins. The arrows in Figure 3.17 indicate the position of the CodB protein, migrating at 33 kDa and 55 kDa. The CodB protein was purified to approximately 70% as estimated by scanning densitometry.

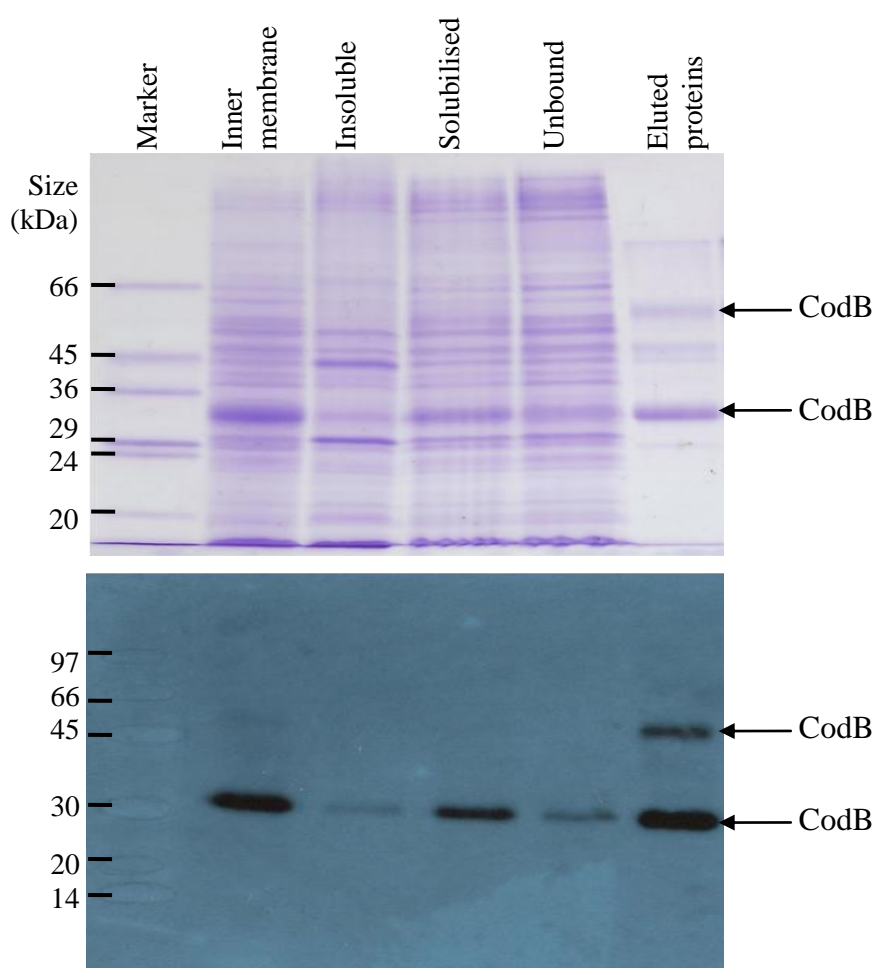


Figure 3.17 SDS-PAGE and Western blot of purification of *E. coli* CodB. CodB was purified from the inner membrane of *E. coli* BL21(DE3) pTTQ18(*codB*)-His₆ cells. The inner membranes were solubilised in 1% DDM. Samples (15 µg), and the final eluted fraction (10 µg) were analysed by SDS-PAGE (top) and Western blotting (bottom). The arrow indicates the position of the CodB protein.

3.2.8. Testing the integrity of CodB by circular dichroism

The secondary structure of the part-purified CodB was examined by circular dichroism (Section 2.6.1). The negative indentations at 208 and 222 nm, and a positive signal at 190 nm, are indicative of a predominantly α -helical content, agreeing with the previous topology analysis (Danielsen *et al.*, 1995) and prediction by TMHMM (Figure 3.9), confirming the retention of the secondary structure of CodB during the purification process (Figure 3.18).

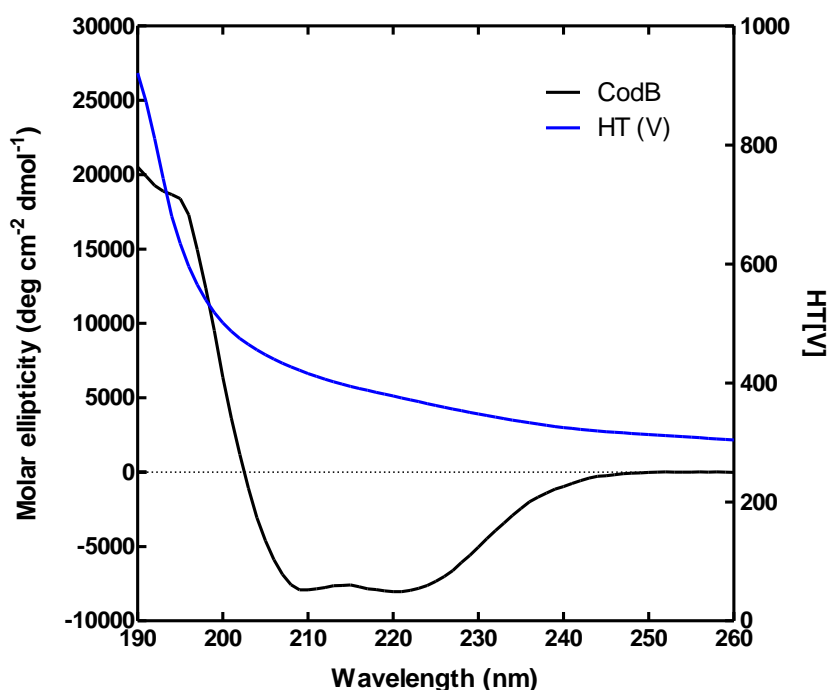


Figure 3.18 Circular dichroism spectrum of purified CodB of *E. coli*. Purified CodB (0.05 mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of CodB was performed using Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample was analysed in Hellma quartz-glass cell of 1mm path length. Spectrum was recorded with 1 nm step resolution at a scan rate of 10 nm/min. Response time was set at 1 second with a sensitivity of 20 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of ten scans, from which the buffer contribution was subtracted.

3.2.9. Discussion: characterisation of CodB

Currently, only one cytosine transporter, CodB (Danielsen *et al.*, 1992) has been identified in *E. coli*. Analysis using FASTA3 showed that the *E. coli* CodB is highly conserved throughout the *E. coli* strains, *Salmonella* and *Vibrio* species (Section 3.2.2.). In the phylogentic tree analysis of the selected NCS1 proteins, CodB clustered very

closely with the *P. aeruginosa* PA0438 protein, suggesting this protein could be a cytosine transporter too (Figure 3.2). The *S. coelicolor* SCO0572 protein also clustered quite closely, indicating this may also be a cytosine transporter (Figure 3.2).

The conserved residues of the NCS1 proteins are highlighted in the topology model of CodB (Figure 3.9); most of these are identical or highly similar within the family (Appendix 2). In CodB, the conserved residues are found in: the N-terminus (R17 and K18); helix I (V26, M27, L28 and G29); the cytoplasmic loop between helices I and II (G46 and L47); the cytoplasmic loop between helices II and III (G74, K76, G78, L79, R86, F89, G90, G93 and S94); helix III (L96, P97, W108 – where a F, G and V residue often follow after it); helix IV (G146 and L150); helix VI and the cytoplasmic loop after it (F204, D213, F214, V215, R216 and F217); and in helix X and the cytoplasmic loop after it (W324, L325, G337, V338, I340, D342, Y343, R347, R348 and Y350) (Figure 3.9 and Appendix 2).

Based on the structures of Mhp1 (Weyand *et al.*, 2008) and the alignment in Appendix 2, the potential residues in CodB that could be involved in substrate recognition and binding were identified. As mentioned earlier in Chapter 1, residue W117 in helix III of Mhp1 forms π -stacking with the hydantoin moiety of its substrate (Figure 1.9). Here, if the alignment in Appendix 2 is correct, residue W108 in helix III of CodB may potentially form π -stacking with the hydantoin moiety of L-benzylhydantoin, a compound which inhibited uptake of cytosine by CodB (Figure 3.13). Residue W220 in helix VI of Mhp1 is predicted to form π -stacking with the benzyl moiety of its substrate (Figure 1.9). Alignment in Appendix 2 suggests residue F204 in helix VI of CodB could potentially form π -stacking with the aromatic moiety of cytosine or benzylhydantoin. Residues S312 and T313 are involved in cation binding in Mhp1; these align with T278 and T279 in CodB, suggesting these two threonine residues have the potential to play a role in cation binding (Appendix 2).

Here, the CodB protein is cloned into plasmid pTTQ18-His₆ (Figure 3.3 – 3.5) and expressed at 9% in the *E. coli* total membrane, migrating at ~33 kDa on the SDS-PAGE (Figure 3.6 and Table 3.3), which is similar to the migration (32 kDa) previously reported by Danielsen *et al.* (1992). The higher migration of CodB in this project could be as a result of the incorporated MNSH and His₆-tag at the N- and C- terminus of the protein, respectively. A slightly slower growth rate was observed in *E. coli* BL21(DE3) pTTQ18(CodB)-His₆ induced with IPTG (Figure 3.10), possibly as a consequence of the cells' expression of CodB and reduced growth rate to compensate. No expression of

CodB was observed in the uninduced cells, indicated by the negative Western blot of the uninduced cells (Figure 3.10), suggesting tight repression of the *tac* promoter.

Transport assays showed cytosine is transported into the cell by CodB (Figure 3.11), and CodB has a narrow substrate specificity, for cytosine, and only benzylhydantoin inhibited the uptake of cytosine by CodB (Figure 3.12). Uptake of cytosine by CodB is concentration dependent, having a K_m and V_{max} of $7.748 \pm 4.894 \mu\text{M}$ and $1.006 \pm 0.0591 \text{ nmol mg}^{-1}$, respectively, in cells expressing CodB (Figure 3.15). Two previously studied transporters, cytosine-purine permease of *S. cerevisiae* (Fcy2) and cytosine-purine scavenger of *A. nidulans* (FcyB), of the NCS1 family have a K_m of $13 \mu\text{M}$ and $20 \mu\text{M}$ respectively (Benoît *et al.*, 1997; Vlanti & Diallinas, 2008), which is higher than that of the *E. coli* CodB. Since the uninduced cells of BL21(DE3) pTTQ18(CodB)-His₆ are shown to take up cytosine (Figure 3.11 and 3.15), the transport experiment should be repeated with an *E. coli codB*-negative strain to confirm the transport activity is carried by the host cell.

The percentage of α -helical or β -sheet contents of CodB was not determined, since it was conventional CD being performed and does not give accurate data at wavelength below 190 nm, due to the limitation of the xenon lamp in the J-715 spectropolarimeter, but it was a quick in-house tool that successfully checked the integrity of CodB (Figure 3.18). HT voltage is roughly proportional to absorbance. If the protein sample is too concentrated, not enough light reaches the detector, which causes saturation and an accurate spectrum will not be recorded. Synchrotron radiation CD (SRCD) could be used, giving reliable spectra measured down to 160 nm, where the complete spectra of α -helical, β -sheet or random coil element could be achieved (Lees & Wallace, 2002).

In this study, CodB was successfully cloned into pTTQ18-His₆ and expressed in *E. coli* BL21(DE3) host to allow purification and production of the protein in milligrams, and activity studies have shown the expressed CodB is functional. This allows the protein to proceed to further studies, which may include mutagenesis to identify residues in CodB that could be important for its substrate binding and recognition, and crystallisation trials for structural studies.

3.3. Characterisation of a *Bacillus subtilis* allantoin transporter – PucI (Bsu36470)

3.3.1. Introduction – nitrogen metabolism & the *puc* genes of *B. subtilis*

Bacillus subtilis is a gram-positive, sporulating soil bacterium that is able to utilise a diverse range of low-molecular-weight compounds as a nitrogen source when the preferred nitrogen sources, glutamate, ammonium, or glutamine, are depleted (Fisher, 1999). Nitrogen metabolism in *B. subtilis* is controlled by the availability of rapidly metabolisable nitrogen sources, rather than by the two-component Ntr regulatory system found in most enteric bacteria (Fisher, 1999). At least three regulatory proteins independently control the expression of genes involved in nitrogen metabolism in response to nutrient availability (Fisher, 1999). Genes expressed at high levels during nitrogen-limited growth are controlled by two related proteins, GlnR and TnrA (Schreier *et al.*, 1993; Wray *et al.*, 1996; Wray *et al.*, 2000) (Table 3.6). CodY, a third regulatory protein, controls expression of several genes involved in nitrogen metabolism and acetate metabolism in response to growth rate (Fisher *et al.*, 1996; Wray *et al.*, 1997) (Table 3.6).

Under the preferred nitrogen source conditions, *B. subtilis* assimilates all of its nitrogen by the glutamine synthetase (GlnA)-catalysed reaction ($\text{NH}_3 + \text{glutamate} + \text{ATP} \rightarrow \text{glutamine} + \text{ADP} + \text{P}_i$). When *B. subtilis* is grown with a less preferred nitrogen source, a number of other enzymes and permeases involved in the assimilation of nitrogen are induced (Fisher, 1999) (Table 3.6). This includes asparaginase (Atkinson & Fisher, 1999), γ -aminobutyric acid permease (*gabP*) (Ferson *et al.*, 1996), urease (*ureABC*), a putative ammonium permease (*nrgA*) (Wray *et al.*, 1994), and the nitrate assimilatory enzymes (*nasABCDEF*) (Nakano *et al.*, 1995; Nakano *et al.*, 1998) (Table 3.6). The transcription of these genes is controlled by TnrA (Wray *et al.*, 1996; Nakano *et al.*, 1998). In addition, TnrA positively regulates its own transcription, and also functions as a negative regulatory protein under nitrogen limited growth, repressing the expression of the ammonium assimilatory enzymes, GlnA, as well as the glutamate synthase operon (*gltAB*) (Wray *et al.*, 1996; Fisher, 1999; Doroshchuk *et al.*, 2006) (Table 3.6).

TnrA and GlnR are homologous, belonging to the MerR family of regulatory proteins, and bind to the same DNA sequence known as the GlnR/TnrA box (5'-TGTTNAN₇TNACA-3'); they are negative regulatory proteins for the glutamine synthase

operon (*glnRA*) (Wray *et al.*, 1996). During growth with excessive nitrogen (glutamine or ammonia plus glutamate), GlnR functions as a repressor and inhibits the expression of the *glnRA* operon encoding GlnR and GlnA (Wray *et al.*, 1996; Fisher, 1999). Whereas, TnrA binds to GlnA, forming an inactive complex (TnrA-GlnA), and consequently, TnrA cannot interact with its specific operator, leading to inhibition of the expression of alternative nitrogen-assimilatory pathways (Wray *et al.*, 1996; Wray *et al.*, 2000; Wray *et al.*, 2001). Under nitrogen-limited conditions, TnrA is released from the TnrA-GlnA complex and GlnR binding activity is inhibited, allowing TnrA to bind to its specific operator, and hence activation of the alternative nitrogen-assimilatory pathways (Fisher, 1999; Wray *et al.*, 2000; Doroshchuk *et al.*, 2006).

CodY is activated under nitrogen-limited conditions where rapid growth occurs on amino acids (Wray *et al.*, 1997). It is a negative regulator of the histidine degradative operon (*hut*) (Fisher *et al.*, 1996), the dipeptide transporter operon (*dpp*) (Slack *et al.*, 1995), the isoleucine and valine degradative operon (*dkd*), *ureABC* and *gabP* (Ferson *et al.*, 1996) (Table 3.6). A summary of the regulatory proteins involved in nitrogen metabolism in *B. subtilis* is listed in Table 3.6.

Regulatory protein	Growth conditions under which regulation occurs	Regulation	Regulated gene/promoter/enzymes	References
TnrA	Nitrogen limitation	Positive	<i>nrgAB</i> , <i>gabP</i> P2, <i>ureABC</i> P3, <i>nasDEF</i> , <i>tnrA</i> , asparaginase, <i>puc</i> genes	Wray <i>et al.</i> (1994) Ferson <i>et al.</i> (1996) Wray <i>et al.</i> (1996) Cruz-Ramos <i>et al.</i> (1997) Nakano <i>et al.</i> (1998) Schultz <i>et al.</i> (2001) Yoshida <i>et al.</i> (2003) Doroshchuk <i>et al.</i> (2006)
	Nitrogen limitation	Negative	<i>glnRA</i> , <i>gltAB</i>	Yoshida <i>et al.</i> (2003)
GlnR	Excess nitrogen	Negative	<i>glnRA</i> , <i>tnrA</i> , <i>ureABC</i> P3	Fisher <i>et al.</i> (1984) Strauch <i>et al.</i> (1988) Cruz-Ramos <i>et al.</i> (1997) Yoshida <i>et al.</i> (2003)
CodY	Rapid growth with amino acids	Negative	<i>Hut</i> , <i>gabP</i> P1, <i>ureABC</i> P3, <i>ureABC</i> P2, <i>dpp</i>	Yoshida <i>et al.</i> (1995) Ferson <i>et al.</i> (1996) Cruz-Ramos <i>et al.</i> (1997) Mathiopoulos <i>et al.</i> (1991) Yoshida <i>et al.</i> (2003)

Table 3.6 The regulatory proteins involved in nitrogen metabolism in *B. subtilis*. Table modified from Fisher, S. (1999).

B. subtilis is capable of taking up purine bases and using them for nucleotide synthesis and as a source of nitrogen. In *B. subtilis*, the natural purine bases all serve as nitrogen sources when the preferred nitrogen source are exhausted, and the *puc* genes (Figure 3.19 and 3.20) are required for the catabolism of purines (Schultz *et al.*, 2001). The *puc* region contains two binding sites for TnrA, and expression of *pucR* is activated by TnrA (Beier *et al.*, 2002). The PurR transcription factor is a pathway-specific regulator of the purine catabolic genes of the *puc* regulon (Schultz *et al.*, 2001; Beier *et al.*, 2002), enabling the cell to utilise purines as an alternative nitrogen source (Doroshchuk *et al.*, 2006). Surprisingly, separate genes of the *puc* regulon were found in *B. lichieniformis*, but they are not regulated by TnrA; also the other bacilli are known to lack purine catabolism genes (Doroshchuk *et al.*, 2006), making *B. subtilis* unique in terms of utilisation of purine as an alternative nitrogen source.

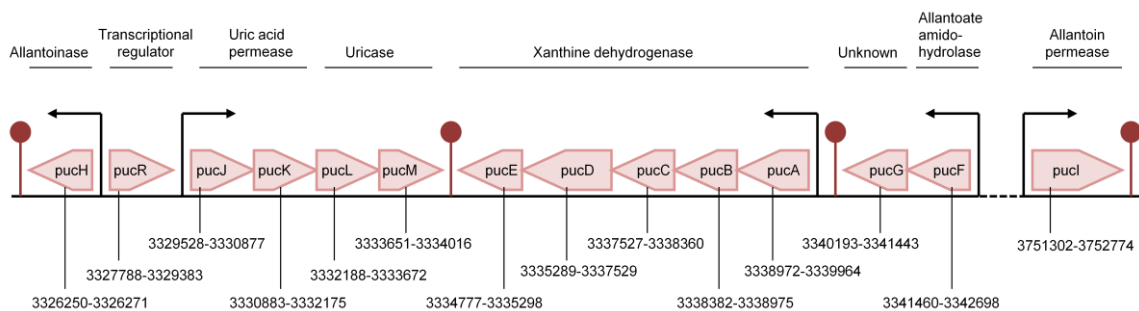


Figure 3.19 Map of the *B. subtilis* *puc* genes region. The bent arrow denotes the promoter site (\curvearrowright). The lollipop indicates the position of transcription termination (\lceil). The pink box represents a *puc* gene (\triangleright), and the number below corresponds to the genome position of the gene. The diagram is drawn based on the information in the DBTBS website (Ishii *et al.*, 2001) and Schultz *et al.* (2001).

The PucR protein of *B. subtilis* regulates expression of fifteen genes involved in purine catabolism (Beier *et al.*, 2002). PucR functions as both a transcriptional activator and repressor. The expression of genes regulated by PucR is enhanced by growth in medium containing intermediates in purine catabolism, for example, uric acid, allantoin or allantoic acid (Schultz *et al.*, 2001; Beier *et al.*, 2002). Analysis using *pucR* deletion mutants, confirmed that PucR induces the expression of *pucFG* (allantoate amidohydrolase and unknown, respectively), *pucH* (allantoinase), *pucI* (allantoin permease), *pucJKLM*, and (uric acid transporter and uricase), *gde* (guanine deaminase), whereas it represses the expression of *pucR* (purine catabolism regulator) and *pucABCDE* (xanthine dehydrogenase) (Beier *et al.*, 2002).

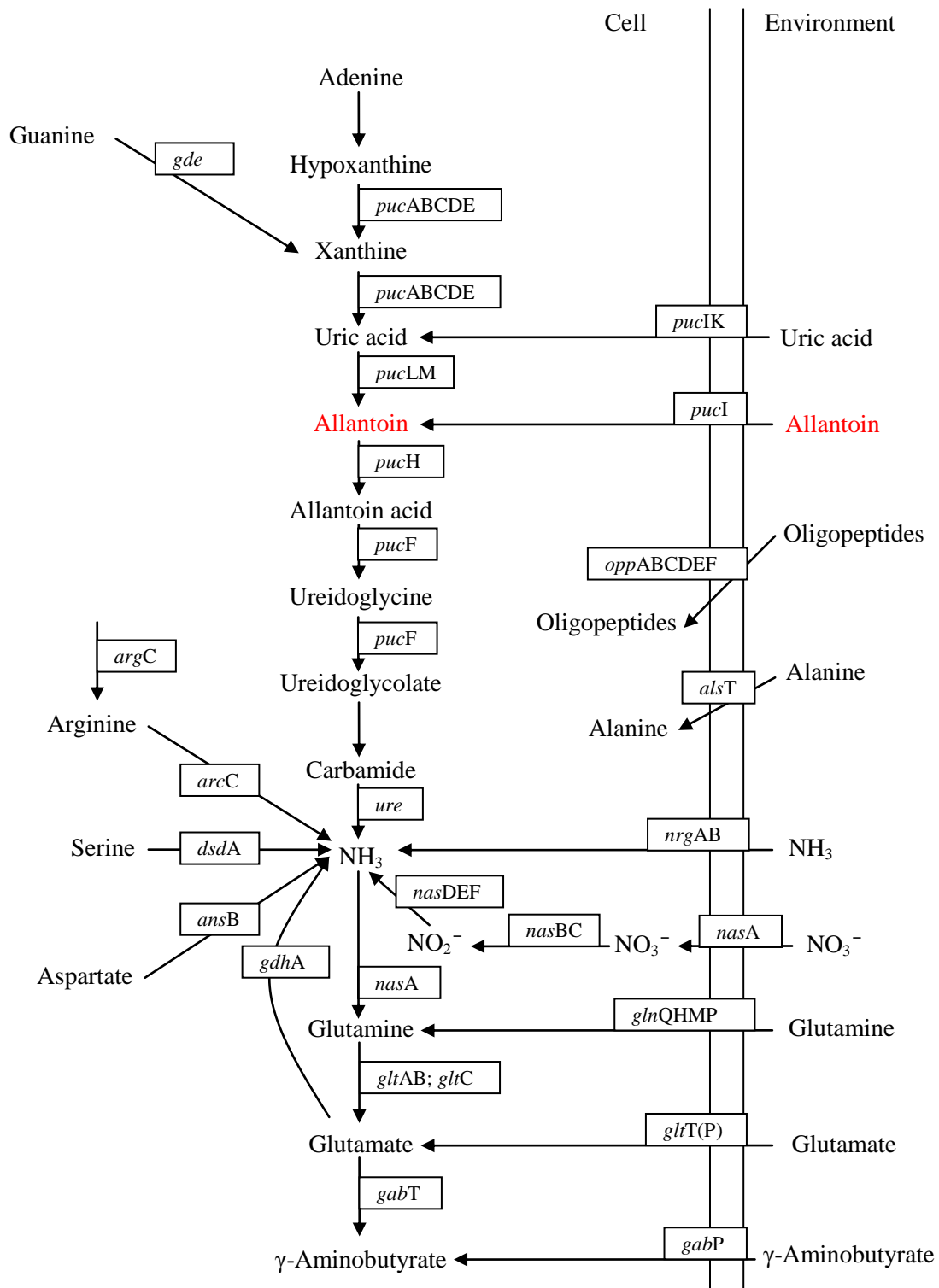


Figure 3.20 Pathway of nitrogen metabolism in *Bacillus subtilis*. Purine degradative genes (*puc*) of *B. subtilis* are shown in the top half of the diagram. The *nas*, *gln*, *glt* and *gab* genes are involved in nitrite oxide, glutamine, glutamate and γ -aminobutyrate metabolism, and are shown in the bottom part, respectively. Diagram taken and redraw from Doroshchuk *et al.* (2006).

Potential PucR boxes were identified in the promoters of the the five induced operons, consisting of the conserved sequence 5'-WWWCNTTGGTTAA-3', and a possible PucR box was also found in the *pucABCDE* promoter region and downstream of the *pucR* transcription start site (Beier *et al.*, 2002). These genes are required for the complete degradation of purine bases to ammonia, as well as for the transport of uric acid and allantoin (Schultz *et al.*, 2001; Beier *et al.*, 2002). All *puc* genes except *pucI* are located in a gene cluster at 284 to 285° on the chromosome. The genome position of *pucI* is much further downstream of all the *puc* genes (Figure 3.19). Intracellular purines, such as guanine, adenine, and hypoxanthine are first converted to xanthine before entering the catabolic pathway (Figure 3.20). These *puc* genes form part of the important components in the pathways of nitrogen metabolism in *B. subtilis* (Figure 3.20). Section 3.3 describes the characterisation of the allantoin permease, PucI (Bsu36470). Firstly, the topology and the conserved sequence in PucI are analysed. The cloning, expression, purification and the transport activity of PucI are also studied.

3.3.2. The *B. subtilis* allantoin transporter (PucI): topology analysis and conserved residues

Topology prediction using TMHMM (Krogh *et al.*, 2001) showed PucI is highly hydrophobic, composed of twelve transmembrane helices (Figure 3.21). Conserved residues in PucI are highlighted in the model and are positioned mainly in helices I, II and X, as well as in the cytoplasmic loops between helices II and III, helices IV and V, and helices X and XI (Figure 3.21).

Studies reported that *B. subtilis* is able to utilise uric acid, allantoin, allantoic acid, adenine, hypoxanthine, xanthine, guanine and urea as sole nitrogen source (Rouf & Lomprey, 1986; Nygaard *et al.*, 1996; Cruz-Ramos *et al.*, 1997; Nygaard *et al.*, 2000). Mutants in the *pucI* gene were demonstrated not to grow on allantoin, suggesting the gene encodes an allantoin transporter (Schultz *et al.*, 2001). PucI is 28.8% identical (61.1% similar) to the allantoin permease of *S. cerevisiae* (DAL4). It is also related to the NCS1 proteins studied in this project; 43.3% identical to SCO6417 (74.4% similar), 38.8% identical to b0511 (75.1% similar) and 37.9% identical to EF3000 (73.9%) (Table 3.1). The phylogenetic tree showed PucI clustered most closely with the putative allantoin permeases, b0511 and EF3000 of the NCS1 family (Figure 3.2). Here, successful cloning, expression and activity studies of PucI revealed the *pucI* gene surely encodes for an allantoin transporter (Section 3.3.3 to 3.3.10). This work also

investigates the kinetics and substrate specificity of allantoin transport by Pucl (Section 3.3.3 to 3.3.10).

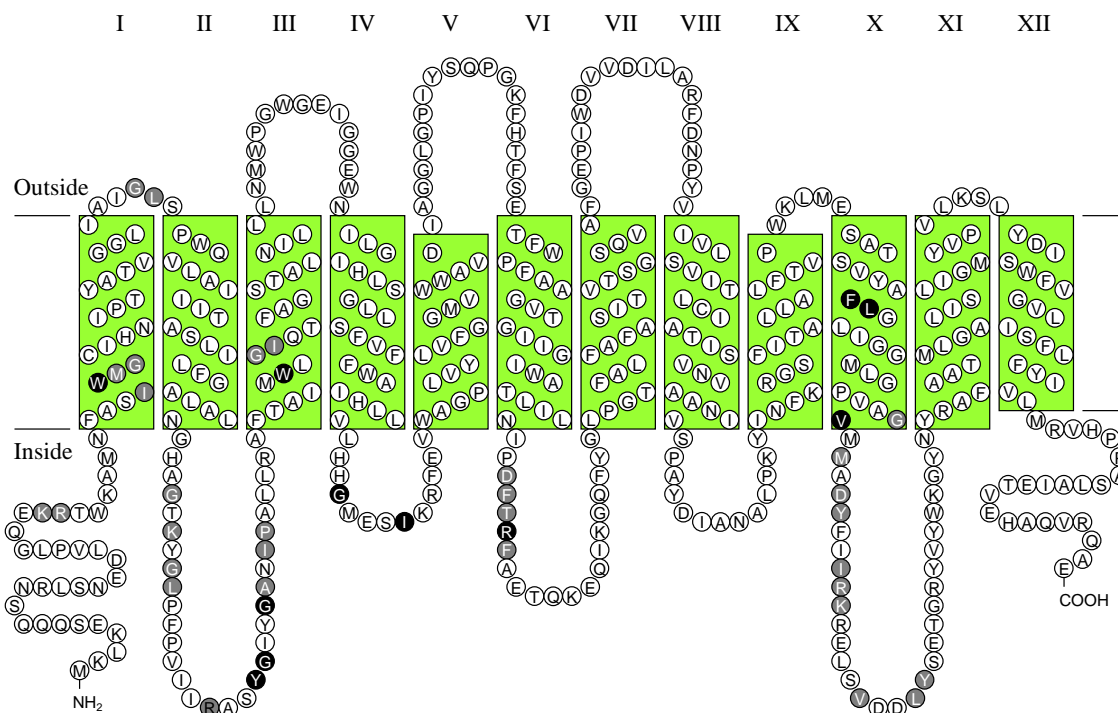


Figure 3.21 Topology model of *B. subtilis* allantoin permease (PucI). Topology prediction by TMHMM (Krogh *et al.*, 2001). The predicted helices are numbered from the N- to C-terminus and are shown in roman numerals. Conserved residues of the NCS1 family proteins are highlighted in grey and black according to the alignment in Appendix 2.

3.3.3. Cloning and expression of the *B. subtilis* PucI

Here, the *pucl* gene (Bsu36470) was amplified from the genomic DNA of *B. subtilis* 168 strain by PCR (Figure 3.3), and cloned into plasmid pTTQ18-His₆ with a His₆-tag added to the C-terminus of the protein. The clone was analysed by PCR screening and restriction digestion, confirming the correct sized gene in the plasmid construct (Figure 3.4 and 3.5). DNA sequencing confirmed no mutations in the cloned *pucl* gene, encoding for the correct amino acid sequence (Appendix 3 and 4). Subsequently, the clone was transformed into *E. coli* BL21(DE3) cells for expression trials, as described in Sections 2.3 and 2.4. The expression of PucI was optimised by testing in LB, 2TY and minimal media. Growth of the cells transformed with pTTQ18(*pucl*)-His₆ showed a detrimental effect when induced with IPTG (Figure 3.22).

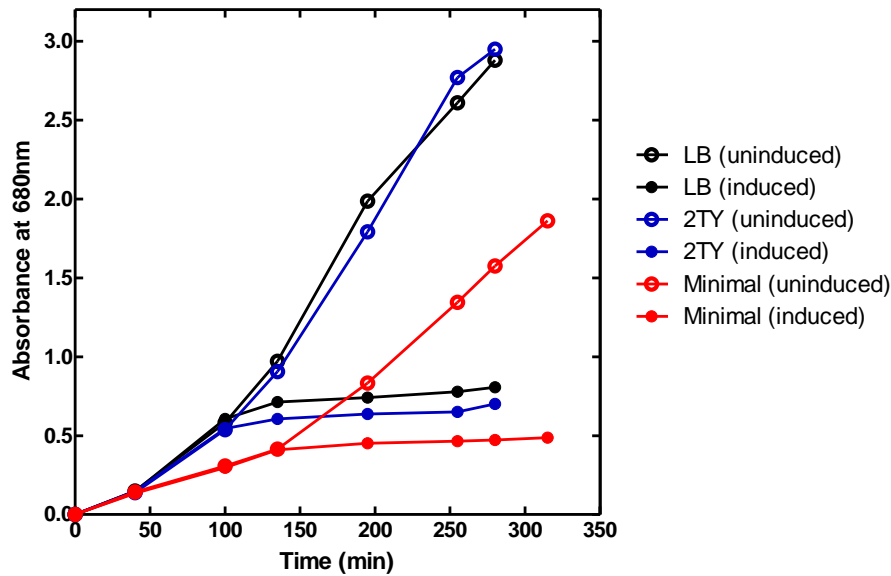


Figure 3.22 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ in LB, 2TY and minimal media. *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ were grown in LB, 2TY and minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction.

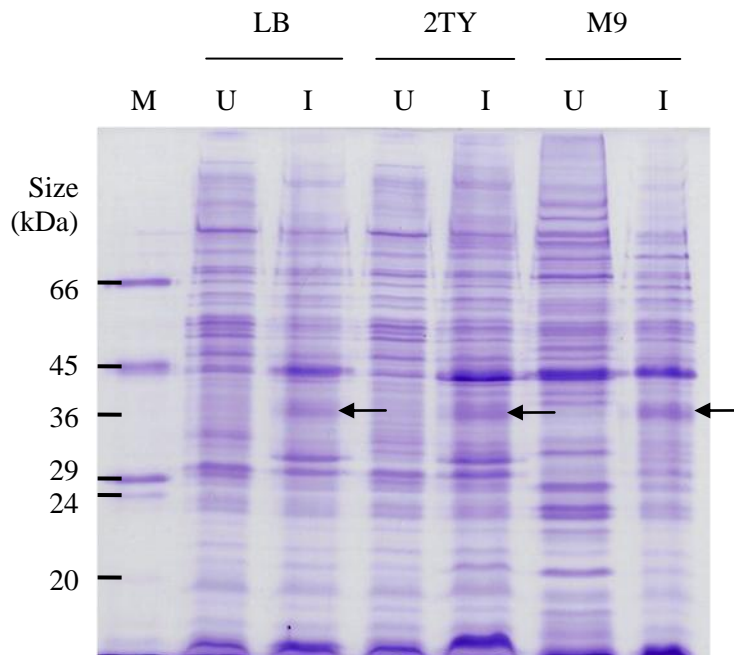


Figure 3.23 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆. BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ were grown in LB, 2TY and minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction, followed by preparation of total *E. coli* membranes and analysis by SDS-PAGE. M (marker), U (uninduced) and I (induced). Arrow indicates position of the PucI protein.

Similar growth rates were observed when cells were grown in either LB or 2TY medium (Figure 3.22). The absorbance of the uninduced cultures at time of harvesting was 2.880, 2.950 and 1.862 for cells grown in LB, 2TY and minimal media, respectively, and the rate of growth in minimal medium is approximately half that of cells grown on the other media (Figure 3.22). The absorbance of the induced cultures at time of harvesting were 0.806, 0.702 and 0.488 for LB, 2TY and minimal media, respectively (Figure 3.22), and densitometry suggests PucI is expressed at approximately 6% in the total prepared membrane for cells grown in LB, 2TY or minimal medium (Figure 3.23). The optimisation test suggests that LB is the choice of medium for expressing PucI, given that the cell density is highest in LB at time of harvesting, and since all media produce PucI at approximately the same percentage in the total *E. coli* membrane (Figure 3.22 and 3.23).

3.3.4. Time course of uptake of radiolabelled allantoin by PucI

A time course of allantoin transport by PucI showed increasing uptake of allantoin in the uninduced and induced cells expressing PucI, up to a maximum of 18.3 and 35.8 nmol mg⁻¹ cells, respectively (Figure 3.24).

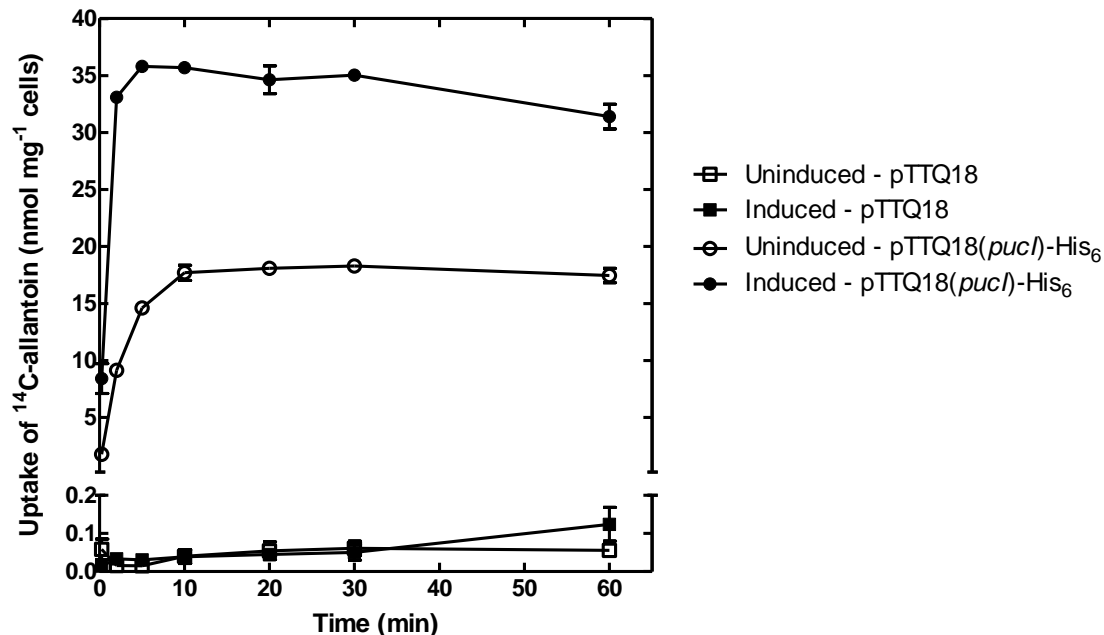


Figure 3.24 Time course of allantoin uptake by PucI. Uptake of ¹⁴C-allantoin in *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ or pTTQ18 – no genes. Cells were cultured in minimal medium in the presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4-0.6, 0.5 mM IPTG was added and cells were harvested 1 hour post-induction, followed by washing. Cells were energised with 20 mM glycerol, and incubated in 50 µM ¹⁴C-allantoin, 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Results represent means of triplicates (±SEM).

The uninduced and induced cells transformed with pTTQ18(*pucI*)-His₆ expressed the His₆-tagged PucI, no matter how long the cells were induced with IPTG, shown by the SDS-PAGE and Western blot in Figure 3.26, indicating leaky expression of PucI in the uninduced cells. As a control, uninduced and induced cells with plasmid containing no *pucI* gene (pTTQ18) were examined for the ability to transport allantoin (Figure 3.24). The results demonstrated relatively little uptake (< 0.13 nmol mg⁻¹ cells) in these cells, further confirming leaky expression from uninduced cells with pTTQ18(*pucI*)-His₆ (Figure 3.24). In induced cells harbouring pTTQ18(*pucI*)-His₆, the maximum uptake was reached at approximately 5 and 10 minutes after addition of the radiolabelled substrate, suggesting rapid uptake of allantoin into the cells (Figure 3.24).

3.3.5. Optimisation of uptake of allantoin by PucI

The transport of allantoin by PucI was optimised by testing a range of IPTG-induction periods. BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ were induced with IPTG for 0.1 to 22 hours when absorbance at 680 nm reached 0.4 – 0.6. Cells were harvested and washed three times in 5 mM MES pH 6.6 and 150 mM KCl, followed by transport assays. In the induced cells, an increasing transport of radiolabelled allantoin by PucI is observed when induction time was increased from 0.1 to 2 hours (Figure 3.25).

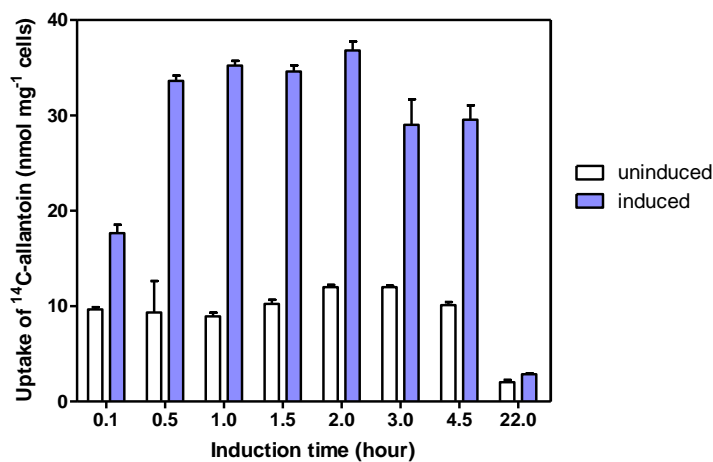


Figure 3.25 Effect of induction time with IPTG on allantoin uptake by PucI. Uptake of ¹⁴C-allantoin in *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆. Cells were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4 – 0.6, 0.5 mM IPTG was added and cells were induced for 0.1 – 22 hours before harvesting. Uninduced cells were grown according to the induced cells, but no IPTG was added. Harvested cells were energised with 20 mM glycerol, and incubated in 50 µM ¹⁴C-allantoin with 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Samples were taken 2 minutes post-addition of radiolabelled allantoin. Results represent means of triplicates (±SEM).

After that the uptake of allantoin by PucI was reduced (Figure 3.25). The cells, 22 hours post-induction, had significantly lost allantoin transport activity (Figure 3.25). After the transport assays the cells were subjected to a total membrane preparation, and analysis by SDS-PAGE and Western blotting, confirming the expression of His₆-tagged PucI in the cells that were tested (Figure 3.26). The uninduced cells showed leaky expression of PucI, indicated by the Western blotting (Figure 3.26), giving rise to the transport activity in the uninduced cells (Figure 3.25). This optimisation test suggested the optimal induction time for transport activity of PucI is approximately between 1 and 2 hours post-induction (Figure 3.25), and optimal expression of PucI is achieved when cells were induced for 1.0 to 4.5 hours with 0.5 mM IPTG (Figure 3.26).

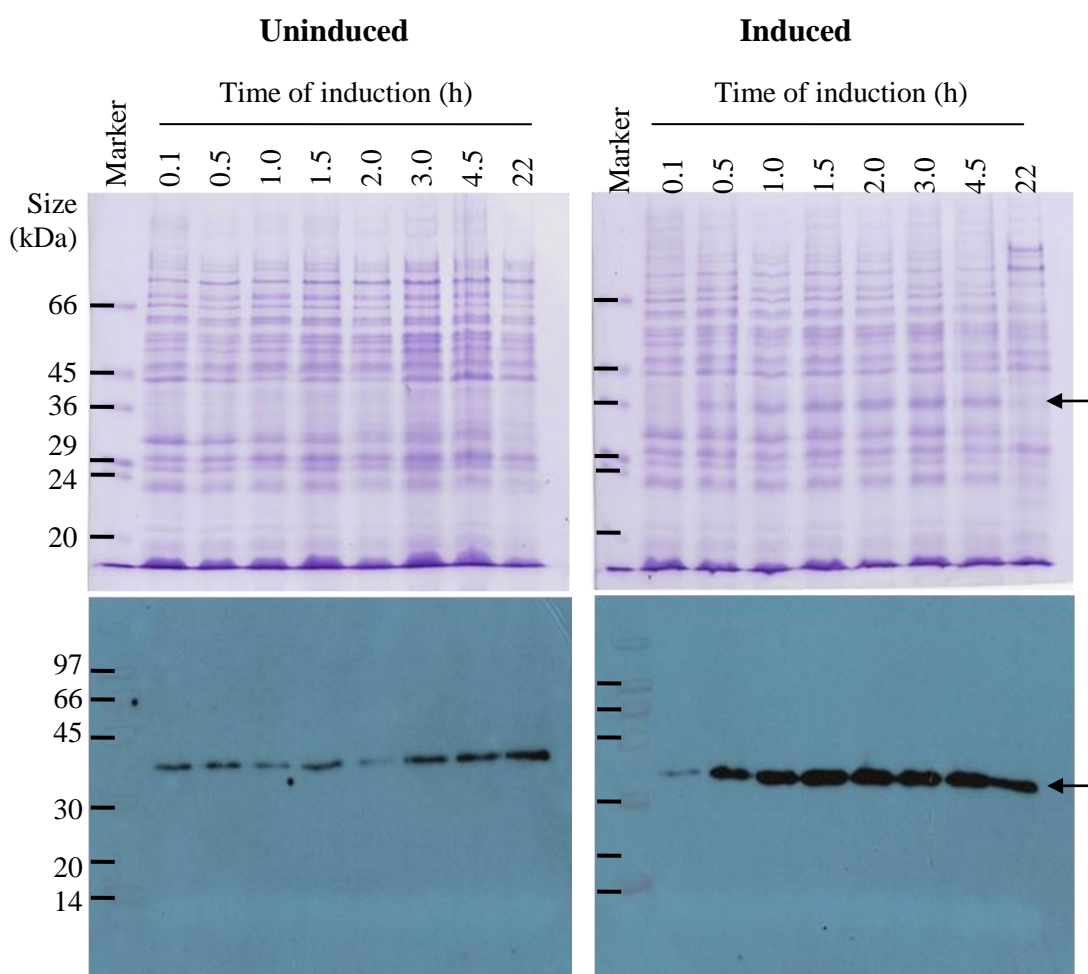


Figure 3.26 SDS-PAGE and Western blot analysis of total membranes of BL21(DE3) pTTQ18(*pucI*)-His₆ cells induced for different periods of time. BL21(DE3) pTTQ18(*pucI*)-His₆ cells were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol, and 0.5 mM IPTG was added when $A_{680nm} = 0.4 - 0.6$. Cells were harvested between 0.1 and 22 hours post-induction. Uninduced cells were grown according to the induced cells, but no IPTG was added. Total membranes were prepared and analysis by SDS-PAGE and Western blotting. Arrow indicates position of PucI.

3.3.6. Assessing the effect of sodium ions on the uptake of allantoin by PucI

To test whether the transport of allantoin by PucI is sodium-dependent, a range of NaCl concentrations in the transport assay buffers was tested. The salt contents of the buffer were adjusted with KCl and kept constant at 150 mM for all experiments. The uptake of allantoin in a range of NaCl concentrations, in cell expressing PucI was 10.87 to 12.8 nmol mg⁻¹ cells at 15 seconds and 35.7 to 39.4 nmol mg⁻¹ cells at 2 minutes (Figure 3.27). Results indicate the transport of allantoin by PucI is not greatly affected by presence of sodium (Figure 3.27).

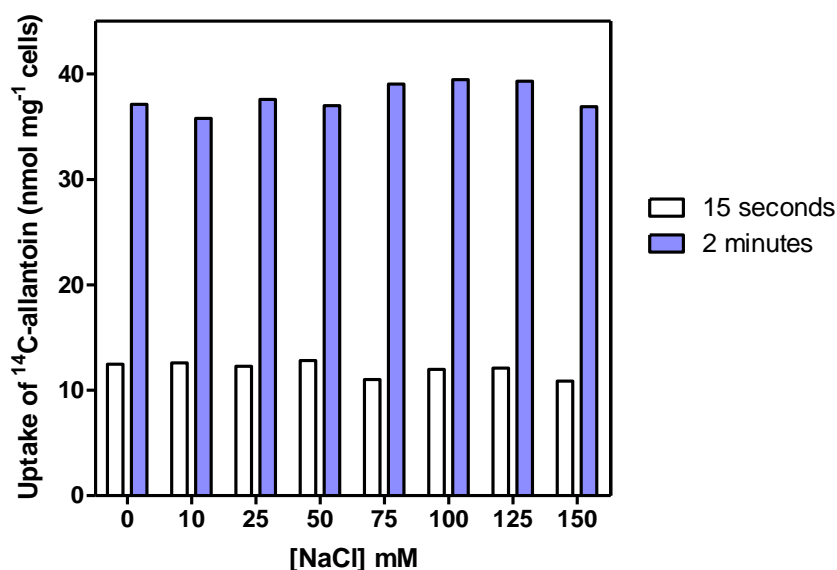


Figure 3.27 Effect of sodium ions on allantoin uptake by cells expressing PucI. *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. Cells were induced with 0.5 mM IPTG when A_{680nm} = 0.4 – 0.6, and harvested 1 hour post-induction. Harvested cells were washed and energised with 20 mM glycerol, and incubated in 50 µM ¹⁴C-allantoin, 5 mM MES pH 6.6 and 0 – 150 mM NaCl for uptake studies. Samples were taken 15 seconds and 2 minutes post-addition of radiolabelled allantoin. Results represent means of duplicates.

3.3.7. Substrate specificity of PucI

To study the substrate specificity of PucI, ¹⁴C-allantoin uptake in *E. coli* cells expressing PucI was determined in the presence of a ten-fold molar excess of potential substrates. This competition assay showed a range of substrates was able to inhibit the transport of allantoin into the cell (Figure 3.28). Competition with an excess of unlabelled allantoin caused a massive reduction of the uptake at 15 seconds and 2 minutes, to 3 and 4 % respectively (Figure 3.28), confirming that PucI mediates the

transport of allantoin. Hydantoin and other hydantoin derivatives, such as benzylhydantoin and hydroxyhydantoin, also reduced allantoin uptake by PucI to approximately 8-36% (Figure 3.28), suggesting these are also possible substrates transported by PucI. The hydantoin moiety and the –NH group of these substrates must therefore be important for their binding to PucI (Figure 3.29). The common structural features are highlighted in the diagram (Figure 3.29). To a lesser degree excessive thymine, allantoinic acid, hypoxanthine and uracil also caused a reduction of allantoin uptake (29-57%) (Figure 3.28). These substrates contain a six-membered ring, but still share some common structural features with allantoin and the hydantoin derivatives (Figure 3.29), further indicating the importance of the highlighted features of the substrates (Figure 3.29).

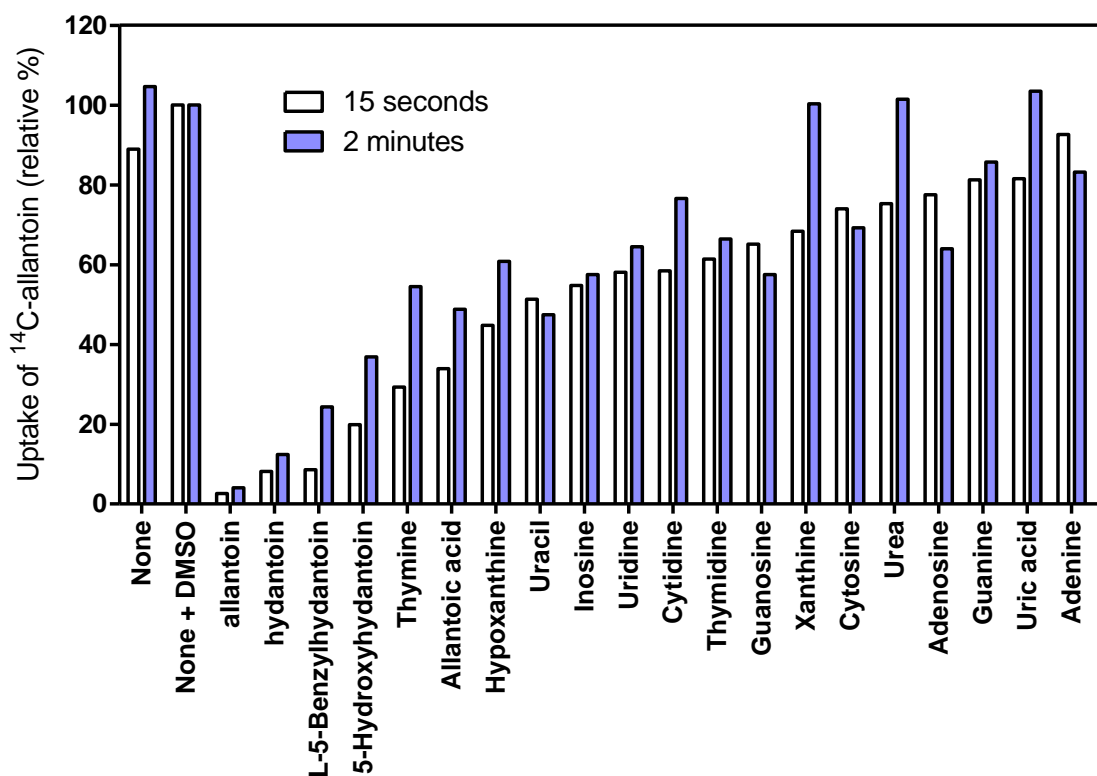


Figure 3.28 Substrate specificity of PucI. Competition of ^{14}C -allantoin ($50\ \mu\text{M}$) uptake into *E. coli* BL21(DE3) cells expressing PucI in the presence of a ten-fold molar excess of potential competitors. The noncompeted uptake rate was taken as 100% corresponding to 14.6 and 30.4 nmol allantoin per mg of cells for 15 seconds and 2 minutes post-addition of ^{14}C -allantoin, respectively. Buffer contains 2% DMSO. Results represent means of two experiments. None = no competitors or DMSO.

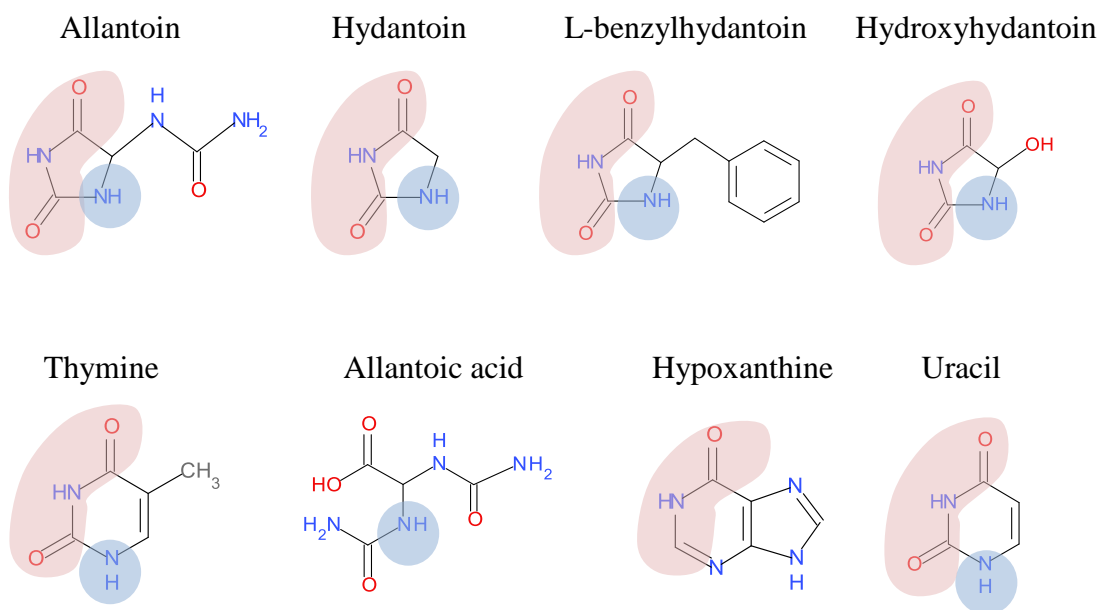


Figure 3.29 Common structural features of PuclI substrates. Arranged in the order of effectiveness in inhibiting transport of allantoin by PuclI.

The substrate specificity results suggest that besides allantoin, PuclI binds hydantoin, the hydantoin derivatives tested and allantoic acid, but not the purines guanine and adenine, or precursors of allantoin synthesis xanthine and uric acids, except for hypoxanthine (Figure 3.20 and 3.28).

3.3.8. Kinetic analysis of allantoin transport by PuclI

A kinetic study of PuclI was performed using *E. coli* BL21(DE3) cells transformed with pTTQ18(*puclI*)-His₆. Cells were incubated in various concentrations of ¹⁴C-allantoin (12.5 – 500 μM) in 5 mM MES pH 6.6 and 150 mM KCl with 20 mM glycerol for energisation. Samples were taken 15 seconds post-addition of ¹⁴C-allantoin, which is the initial rate of uptake. Transport of allantoin by *E. coli* BL21(DE3) cell expressing PuclI is concentration dependent, showing a K_m for allantoin of 24.44 ± 3.072 μM in the induced cells (Figure 3.30), and a V_{max} of 14.83 ± 0.4592 nmol mg⁻¹ cells (Figure 3.30). Uninduced cells demonstrated leaky expression of PuclI (Figure 3.26), which also displayed saturation kinetics (K_m of 15.36 ± 2.970 μM and V_{max} of 3.071 ± 0.1216 nmol mg⁻¹ cells (Figure 3.30).

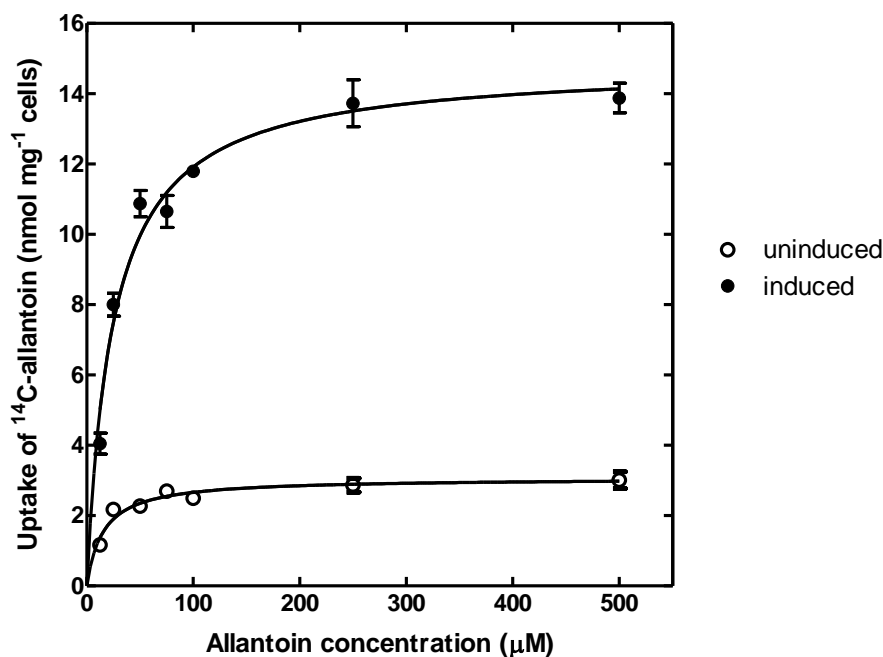


Figure 3.30 Michaelis-Menten kinetics of ¹⁴C-allantoin uptake by PucI. *E. coli* BL21(DE3) cells transformed with pTTQ18(*pucI*)-His₆ were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4 – 0.6, the induced sample received 0.5 mM IPTG and cells were harvested 1 hour post-induction, followed by washing. Cells were energised with 20 mM glycerol, and incubated in various concentrations of ¹⁴C-allantoin (12.5 – 500 µM) in 5 mM MES pH 6.6 and 150 mM KCl for kinetic studies. Samples were taken 15 seconds post-addition of ¹⁴C-allantoin. Results represent means of triplicates (±SEM).

3.3.9. Purification of PucI from the inner membrane

A six litre culture of *E. coli* BL21(DE3) cells expressing the His₆-tagged PucI were grown aerobically, described in Section 2.4.1. The cells harvested were lysed by explosive decompression and the total membrane obtained was separated on sucrose density gradients (Section 2.5.1). The membrane fractions obtained were analysed by SDS-PAGE, and showed the presence of PucI in the prepared inner membrane (Figure 3.31). Some PucI was also found in the mixed membrane, and a little in the outer membrane (Figure 3.31), indicating reasonable separation of the membranes. In section 3.1.4, the expression test showed PucI is present at 5% in the total *E. coli* membrane (Figure 3.6 and Table 3.3). Here, after the sucrose gradient step, PucI is present at 10% in the inner membrane, and 5% and 2% in the mixed and outer membrane respectively (Figure 3.31), as determined by scanning densitometry. This indicates the sucrose gradients step is important for concentrating PucI in the membrane.

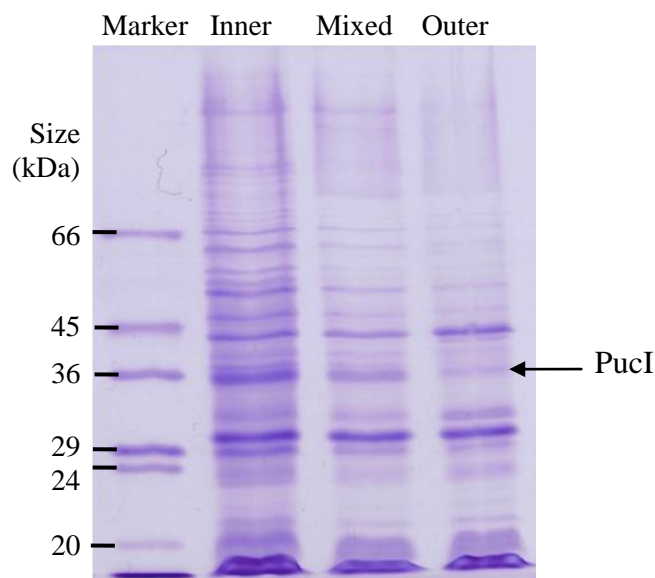


Figure 3.31 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) cells expressing PucI. Total *E. coli* membranes were prepared from *E. coli* BL21(DE3) pTTQ18(*pucI*)-His₆ grown on LB medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes were separated on 25-55% sucrose gradients and membrane fractions extracted from the gradients were analysed by SDS-PAGE. The black arrow indicates the position of the PucI protein.

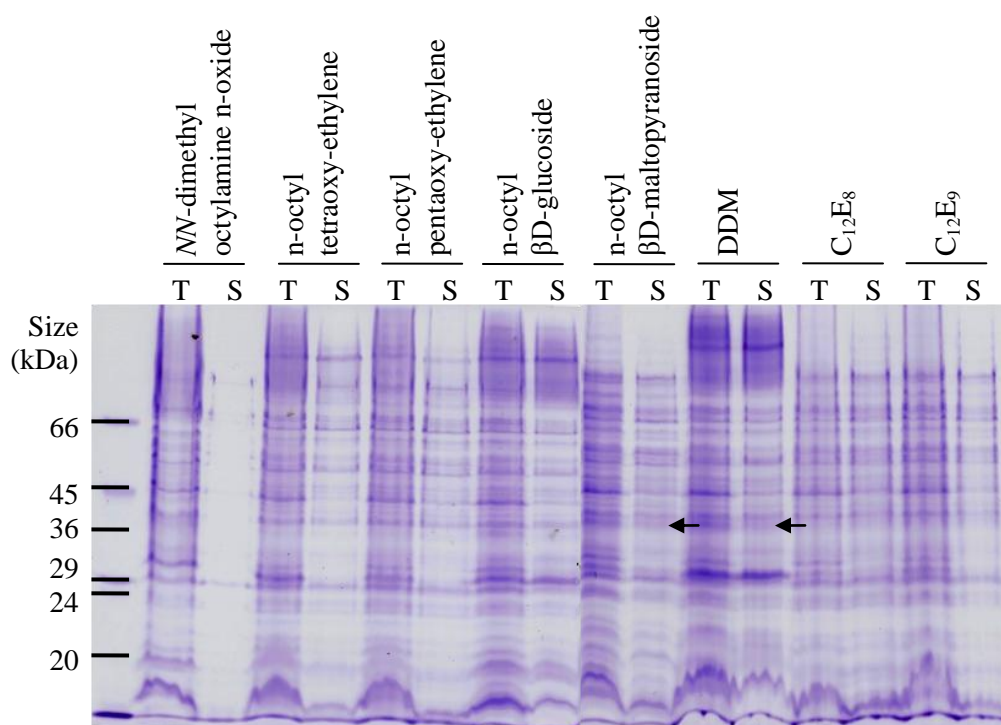


Figure 3.32 SDS-PAGE of total and solubilised inner membranes of *E. coli* BL21(DE3) cells expressing PucI. Total *E. coli* inner membranes transformed with pTTQ18(*pucI*)-His₆ were tested for solubilisation in eight detergents. Membranes were solubilised in 1% detergent for 1 hour and subjected to centrifugation to remove the insoluble fraction (Section 2.5.2). Total inner membranes (T) and soluble fractions (S) were then analysed by SDS-PAGE. Arrows indicate position of PucI.

The inner membranes were subjected to solubilisation trials, described in Section 2.5.2. *NN*-dimethyloctylamine *N*-oxide is a poor detergent for PucI (Figure 3.32). *N*-octyl- β D-maltopyranoside and *N*-dodecyl- β D-maltopyranoside (DDM) were the best for solubilising PucI (Figure 3.32). Therefore purification was performed using DDM (Section 2.5.3). Fractions from the purification of PucI were analysed by SDS-PAGE and Western blotting (Figure 3.33); the majority of PucI was solubilised using 1% DDM, and only a very small amount of PucI remains in the insoluble fraction (Figure 3.33). Almost all the PucI was successfully bound to the Ni-NTA resin, indicated by the faint band of PucI in the unbound fraction (Figure 3.33).

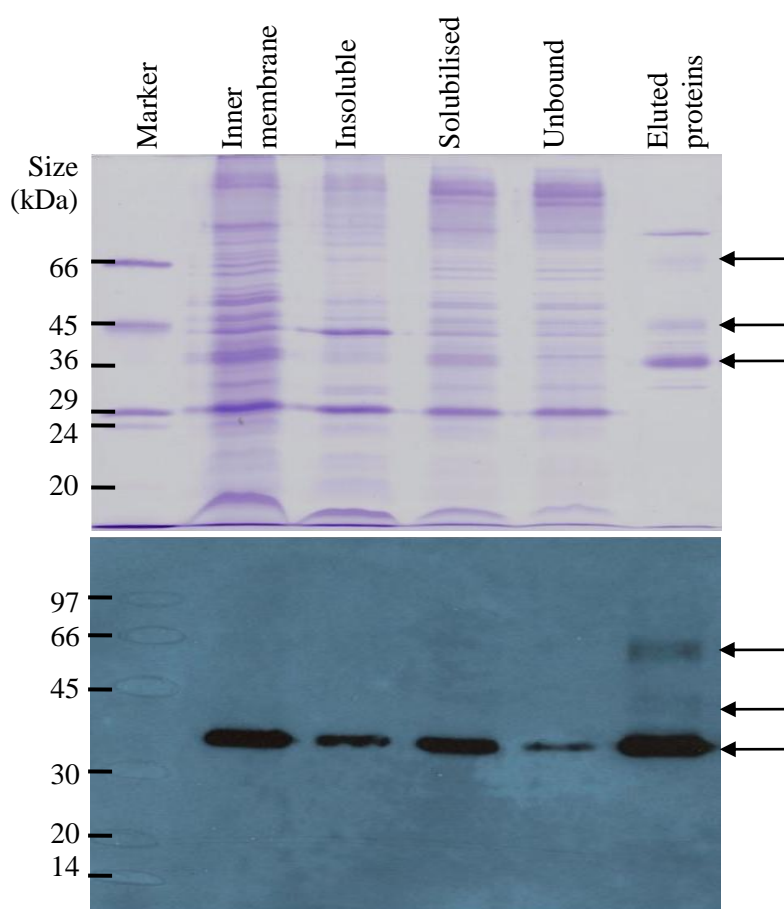


Figure 3.33 SDS-PAGE and Western blot of purification of *B. subtilis* PucI. PucI was purified from the inner membrane of *E. coli* BL21(DE3) pTTQ18(*pucI*)-His₆ cells. The inner membranes were solubilised in 1% DDM. Fractions (15 μ g), and eluted fraction (10 μ g) from the purification were analysed by SDS-PAGE (top) and Western blotting (bottom). The arrows indicate the position of the PucI protein.

PucI was purified to 77%, determined by scanning densitometry, and migrates at 37 kDa and 66 kDa on the SDS-PAGE gel (Figure 3.33). A contaminant was co-eluted with PucI, migrating at >66 kDa on the SDS-PAGE. The >66 kDa protein is commonly recognised as an AcrB contaminant in *E. coli* expression systems, and could possibly be removed by size-exclusion chromatography.

3.3.10. Testing the integrity of PucI by circular dichroism

The secondary structure of PucI was examined by circular dichroism (Section 2.6.1). The negative indentation at 208 and 222 nm, and a positive signal at ~ 190 nm, are indicative of a predominantly α -helical content (Figure 3.34). This agrees with the topology analysis by TMHMM, which predicts twelve transmembrane helices (Figure 3.21), and confirms retention of the secondary structure of PucI during the purification process.

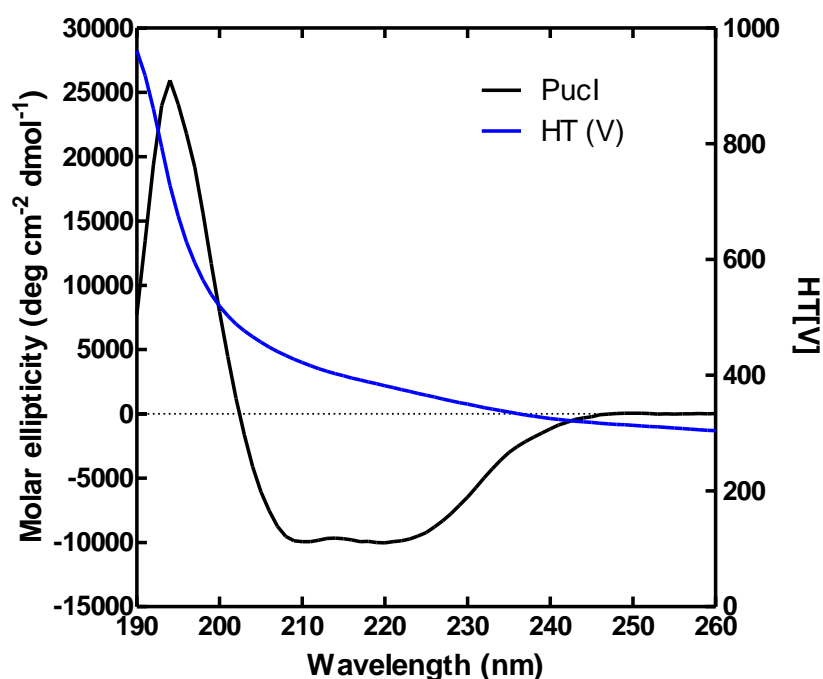


Figure 3.34 Circular dichroism spectrum of purified PucI of *B. subtilis*. Purified PucI (0.05 mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of PucI was performed using Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample were analysed in Hellma quartz-glass cell of 1mm path length. Spectrum was recorded with 1 nm step resolution at a scan rate of 10 nm/min. Response time was set at 1 second with a sensitivity of 20 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of ten scans, from which the buffer contribution was subtracted. The blue line represents the voltage applied to the photomultiplier.

3.3.11. Discussion: characterisation of PucI

The *pucI* gene of *B. subtilis* encodes an allantoin transporter (Danielsen *et al.*, 1992), important in providing the bacterium with a nitrogen source when the preferred nitrogen sources such as glutamine and ammonium, or glutamate, are depleted (Fisher, 1999). *pucI* is one of the 14 *puc* genes (Figure 3.19) that constitute the purine catabolism pathway in *B. subtilis* (Schultz *et al.*, 2001), allowing the bacterium to utilise a wide range of low molecular weight compounds as nitrogen source (Schultz *et al.*, 2001) (Figure 3.20).

In this study, the phylogenetic tree of the selected NCS1 proteins showed the PucI protein clustered closely with b0511 and EF3000 (Figure 3.2). The b0511 and EF3000 proteins were tested for allantoin uptake in Section 3.1.6, which revealed that these were not transporters for allantoin under the test conditions (Table 3.4). The conserved residues in PucI are highlighted in Figure 3.21. These are in: the N-terminus (K26 and R27); helix I (I36, W37, M38 and G39); the periplasmic loop between helices I and II (G57 and L58); the cytoplasmic loop between helices II and III (G85, K87, G89, L90, R97, Y100, G101, G104, A105, I107 and P108); helix III (W119, G121 and I122); the cytoplasmic loop between helices IV and V (G176 and I180); the cytoplasmic loop between helices VI and VII (D249, F250, T251, R252 and F253); helix X (F386, L387, G399 and V400); and in the cytoplasmic loop between helices X and XI (M402, D404, Y405, I408, R409, K410, V415, L418 and Y419) (Figure 3.21 and Appendix 2).

Based on the structures of Mhp1 (Weyand *et al.*, 2008) and the alignment in Appendix 2, potential residues in PucI that could be involved in substrate recognition and binding were identified. The W117 in Mhp1 aligns with W119 in helix III of PucI and therefore this residue could potentially form π -stacking with the hydantoin moiety of allantoin, hydantoin, L-benzylhydantoin and hydroxyhydantoin (Figure 3.28). Also, W240 in helix VI of PucI aligns with W220 in Mhp1 and could suggest this residue may form π -stacking with the aromatic moiety of L-benzylhydantoin (Appendix 2). Residues S323 and V324 in PucI could be involved in cation binding since these align with S312 and T313 of Mhp1 which binds sodium (Appendix 2). Further investigations, e.g. by producing PucI mutants, will be needed to completely understand the function of these residues.

Previous studies demonstrated that mutants in the *pucI* gene were unable to grow on allantoin and therefore it could encode for an allantoin transporter (Schultz *et al.*, 2001). Here, cloning of the *pucI* gene into plasmid pTTQ18-His₆ (Figure 3.2 to 3.5)

and transformation into *E. coli* BL21(DE3) cells (Figure 3.6), followed by transport assay, demonstrated PucI does transport allantoin (Table 3.4 and Figure 3.24). The induction time test showed the optimal transport activity of PucI is between 1 and 2 hours after induction with IPTG (Figure 3.25 and 3.26). The transport characteristic of PucI were also investigated, showing K_m and V_{max} of $24.44 \pm 3.072 \mu\text{M}$ and $14.83 \pm 0.4592 \text{ nmol mg}^{-1} \text{ cells}$, respectively (Figure 3.30), and indicating that transport activity, under the test condition is not sodium-dependent (Figure 3.27). PucI is able to interact with a wide range of substrates, being most specific for allantoin, hydantoin and hydantoin derivatives, therefore suggesting the hydantoin moiety is important for substrate recognition by PucI (Figure 3.28 and 3.29). Further tests are required to determine if these compounds are transported by, or are inhibitors of, allantoin uptake by PucI. Therefore radiolabelled forms of these compounds are needed. The transport of allantoin by PucI could also be tested for pH-dependence by preincubation of vesicles with 2,4-dinitrophenol, a proton ionophore that can shuttle protons across biological membranes, uncoupling oxidative phosphorylation, and hence destroying active transport.

PucI was overexpressed to 5% in the total *E. coli* membranes, migrating at 37 kDa on the SDS-PAGE gel (Figure 3.6 and Table 3.3). Tests for the expression of PucI in LB, 2TY and M9 minimal media suggested LB was the best for expressing PucI (Figure 3.23). After the sucrose density gradient step, PucI increased from 5-6% in the total membrane to 10% in the inner membrane, successfully concentrating PucI (Figure 3.31). The expression of PucI could be further optimised by testing in various IPTG concentrations, and different expression hosts and plasmids with different promoter or features, and using different affinity tags for purification. The circular dichroism confirmed the integrity of the secondary structure of PucI (Figure 3.34), and the transport assay verified the expressed protein is active (Figure 3.24 to 3.29). PucI was purified to 77% using Ni-NTA, and migrates at 37 and 66 kDa on the SDS-PAGE gel and Western blotting (Figure 3.33). Further optimisation of the purification, using gel filtration and varying the buffer components, might further increase the purity of PucI. The purified protein could be reconstituted into liposomes (Ward *et al.*, 2000), for functional tests, as well as to test accurately for cation-dependence. Crystallisation could then be undertaken to elucidate the structure of PucI, which might help to understand the structure-activity relationship of the NCS1 family of transporters, and the fundamentals of the alternating access model.

3.4. Characterisation of a *Pseudomonas aeruginosa* uracil transporter (PA0443)

3.4.1. Introduction – *Pseudomonas aeruginosa* and PA0443

Pseudomonas aeruginosa is a versatile Gram-negative bacterium that grows in soil, coastal marine habitats, and on plant and animal tissues. It is one of the top three causes of opportunistic human infections, including bacteraemia in burn victims, urinary-tract infections in catheterised patients and hospital-acquired pneumonia in patients on respirators. The emergence of this bacterium as a major opportunistic human pathogen could be a consequence of its resistance to the antibiotics and disinfectants that eliminated other environmental bacteria. The whole genome sequence of this bacterium was completed (Stover *et al.*, 2000), showing high proportion of regulatory genes and a large number of genes involved in the catabolism, transport and efflux of organic compounds, as well as potential chemotaxis systems, which may give rise to the environmental adaptability of this bacterium (Stover *et al.*, 2000).

Here the PA0443 protein of *P. aeruginosa* is studied, for its high similarity and identity to the Mhp1 protein (Table 3.1), although of unknown function. PA0443 lies next to a putative N-carbamoyl-beta-alanine amidohydrolase (*hyuC*) gene, also referred to as beta-ureidopropionase (Figure 3.35). The transcription start and termination sites are not yet annotated in this part of the *Pseudomonas* genome and are therefore not labeled in the diagram (Figure 3.35).

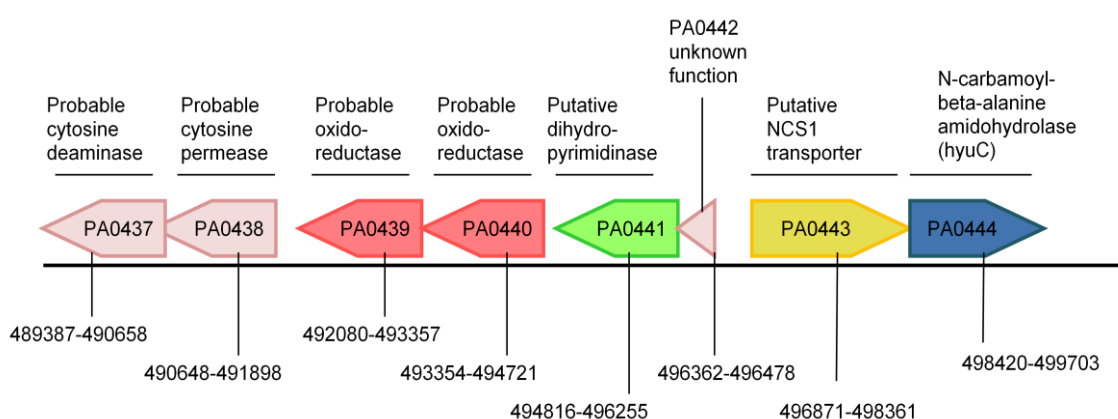


Figure 3.35 Map of the *P. aeruginosa* PA0443 gene region. Each arrow represents a gene, and the number below corresponds to the genome position of the gene. The diagram is modified from the *Pseudomonas* genome database (Stover *et al.*, 2000).

As can be seen, the PA0438 (putative cytosine transporter), which was also studied in this NCS1 transporters project (Section 3.1), lies very close to the PA0443 protein, and both are located near the putative purine catabolism genes of *P. aeruginosa* (Figure 3.35), important in providing an alternative nitrogen source for the bacterium. In *Pseudomonas*, adenine, hypoxanthine, xanthine and guanine are broken down to allantoin by the enzymes adenine deaminase, guanine deaminase, xanthine dehydrogenase and uricase (Bongaerts *et al.*, 1997), the same as in *B. subtilis* (Figure 3.20). Pyrimidine catabolism in *P. aeruginosa* is induced by uracil, and this bacterium was able to utilise uracil and thymidine as well as their respective reductive catabolic products as sole sources of nitrogen, indicating the bacterium contains genes for the catabolism and utilisation of these compounds (Kim & West, 1991). The reductive catabolic pathway enzymes (Figure 3.12) dihydropyrimidine dehydrogenase, dihydropyrimidinase and ureidopropionase were all detected in minimal medium grown *Pseudomonas aeruginosa* cells (Kim & West, 1999), suggesting *P. aeruginosa* is able to utilise pyrimidine as well as purine, as sole sources of nitrogen. This section studies the PA0443 protein of *P. aeruginosa*, a NCS1 family protein shown to transport uracil (Section 3.1.6). Firstly, the topology and conserved residue of this transporter are analysed, followed by purification of the cloned His₆-tagged protein, and kinetic and substrate specificity studies.

3.4.2. The *P. aeruginosa* uracil transporter (PA0443): topology analysis and conserved residues

Topology prediction by TMHMM (Krogh *et al.*, 2001) showed PA0443 is highly hydrophobic, consisting of twelve transmembrane α -helices. A topology model of PA0443 is shown in Figure 3.36, with the conserved residues highlighted, which are positioned mostly in helices I, II and X, as well as in the cytoplasmic loops between helices II and III, helices IV and V, and helices X and XI (Figure 3.36). PA0443 is 29.3% identical and 59.8% similar to Mhp1 (Table 3.1), and clustered most closely with Pden1111 and SCO6417 on the phylogenetic tree (Figure 3.2). Both PA0443 and Pden1111 were shown to transport uracil in Section 3.1.6. (Table 3.4), therefore suggesting SCO6417 could also be an uracil transporter.

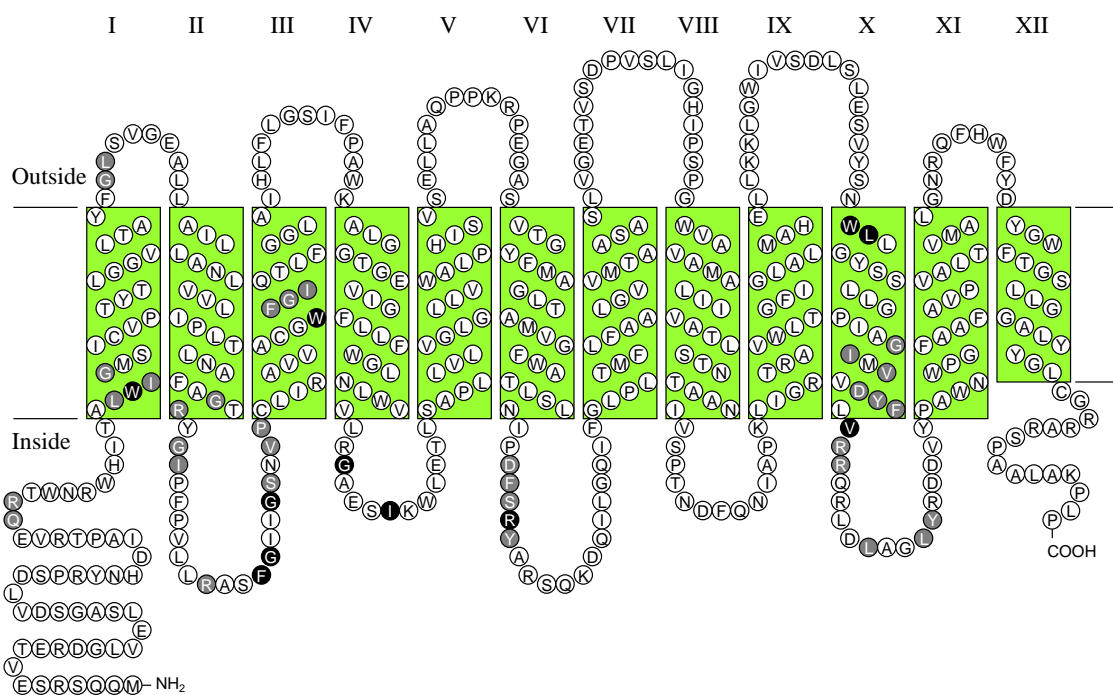


Figure 3.36 Topology model of *P. aeruginosa* uracil permease (PA0443). Topology prediction by TMHMM. The predicted helices are numbered from the N- to C-terminus and are shown in roman numerals. Conserved residues in the NCS1 family proteins are highlighted in black and grey according to the alignment in Appendix 2.

3.4.3. Cloning and expression of PA0443

In order to study the PA0443 transporter of *P. aeruginosa*, the PA0433 gene was amplified from the genomic DNA of *P. aeruginosa* PAO1 strain by PCR (Figure 3.2), and cloned into plasmid pTTQ18-His₆ that introduces a C-terminal His₆-tag (Section 2.3.1). The clone was analysed by colony PCR and restriction digestion, confirming the correct gene in the plasmid construct (Figure 3.3 and 3.4). DNA sequencing revealed no mutations in the cloned PA0443 gene and the correct protein sequence (Appendix 3 and 4). Subsequently, the clone was transformed into *E. coli* BL21(DE3) cells showing overexpression of the PA0443 protein in the IPTG-induced *E. coli* total membranes, migrating at 41 kDa on the SDS-PAGE gel (Figure 3.6 and Table 3.3).

3.4.4. Optimisation of expression of PA0443

Expression tests showed PA0443 of *P. aeruginosa* was expressed to 24% in the total *E. coli* membrane (Figure 3.6 and Table 3.3). The expression of PA0443 was optimised by testing in 2TY and minimal media, as well as in a range of IPTG concentrations. Growth of the *E. coli* BL21(DE3) cells transformed with

pTTQ18(PA0443)-His₆ showed little effect when induced with IPTG (Figure 3.37). Cell growth was higher and similar in LB and 2TY, than in M9 minimal medium (Figure 3.37). The absorbance of the induced cultures at time of harvesting was 2.040, 2.345 and 0.913 for cells grown in LB, 2TY and minimal media, respectively (Figure 3.37). Harvested cells were subjected to total membrane preparation and analysis by SDS-PAGE and Western blotting, showing the presence of the His₆-tag PA0443 in the induced cells, and absence in the uninduced cells (Figure 3.38), confirming tight regulation of the expression of PA0443. Scanning densitometry showed PA0443 was expressed to 10% and 26% in the total membranes of cells grown in 2TY and minimal medium, respectively. Given the expression of PA0443 is highest in minimal medium, a range of IPTG concentrations was tested; all slowed growth (Figure 3.39A), and scanning densitometry indicated induction with 0.5 mM IPTG is optimal for the expression of PA0443 in minimal medium (Figure 3.39B). The cells grown in LB medium start to reach stationary phase towards the time of harvesting, whereas cells grown in minimal medium are still in the exponential phase (Figure 3.37); therefore, further tests in minimal medium to test for longer induction times is required to determine the optimal expression level of PA0443, i.e. harvesting cells 4, 6, 8 and 10 hours post-induction should be tested.

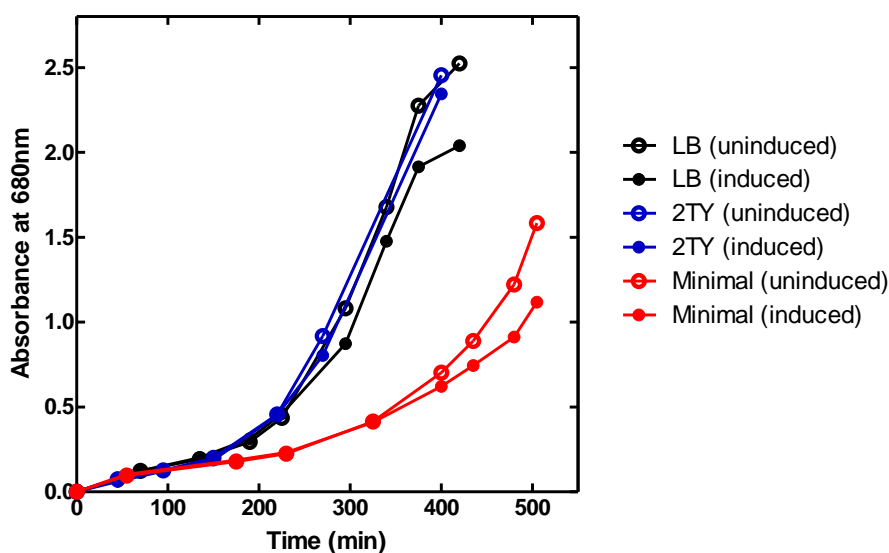


Figure 3.37 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ in LB, 2TY and M9 minimal media. *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ were grown in LB, 2TY and minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when A_{680nm} reached 0.4 – 0.6 and harvested 3 hours post-induction.

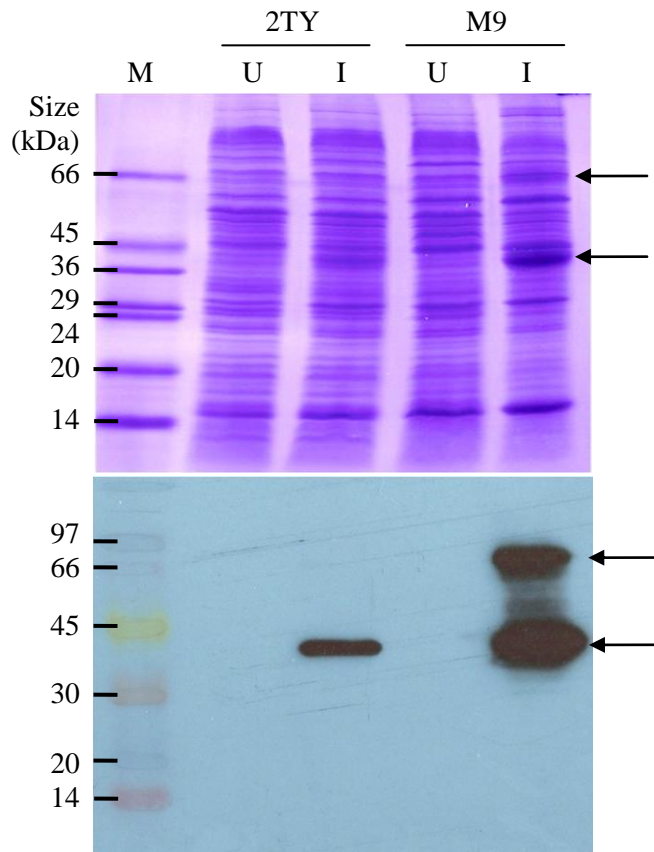


Figure 3.38 Expression of *P. aeruginosa* uracil permease (PA0443) in cells grown in 2TY and M9 minimal medium. BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ were grown in 2TY and M9 minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction, followed by preparation of total *E. coli* membranes and analysis by SDS-PAGE and Western blotting. M (marker), U (uninduced) and I (induced). Arrows indicate position of the PA0443 protein.

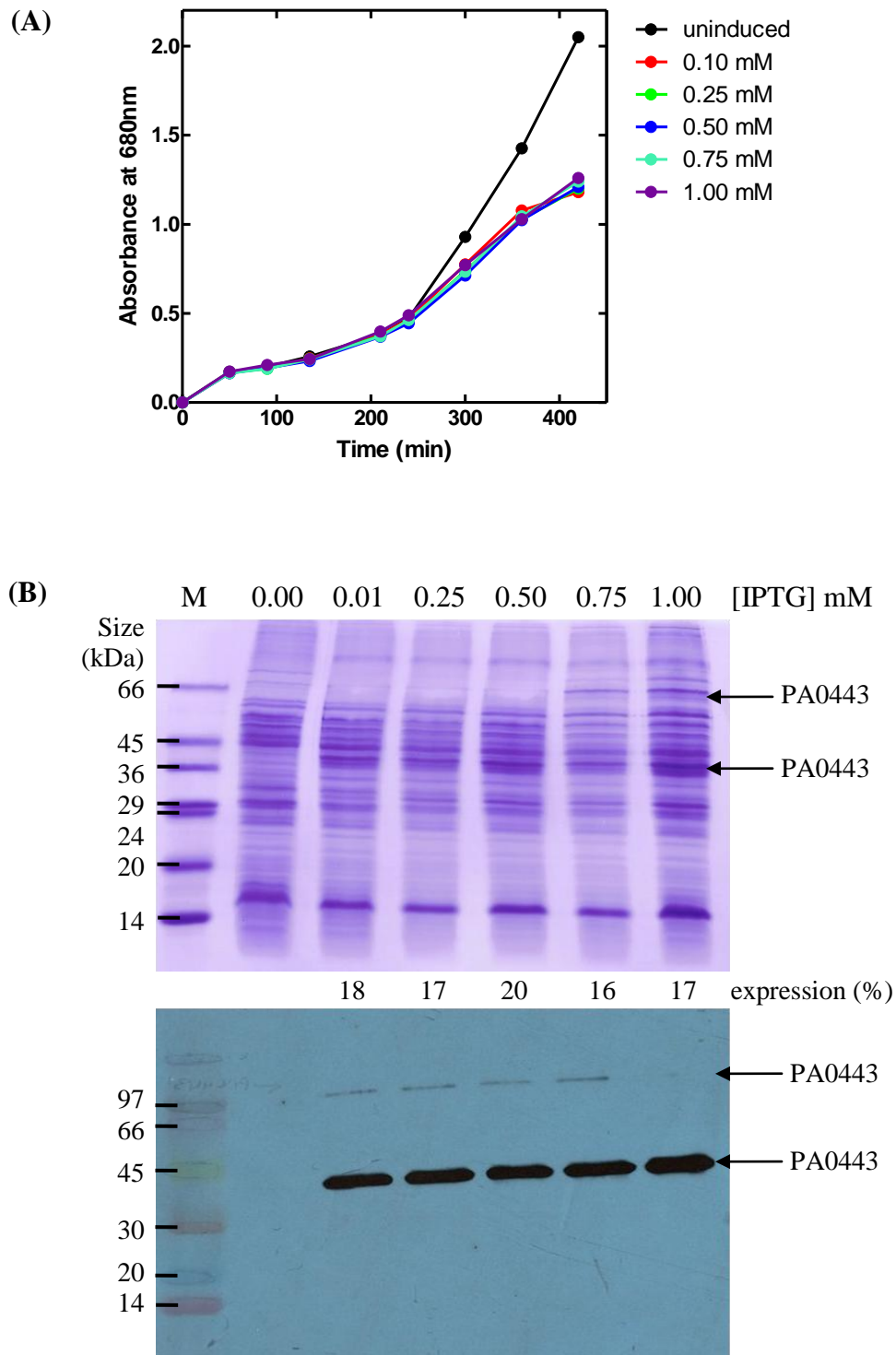


Figure 3.39 Growth and expression of *P. aeruginosa* uracil permease (PA0443) using a range of IPTG concentrations in minimal medium. *E. coli* BL21(DE3) pTTQ18(PA0443)-His₆ cells were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol, and samples were induced with a range of IPTG concentrations (0 – 1 mM) when $A_{680\text{nm}}$ reached 0.4 – 0.6, and harvested 3 hours post-induction. (A) Growth curves. Total membranes were prepared from the harvested cells and analysed by (B) SDS-PAGE (top) and Western blotting (bottom). Arrows indicate position of PA0443.

3.4.5. Uptake of radiolabelled uracil by PA0443

A time course of uracil transport by PA0443 showed an increase uptake of uracil in the induced cells expressing PA0443, up to a maximum of 2.0959 nmol mg⁻¹ cells, from 0 to 3 minutes (Figure 3.40). The amount of radioactivity then decreased to 0.8040 nmol mg⁻¹ cells, which could indicate metabolism of uracil inside the cell, followed by radiolabelled metabolites leaving the cell. The radiolabelled uracil used in this work is ³H-5,6-uracil. Uptake of uracil into the cell is either incorporated into RNA or metabolised in the reductive and oxidative pathways to products shown in Figure 3.12. The decrease in radioactivity in the cells could be due to excretion of radiolabelled water from the cells. At 5 to 10 minutes of the time course, an increasing amount of uracil is shown inside the cells, which could suggest that, when the cellular level of uracil becomes low, the cells start to take up uracil again (Figure 3.40). A similar uracil uptake pattern is shown in the uninduced cells (Figure 3.40), previously indicated by the SDS-PAGE gel and Western blotting not to express PA0443 (Figure 3.38 and 3.39B), therefore the activity observed may be the uracil transporter of *E. coli*, UraA, which was previously characterised and demonstrated to be necessary for uracil uptake at low exogenous uracil concentrations (Andersen *et al.*, 1995).

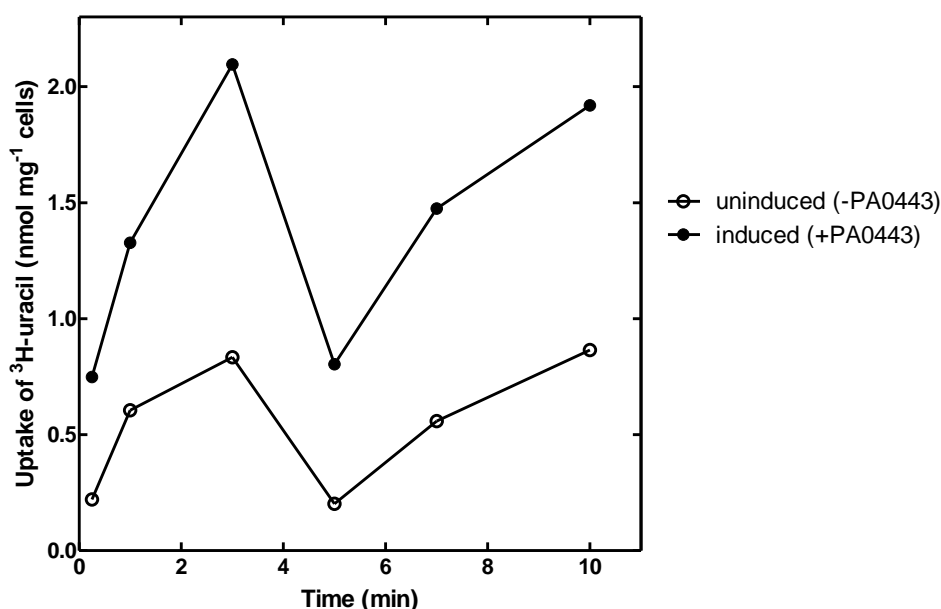


Figure 3.40 Time course of uracil uptake by PA0443. Uptake of ³H-uracil in *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆. Cells were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4-0.6, 0.5 mM IPTG was added to the induced sample and cells were harvested 1 hour post-induction, followed by washing of cells. Cells were energised with 20 mM glycerol, and incubated in 50 µM ³H-uracil in 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Results represent means of duplicates.

3.4.6. Substrate specificity of PA0443

The substrate specificity of PA0443 was studied using radiolabelled uracil and testing uptake in *E. coli* cells expressing PA0443, in presence of ten-fold molar excess of potential substrates (Figure 3.41). The competition assay showed PA0443 has a narrow substrate specificity, being specific for transport of uracil and thymine into the cell (Figure 3.41). Competition with an excess of unlabelled uracil caused a reduction of uptake to 20.4 and 50.1 % at 15 seconds and 2 minutes respectively (Figure 3.41), confirming that PA0443 mediates the uptake of uracil. Excessive thymine also caused reduction of radiolabelled uracil uptake to 29.9 and 50.3 % at 15 and 2 minutes respectively (Figure 3.41), confirming this could also be transported by PA0443.

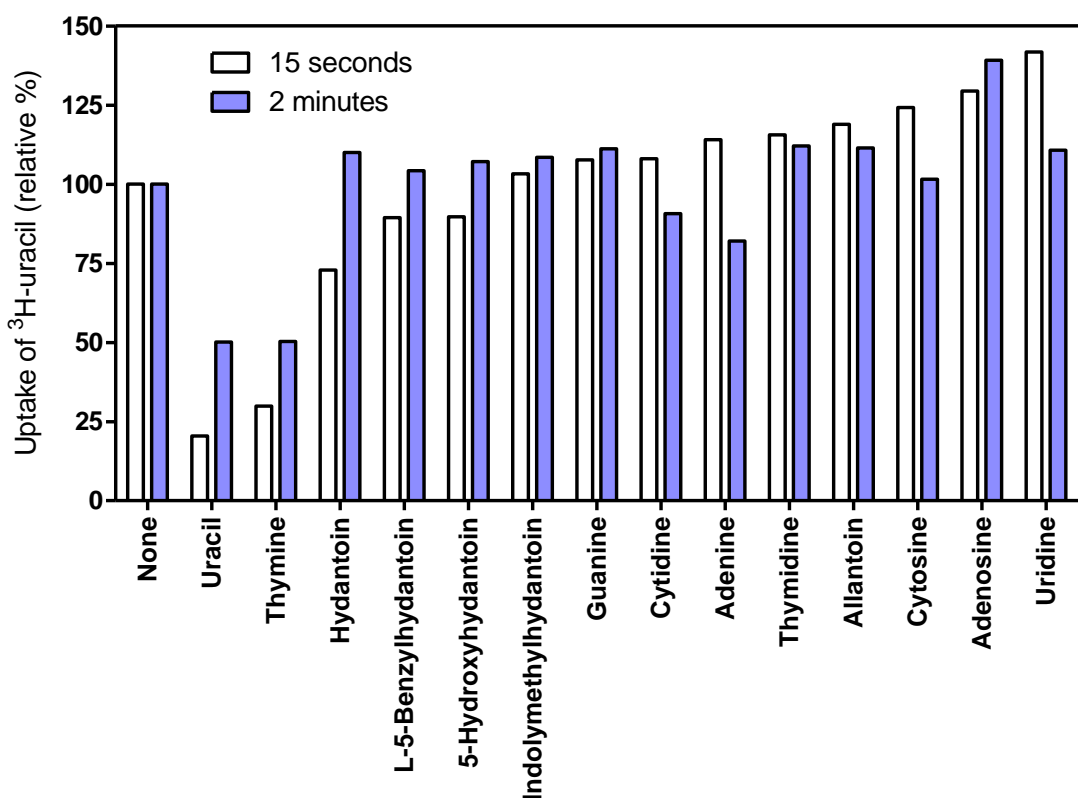


Figure 3.41 Substrate specificity of PA0443. Competition of ^3H -uracil ($50\ \mu\text{M}$) uptake into *E. coli* BL21(DE3) cells expressing PA0443 in the presence of a ten-fold molar excess of potential competitors. The non-competed uptake rate was taken as 100% corresponding to 0.18 and 0.25 nmol uracil per mg of cells for 15 seconds and 2 minutes post-addition of ^3H -uracil, respectively. Results represent means of two experiments.

Both uracil and thymine are six-membered ring compounds, and are very similar in structure (Figure 3.42). Thymine is also known as 5-methyluracil, derived by methylation of uracil at the 5th carbon. The two compounds inhibited uptake of the

radiolabelled uracil in the competition assay (Figure 3.41), suggesting the six-membered ring and the highlighted features in Figure 3.42 are important for transport by PA0443.

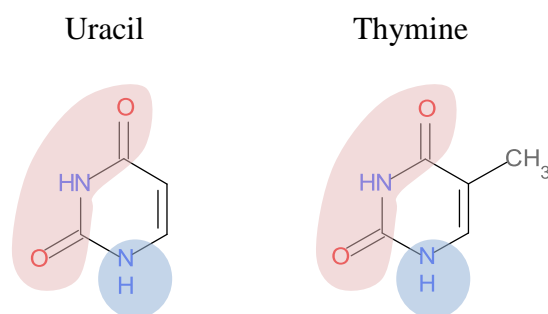


Figure 3.42 Common structural features of PA0443 substrates.

Hydantoin is structurally similar to uracil, but a five-membered ring, and excessive hydantoin in the competition assay caused a relative uptake of 72.9 and 110.1% (Figure 3.41), showing this substrate is not favoured for transport by PA0443, and indicates the size of substrate is important for transport by PA0443.

3.4.7. Kinetic analysis of uracil transport by PA0443

Kinetic studies of PA0443 were performed using *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆. Cells were incubated in various concentrations of ³H-uracil (40 – 400 μM) in 5 mM MES pH 6.6 and 150 mM KCl with 20 mM glycerol for kinetic study. Samples were taken at 15 seconds post-addition of ³H-uracil, which is the initial rate of transport. Transport of uracil by *E. coli* BL21(DE3) cells expressing PA0443, induced cells, is concentration dependent, showing a K_m for uracil of $26.72 \pm 3.219 \mu\text{M}$ (Figure 3.43), and displayed saturation kinetics with a V_{max} of $1.181 \pm 0.03442 \text{ nmol mg}^{-1} \text{ cells}$ (Figure 3.43). Uninduced cells did not express PA0443 (Figure 3.38 and 3.39), the kinetics observed in the uninduced cells (Figure 3.43) are likely to be the uracil transporter of the host cells, possibly activity of UraA of *E. coli* (Andersen *et al.*, 1995). As a control, uninduced and induced cells with plasmid pTTQ18-His₆ containing no PA0443 gene could be tested.

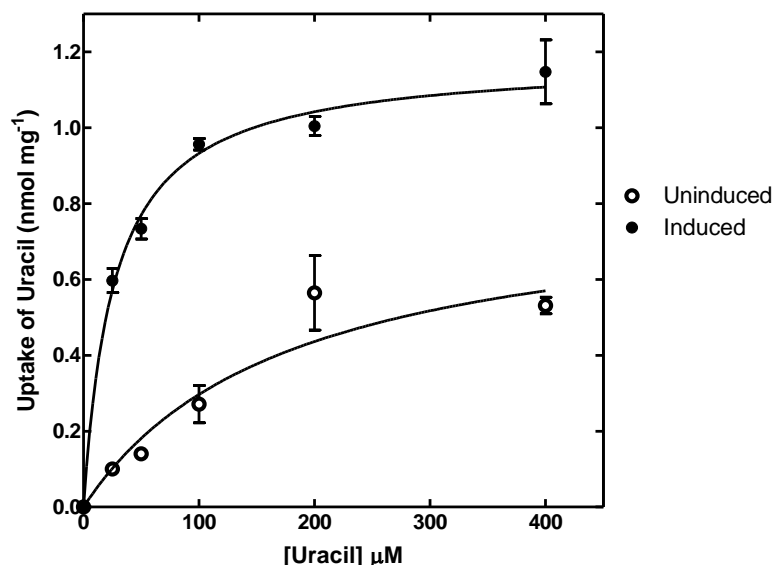


Figure 3.43 Michaelis-Menten kinetics of ³H-uracil uptake by PA0443. *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ were cultured in minimal medium in presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4-0.6, 0.5 mM IPTG was added to the induced sample and cells were harvested 1 hour post-induction, followed by washing. Cells were energised with 20 mM glycerol, and incubated in various concentrations of ³H-uracil (40 – 400 µM) in 5 mM MES pH 6.6 and 150 mM KCl for uptake studies. Samples were taken 15 seconds post-addition of ³H-allantoin. Results represent means of triplicates (±SEM).

3.4.8. Purification of PA0443 from the inner membrane

A six litre culture of *E. coli* BL21(DE3) cells expressing the His₆-tagged PA0443 was grown aerobically (Section 2.4.1). The cells harvested were lysed by explosive decompression and the total membrane obtained was separated on sucrose density gradients, described in Section 2.5.1. From the amino acid sequence and analysis by TMHMM, the PA0443 protein is predicted to be an integral membrane protein; therefore the membrane fractions obtained were analysed by SDS-PAGE and Western blotting, and showed the PA0443 protein enriched in the membrane fractions (Figure 3.44).

Cells lysed by explosive decompression were subjected to centrifugation to remove the cell debris from the cytosol + membrane fractions, followed by separation of the membrane from the cytosol fraction at high speed centrifugation, described in Ward *et al.* (2000). Using scanning densitometry, the percentage of expression of PA0443 in each cellular fraction was determined. The PA0443 protein is present at ~15% in the cell debris, and 9 % in the cytosol + membrane fraction (Figure 3.44); indicating loss of the PA0443 protein in the cell debris. The percentage of PA0443 in the cell

debris is high, but nevertheless, the loss of PA0443 is very little in comparison to the volume of cytosol + membrane fraction obtained. An intense band migrating at ~31 kDa in the cell debris, which is also present in the outer membrane fraction, suggest this protein band could be porin which is abundant in outer membranes of bacteria. The PA0443 protein is present in the outer, mixed and inner membranes at 22, 29 and 30%, respectively (Figure 3.44), so the fractionation was probably inefficient.

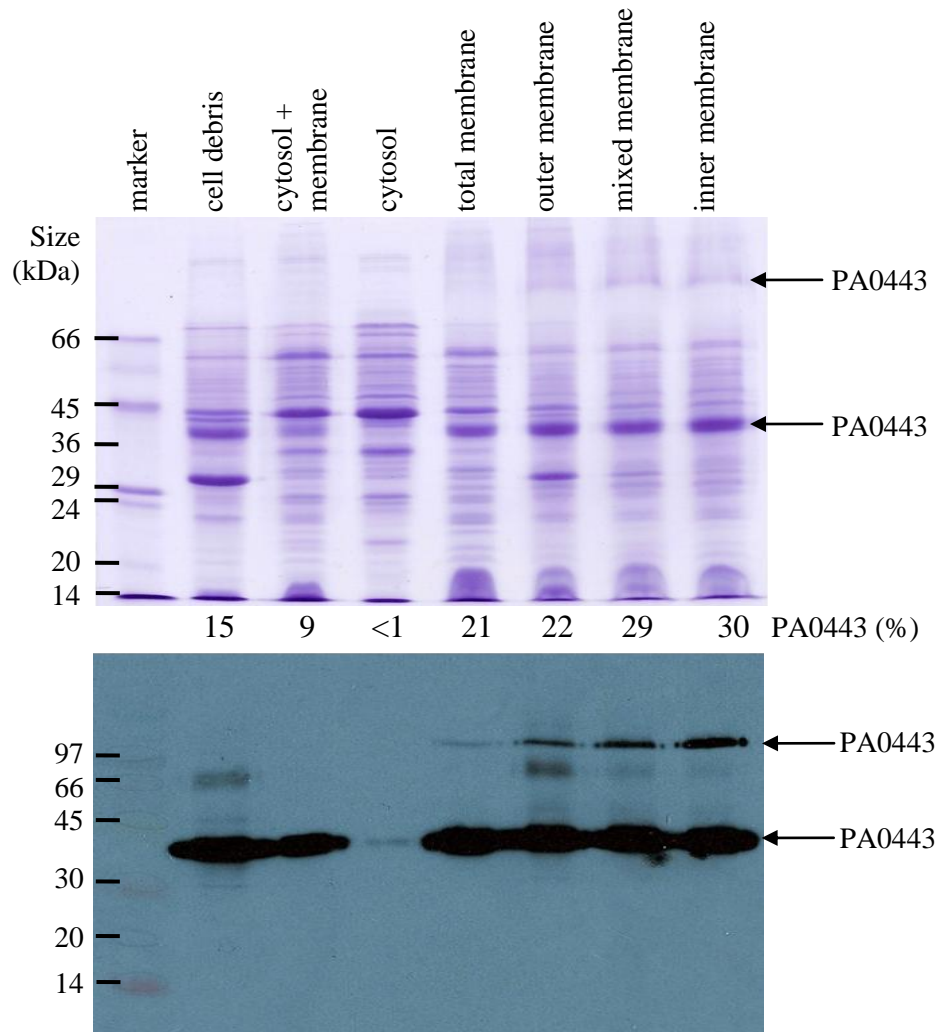


Figure 3.44 SDS-PAGE and Western blotting of fractions obtained from the preparation of inner membranes of *E. coli* expressing PA0443. *E. coli* BL21(DE3) cells transformed with pTTQ18(PA0443)-His₆ were grown on M9 minimal medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes prepared were separated on 25-55% sucrose density gradients and fractions from the preparation were analysed by SDS-PAGE and Western blotting. The arrows indicate position of PA0443.

Inner membranes expressing PA0443 were solubilised in 1% DDM and purification was performed using NiNTA, described in Section 2.5.3. Fractions from the purification were analysed by SDS-PAGE, showing poor solubilisation of the PA0443 protein (Figure 3.45). The PA0443 protein was successfully bound to the NiNTA resin, indicated by the absence or small amount of PA0443 in the unbound fraction (Figure 3.45). PA0443 was purified to 89%, determined by scanning densitometry, and migrates at 41 kDa on the SDS-PAGE gel (Figure 3.45).

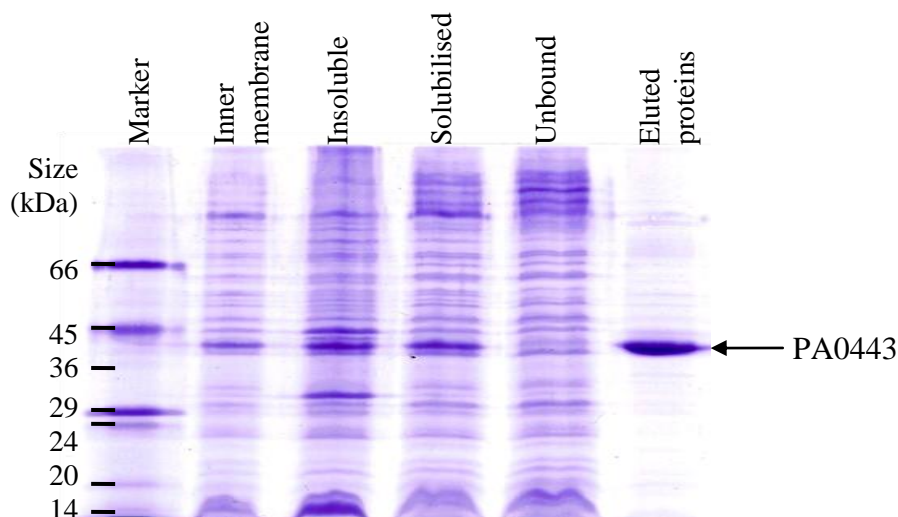


Figure 3.45 SDS-PAGE of purification of *P. aeruginosa* PA0443. PA0443 was purified from the inner membrane of *E. coli* BL21(DE3) pTTQ18(PA0443)-His₆ cells. The inner membranes were solubilised in 1% DDM. Fractions from the purification were analysed by SDS-PAGE. The arrow indicates the position of the PA0443 protein.

3.4.9. Testing the integrity of PA0443 by circular dichroism

The secondary structure of PA0443 was examined by circular dichroism (Section 2.6.1). The negative indentation at 208 and 222 nm, and a positive signal at ~190 nm, are indicative of a predominantly α -helical content (Figure 3.45), agreeing with the prediction by TMHMM that this protein is highly hydrophobic, confirming the retention of the secondary structure of PA0443 during purification.

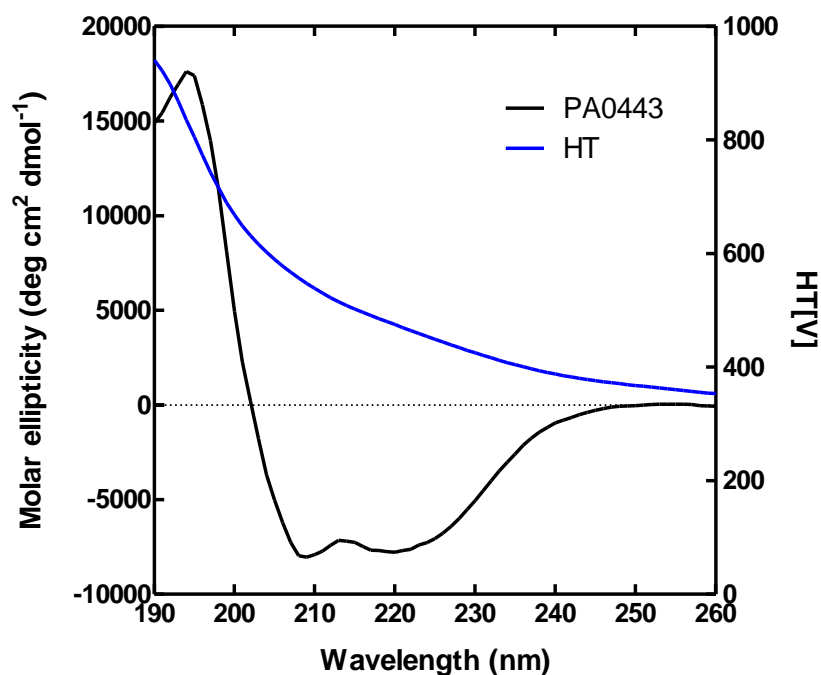


Figure 3.46 Circular dichroism spectrum of purified PA0443 of *P. aeruginosa*. Purified PA0443 (0.05 mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of PA0443 was performed using Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample were analysed in Hellma quartz-glass cell of 1mm path length. Spectrum was recorded with 1 nm step resolution at a scan rate of 20 nm/min. Response time was set at 1 second with a sensitivity of 50 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of thirty scans, from which the buffer contribution was subtracted.

3.4.10. Discussion: characterisation of PA0443

Prior to this study, the function of the PA0443 protein of *P. aeruginosa* was unknown, but this protein shows high identity and similarity to the Mhp1 protein (Table 3.1), potentially transporting nucleobases, nucleosides, allantoin and thiamine. It was therefore chosen as a target of the NCS1 family protein study. Here, the PA0443 gene was cloned into the membrane protein expression vector pTTQ18-His₆ and transformed into the expression host BL21(DE3) cells, successfully identifying this as an uracil and thymine transporter of *P. aeruginosa* (Section 3.1).

The PA0443 of *P. aeruginosa* clustered most closely to the Pden1111 and SCO6417 proteins on the phylogenetic tree of the selected NCS1 proteins (Figure 3.2). Pden1111 was identified to uptake uracil, adenine and thymidine in the signal screening test (Table 3.4). PA0443 transported uracil, but also thymine (Table 3.4 & Figure 3.41). The conserved residues in PA0443 are highlighted in Figure 3.36. These are in: the N-

terminus (Q40 and R41); helix I (L51, W52, I53 and G54); the periplasmic loop between helices I and II (G73 and L74); helix II and the periplasmic loop after it (G101, R103, G105, I106, R113, F116, G117, G120, S121, V123 and P124); helix III (W135, F136, G137 and I138); the cytoplasmic loop between helices IV and V (G184 and I188); the cytoplasmic loop between helices VI and VII (D256, F257, S258, R259 and Y260); and helix X and the cytoplasmic loop after it (W393, L394, G406, I407, V409, D411, Y412, F413, V415, R416, R417, L422, L425 and Y426) (Figure 3.36 and Appendix 2).

Based on the structures of Mhp1 (Weyand *et al.*, 2008) and the alignment in Appendix 2, the conserved residues in PA0443 that could potentially be involved in substrate recognition and binding were identified. For example, residues W135 and W247 in PA0443 align with residue W117 and W220 in Mhp1, respectively, suggesting these residues may potentially be involved in substrate binding, forming π -stacking with the six-membered ring substrates of PA0443. Residues in Mhp1 that were involved in cation binding were also found conserved in the PA0443 protein; these are S329, T330 and N331 in PA0443 (Appendix 2), suggesting PA0443 could potentially be cation dependent, and therefore this could be tested by reconstitution of PA0443 into liposomes and testing the uptake in presence of cations. Mutagenesis on these residues and testing the substrate specificity of the mutants will be needed to fully understand the functions of these residues.

In the substrate screening test, PA0443 was identified to transport uracil (Table 3.4). The transport characteristics of this transporter were therefore investigated, showing K_m and V_{max} of $26.72 \pm 3.219 \mu\text{M}$ and $1.181 \pm 0.03442 \text{ nmol mg}^{-1} \text{ cells}$, respectively (Figure 3.43). PA0443 has a narrow substrate specificity, being specific for the transport of uracil and thymine (Figure 3.41), showing the substrates transported by PA0443 are relatively small in size compared to CodB and PucI's substrate. It is not known whether the transport of uracil and thymine by PA0443 is pH-dependent, therefore this could be tested by using 2,4-dinitrophenol on membrane vesicle expressing PA0443.

PA0443 was overexpressed to 24% in the total *E. coli* membranes, migrating at 41 kDa on the SDS-PAGE gel (Figure 3.6 and Table 3.3). Tests for the expression of PA0443 in LB, 2TY and minimal medium suggested M9 minimal medium gave the highest expression of PA0443 under the test conditions (Figure 3.38), but time of harvesting should also be tested. The IPTG concentration test showed the highest level of expression of PA0443 when cells were induced with 0.5 mM IPTG (Figure 3.39).

The inner membranes prepared showed a higher level of PA0443 than in the total *E. coli* membrane, 30% compared to 24% (Figure 3.44), indicating this step is important for concentrating the protein. Some PA0443 proteins in the cell debris (Figure 3.44) could suggest this protein is forming inclusion bodies. Inclusion bodies are very dense and they tend to sediment readily under the influence of a low centrifugal force, and therefore they often cosedimented with cell debris (Walsh, 2001). Both inner and outer membrane fractions showed presence of the PA0443 protein (Figure 3.44). Some studies have shown aggregated proteins, being more dense, tend to cosediment with the outer membrane proteins (Laskowska *et al.*, 2004); therefore this could indicate the PA0443 protein is forming aggregations. The PA0443 protein was successfully purified from the inner membranes (Figure 3.45), and circular dichroism confirmed the integrity of the secondary structure of PA0443, showing predominantly α -helical content of the protein (Figure 3.46). Next, size-exclusion chromatography could be used to test the monodispersity and the suitability of PA0443 for crystallisation.

Here, the PA0443 gene, encoding a hydrophobic membrane protein, was confirmed to transport uracil and thymine (Figure 3.41). This is the first uracil/thymine transporter in *Pseudomonas* to be studied and characterised. Pyrimidine salvage enzymes are present in *Pseudomonas*, which enables the bacterium to utilise these compounds as sole source of nitrogen (Kim & West, 1999). Also, the PA0443 protein is important for the transport of pyrimidine into the cell when the preferred nitrogen source is depleted. The PA4647 gene of *P. aeruginosa*, belonging to the NCS2 family, and an orthologue of the UraA of *E. coli* (Andersen *et al.*, 1992), is also a putative uracil transporter but there are no experimental data to prove this at present.

Prokaryotic cells use RNA-based intracellular signals, cAMP, cGMP and c-diGMP for diverse biological functions. For example, c-diGMP regulates virulence factors and biofilm formation in *P. aeruginosa* (Cotter & Stibitz, 2007), and cAMP regulates various *in vivo* bacterial functions by binding to intracellular receptors and protein kinases (Baker & Kelly, 2004). A recent study showed that uracil could influence the quorum-sensing pathways of *P. aeruginosa* (Ueda *et al.*, 2009); biofilms of this organism are notoriously difficult to control. Whole-transcriptome analysis showed *Pseudomonas* with an altered uridine monophosphate (UMP) synthesis pathway had hundreds of quorum-sensing genes repressed; the transcription of these genes was restored by exogenous uracil, and an anticancer uracil analogue, 5-fluorouracil was able to repress biofilm formation (Uede *et al.*, 2009), so identifying UMP as an important

signal for biofilm formation in *Pseudomonas*. UMP is synthesised by several pathways in *P. aeruginosa*. One way includes the conversion of L-glutamine to UMP by the carbamoylphosphate synthetase (*carAB*) and *pyr* genes, or the conversion of uracil to UMP by uracil phosphoribosyltransferase (Kanehisa & Goto, 2000). The extracellular uracil transported into *Pseudomonas* is important for UMP synthesis, and hence biofilm formation (Uede *et al*, 2009); therefore, the PA0443 protein may be a potential target for inhibiting biofilm formation in *Pseudomonas*. Further studies, including creating mutants to understand the residues important for substrate binding, and solving the structure of the PA0443 uracil transporter of *Pseudomonas* will be crucial for understanding the fundamentals of uracil transport by this protein.

Chapter 4

Purification and substrate binding
studies of *Microbacterium liquefaciens*
hydantoin permease-1 (Mhp1)

4.1. Introduction – the hydantoin permease of *M. liquefaciens*

As previously described in the Introduction (Section 1.3.5), Mhp1 is composed of twelve transmembrane helices, ten of which are arranged in two inverted repeats of five helices (Figure 1.6). Based on this structure, a topology diagram of Mhp1 is made, highlighting the conserved residues amongst the NCS1 family of transporters (Figure 4.1). The phylogenetic tree (Figure 3.2) showed Mhp1 clustered most closely to HyuP of *Arthrobacter aurescens* (Table 1.1), a hydantoin transporter, and to Pden4351 and Pden0678, which were the NCS1 targets in this project, indicating that these two targets could also transport hydantoin.

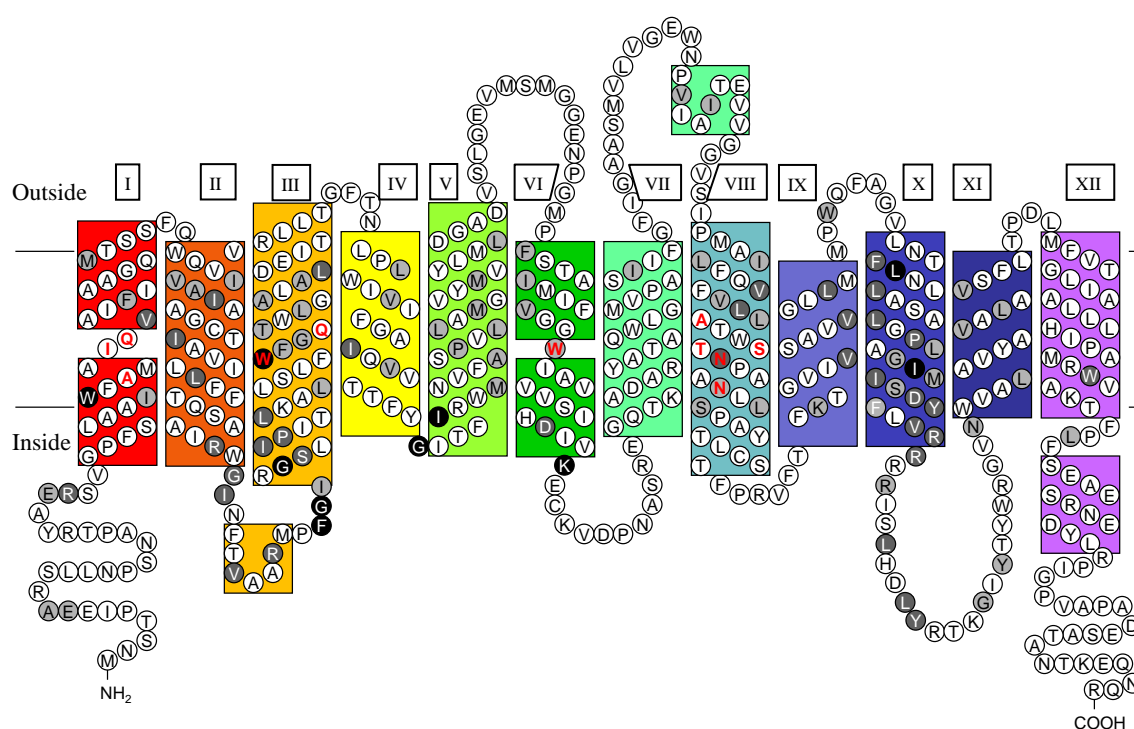


Figure 4.1 Topology diagram of *Microbacterium liquefaciens* hydantoin permease (Mhp1). Transmembrane segments are coloured and drawn based on the published structure of Mhp1 in Weyand *et al.* (2008). The helices are numbered from the N- to C-terminus and are shown in roman numerals. Conserved residues of the NCS1 family proteins are highlighted in grey and black according to the alignment in Appendix 2. Residues involved in substrate or cation binding are highlighted in red.

In the diffracting crystal of Mhp1, the electron density map at 2.85 Å showed a possible cation-binding site between helix I and helix VIII (Weyand *et al.*, 2008). The site is comprised of carbonyl-oxygen-atoms of amino acids A38, I41 and A309, and the hydroxyl-oxygen-atoms of the side chains of S312 and T313, which are predicted to form a cation-binding pocket. In other similar structures with also two inverted repeats of five helices, LeuT and vSGLT (Yamashita *et al.*, 2005; Faham *et al.*, 2008), an

equivalent sodium site was reported. The cation-binding site in close proximity to the substrate binding site presents a basis for the coupling of cation to substrate translocation. The work in this chapter was performed prior to publication of the structure of Mhp1 (Weyand *et al.*, 2008), and uses fluorimetry and the radiolabelled transport assay, to elucidate the cation that coupled to L-benzylhydantoin uptake.

4.2. Assessing the effect of sodium ions on the uptake of 5-L-benzylhydantoin by Mhp1

In a previous whole cell uptake experiment of Mhp1 for sodium dependency, a single concentration of 10 mM NaCl was tested and showed no evident effect on uptake (Suzuki & Henderson, 2006). Therefore it was repeated in this work using a wider range of NaCl concentrations (0 – 150 mM). The salt contents of the buffer were adjusted accordingly with KCl and kept constant at 150 mM for all experiments, and performed as described (Section 2.5.4). The greatest reduction in the uptake of L-benzylhydantoin (L-BH) was observed in the presence of 10 mM NaCl (Figure 4.2).

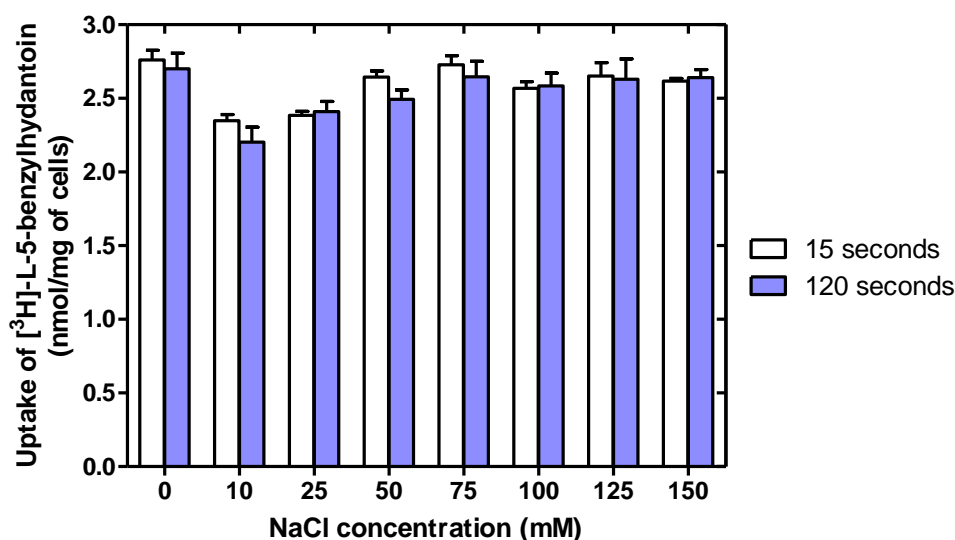


Figure 4.2 Effect of sodium ions on L-benzylhydantoin uptake by cells expressing Mhp1. *E. coli* BLR cells transformed with pTTQ18(*mhp1*)-His₆ were cultured in minimal medium in the presence of 100 µg/ml carbenicillin and 20 mM glycerol. When cells reached an A_{680nm} of 0.4 – 0.6, 0.5 mM IPTG was added, and cells were harvested 1 hour post-induction. Harvested cells were washed and energised with 20 mM glycerol, and incubated in 50 µM [³H]-L-5-benzylhydantoin, 5 mM MES pH 6.6 and 0 – 150 mM NaCl for uptake studies. Samples were taken 15 seconds and 2 minutes post-addition of radiolabelled L-benzylhydantoin. Results represent means of triplicates ±SEM.

Further increases in NaCl concentrations showed subtle increase in uptake (Figure 4.2), and suggest sodium could exert an effect on Mhp1 substrate uptake. Therefore this was further investigated using fluorimetry (Section 4.4).

4.3. Purification of Mhp1 from the inner membrane

Firstly for fluorimetry studies, Mhp1 was purified. The *E. coli* BLR strain transformed with pTTQ18(*mhp1*)-His₆ for this work is taken from Suzuki & Henderson (2006). Cells were cultured by Mr David Sharples, and the inner membranes he subsequently prepared were used in this work. Mhp1 was solubilised in DDM and purified from the inner membrane, and migrates at ~37 kDa on the SDS-PAGE (Figure 4.3), which is the same as previously described (Suzuki & Henderson, 2006). The purified protein at ~98% purity, determined by scanning densitometry, was confirmed intact by mass spectrometry showing the correct mass (Figure 4.4). Samples were also prepared for N-terminal sequencing (Section 2.6.4) and analysis by Dr Jeff Keen indicated the correct amino acids sequence for Mhp1 (MNSTPIEEAR). The purified protein was subjected to spectrophotofluorimetry experiments (Section 4.4), to assess the binding of ligands.

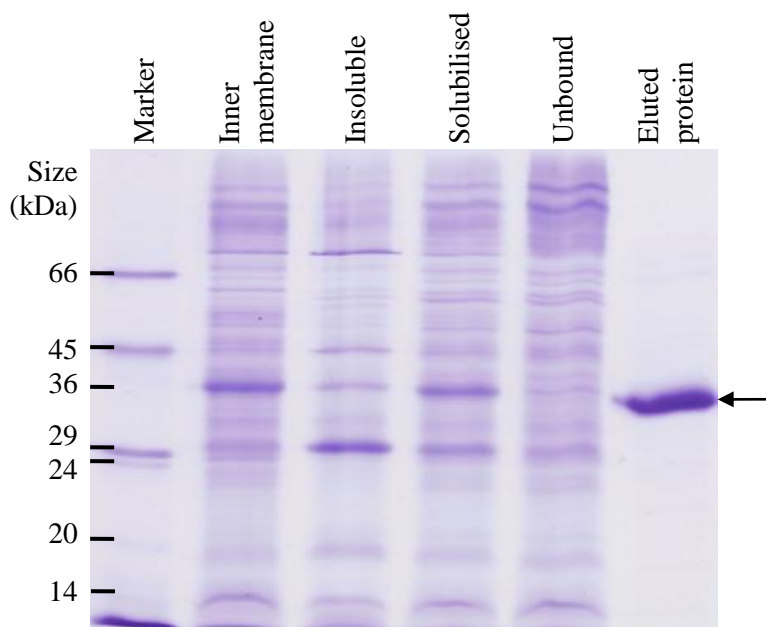


Figure 4.3 SDS-PAGE of purification of Mhp1. Mhp1 was purified from the inner membrane of *E. coli* BLR pTTQ18(*mhp1*)-His₆ cells (Section 2.5.3.1). The inner membranes were solubilised in 1% DDM. Fractions (15 µg), and eluted protein (10 µg) from the purification were analysed by SDS-PAGE. The arrow indicates the position of Mhp1.

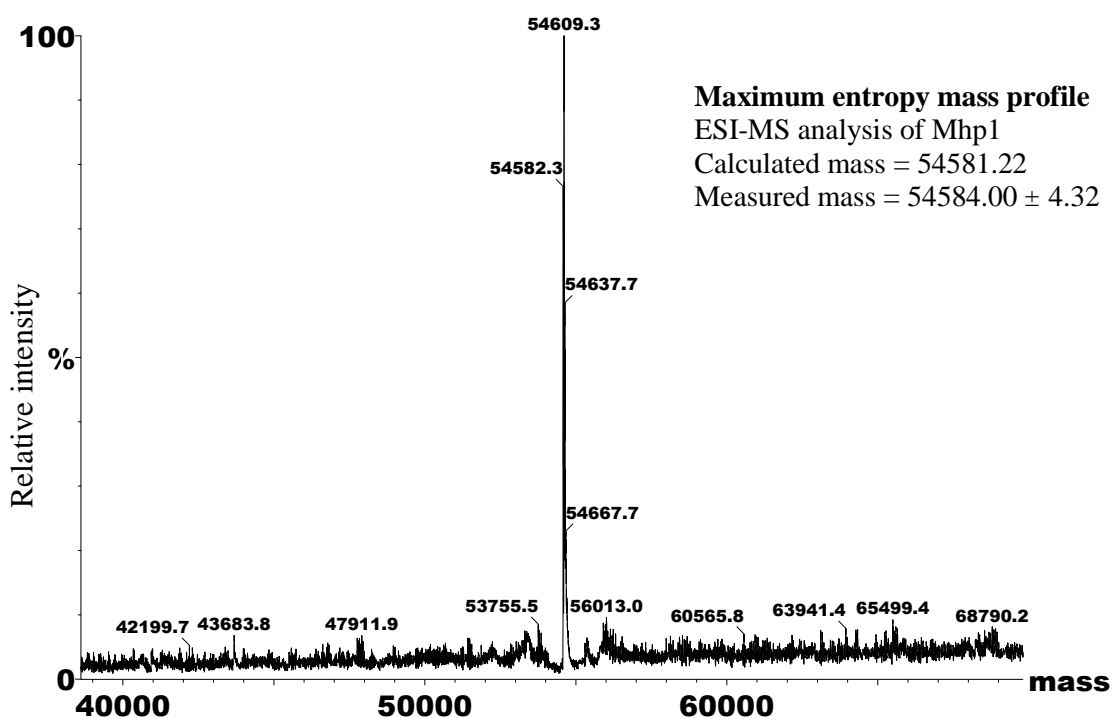
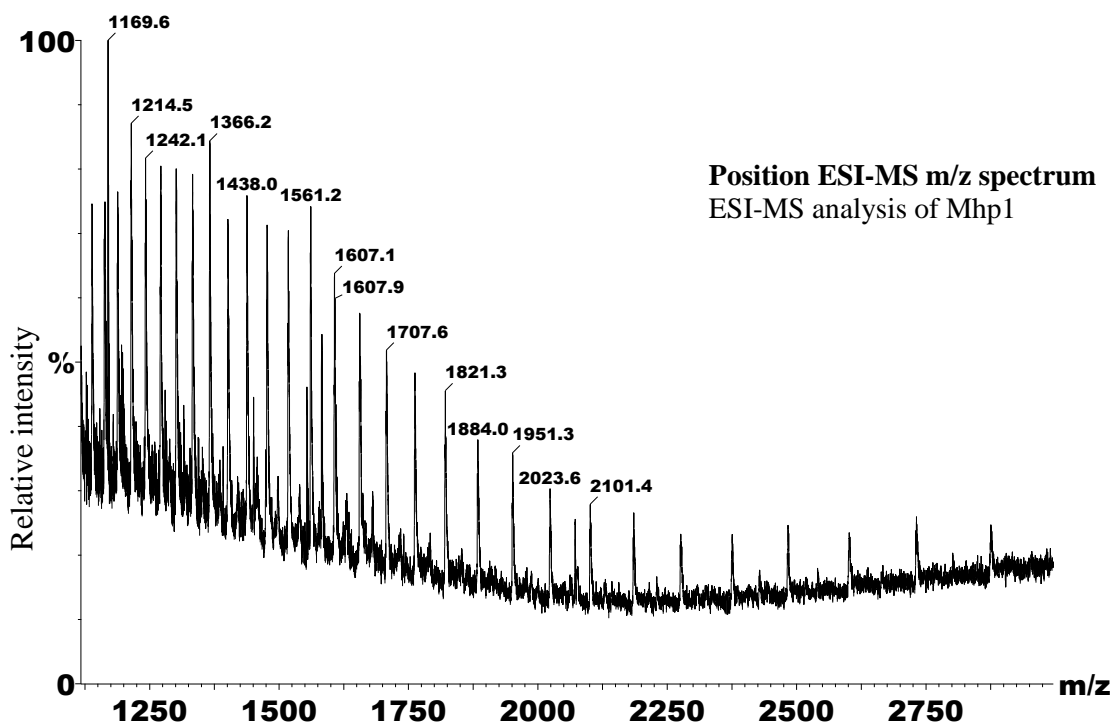


Figure 4.4 Electropray ionisation mass spectrometry of Mhp1. Purified protein (5 mg/ml) in 10 mM Tris pH 8.0 and 5% glycerol was prepared (Section 2.6.2). SEC, mass spectrometry and analysis were performed by Miss Lynsey Jones. There is evidence of formylation of the protein during its preparation for mass spectrometry.

4.4 Fluorimetry studies of Mhp1

Mhp1 has fifteen tryptophan residues, two of which are involved in substrate binding (W117 and W220) (Weyand *et al.*, 2008). Spectrophotofluorimetry experiments were performed to measure the intrinsic tryptophan fluorescence activity of Mhp1 upon titration with benzylhydantoin, NaCl or HgCl₂ (Section 2.6.3). Purified Mhp (140µg) protein in 10 mM potassium phosphate or Tris-HCl pH 7.6 buffer containing 0.05% DDM at 18°C was mixed for 4 minutes and titrated with one of the above compounds. Two minutes after each volume addition, a spectrum was scanned through 300 – 400 nm with an excitation wavelength of 285 nm, and the decrease in fluorescence at 348 nm was expressed as a percentage of the peak height.

Titration of 5-indolymethylhydantoin and 5-benzylhydantoin, substrates of Mhp1, were both initially tested. 5-indolymethylhydantoin was very insoluble and precipitated at the tested concentrations, whereas 5-benzylhydantoin remained in solution and was therefore used in the following fluorimetry experiments.

Firstly, the effect of sodium ions on the affinity of benzylhydantoin to the protein was investigated using fluorescence quenching (Section 2.6.3.1). There was an over 10-fold increase in affinity in the presence of sodium, with an apparent dissociation constant (K_d) 0.054 ± 0.007 mM compared to 0.88 ± 0.27 mM (Figure 4.5).

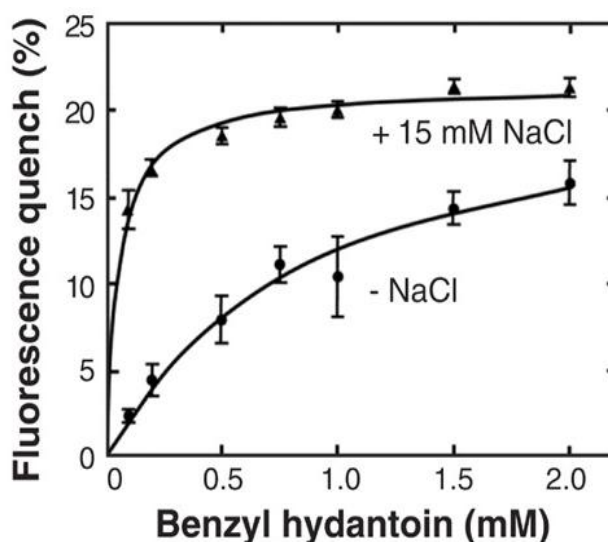


Figure 4.5 Quenching of tryptophan fluorescence in Mhp1 with benzylhydantoin. Purified Mhp1 was buffer-exchanged into 10 mM potassium phosphate pH 7.6 with 0.05% DDM, and titrated with benzylhydantoin (Section 2.6.3.1). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent means of triplicates \pm SEM. This work was performed with help from Dr Massoud Saidijam and is published in Weyand *et al.* (2008).

Also, an increase in affinity for sodium was observed in the presence of 2 mM benzylhydantoin (Figure 4.6); in the absence and presence of benzylhydantoin, the apparent K_d for sodium was 1.15 ± 0.28 mM and 0.15 ± 0.04 mM respectively, which is an approximately 8-fold increase. These results indicate that the binding of benzylhydantoin and sodium are tightly coupled in Mhp1.

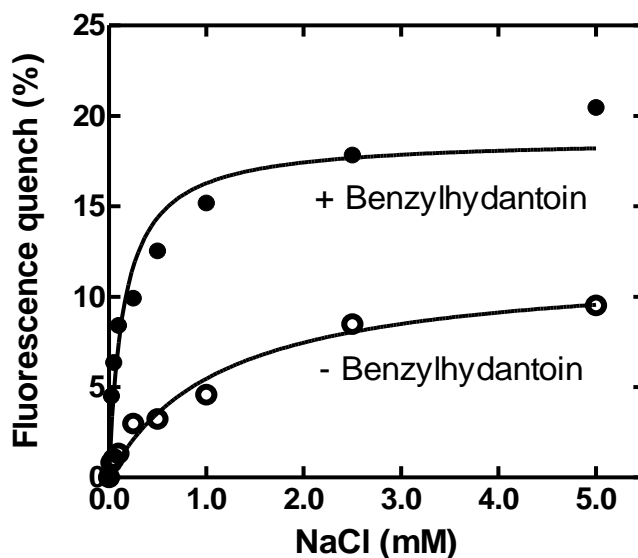


Figure 4.6 Quenching of tryptophan fluorescence in Mhp1 with NaCl. Purified Mhp1 was buffer-exchanged into 10 mM potassium phosphate pH 7.6 with 0.05% DDM, and titrated with 0 – 5 mM NaCl (Section 2.6.3.2). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent singlet data. This work was performed together with Dr Massoud Saidijam.

To examine at the cation specificity of Mhp1, a range of different cations were tested, this including lithium, sodium, potassium, rubidium and caesium (Figure 4.7). No indication was found that chloride ions interacted with the protein (Figure 4.7). Choline chloride was used to balance the osmolarity, so total salt concentration is 150 mM in the reaction volume, which is the same as in the transport assay (Section 2.5.4), i.e. where 15 mM chloride salt was added, 125 mM choline chloride and 10 mM Tris-HCl were also included. Similar to earlier observations, the presence of sodium caused an increased affinity of Mhp1 for benzylhydantoin, apparent K_d for benzylhydantoin was 0.0825 ± 0.0121 mM (Figure 4.7). The slight difference in the K_d could be due to Tris-HCl buffer used instead of potassium phosphate. The presence of other cations allowed a reduced fluorescence quench, and hence the affinity for benzylhydantoin was decreased (Figure 4.7), suggesting these cations are not stimulatory.

The effect of anions was also tested (Figure 4.8). Potassium acetate (CH_3COOK) caused an increase in fluorescence quench, possibly due to pH of the buffer (apparent

K_d of 0.1451 ± 0.01677). Other anions (CO_3^{2-} , SO_4^{2-} and PO_4^{2-}) had relatively very little effects, as the fluorescence quench is similar to the control.

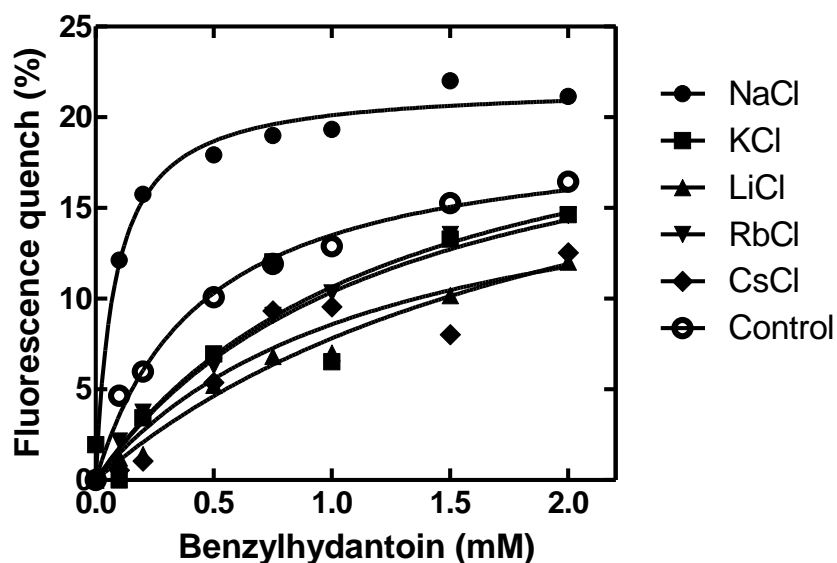


Figure 4.7 Quenching of tryptophan fluorescence in Mhp1 by different cations. Purified Mhp1 was buffer-exchanged into 10 mM Tris-HCl pH 7.6 with 125 mM choline chloride and 0.05% DDM, and titrated with benzylhydantoin (0 – 2 mM) in presence of different metal cations (15 mM) (Section 2.6.3.1). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent a single experiment. Control contains an additional 15 mM choline chloride.

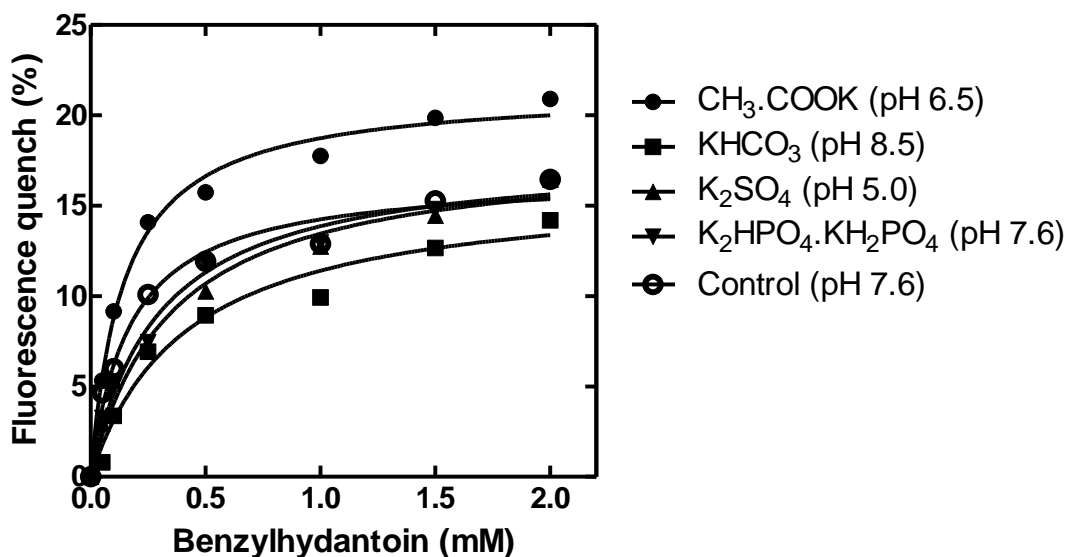


Figure 4.8 Quenching of tryptophan fluorescence in Mhp1 by different anions. Purified Mhp1 was buffer-exchanged into 10 mM Tris-HCl pH 7.6 with 0.05% DDM, and titrated with benzylhydantoin (0 – 2 mM) in presence of different anions (130 mM CH_3COOK ; 130 mM KHCO_3 ; 90 mM K_2SO_4 ; or 90 mM K_2HPO_4 KH_2PO_4) (Section 2.6.3.1). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent a single experiment. Control = 140 mM choline chloride.

Mhp1 crystals were phased by soaking in 2.5 mM methyl mercury acetate, 10 mM $K_2Pt(NO_2)_4$ or 0.5 mM ethyl mercury thiosalicylate (supplementary material for Weyand *et al*, 2008). Mercury (Hg^{2+}) has a high affinity for functional groups like $-SH$, $-OH$ and $-COOH$. A negative fluorescence quenching is shown in the presence of 2 mM $HgCl_2$ (Figure 4.9), suggesting mercury could possibly affect substrate binding of Mhp1.

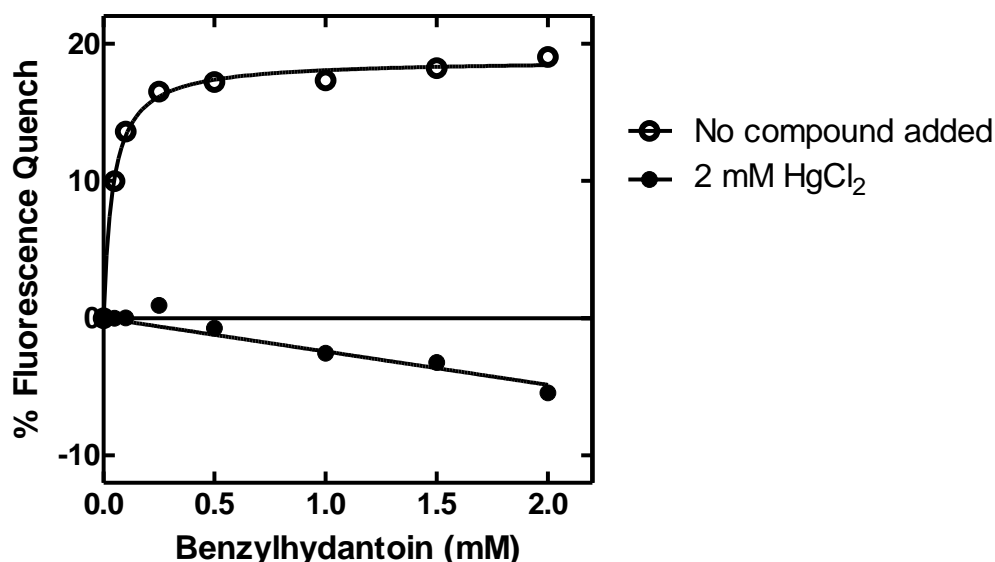


Figure 4.9 Quenching of tryptophan fluorescence in Mhp1 by benzylhydantoin in the presence and absence of $HgCl_2$. Mhp1 (140 μ g) was added to a reaction volume (980 μ l) containing 10 mM Tris-HCl pH 7.6, 135 mM Choline chloride, 0.05% DDM and 5 mM NaCl, and titrated with benzylhydantoin (0 – 2 mM) (Section 2.6.3.1). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent duplicates.

Further experiments involving titrations with $HgCl_2$ in the presence and absence of NaCl and/or benzylhydantoin were performed (Figure 4.10). Results indicate in the presence of these compounds a reduction in the fluorescence quench was observed (Figure 4.10). Sodium caused the most dramatic reduction. Benzylhydantoin caused a smaller reduction than sodium, with the presence of these two compounds showing an intermediate effect (Figure 4.10). When mercury was present prior to titration with benzylhydantoin (Figure 4.9), the fluorescence quench became negative and there was an increase in fluorescence intensity. Looking at it the other way, when benzylhydantoin was present and the solution was titrated with mercury (Figure 4.10), the fluorescence quench stay positive. This could suggest benzylhydantoin protects the protein against $HgCl_2$, as does sodium.

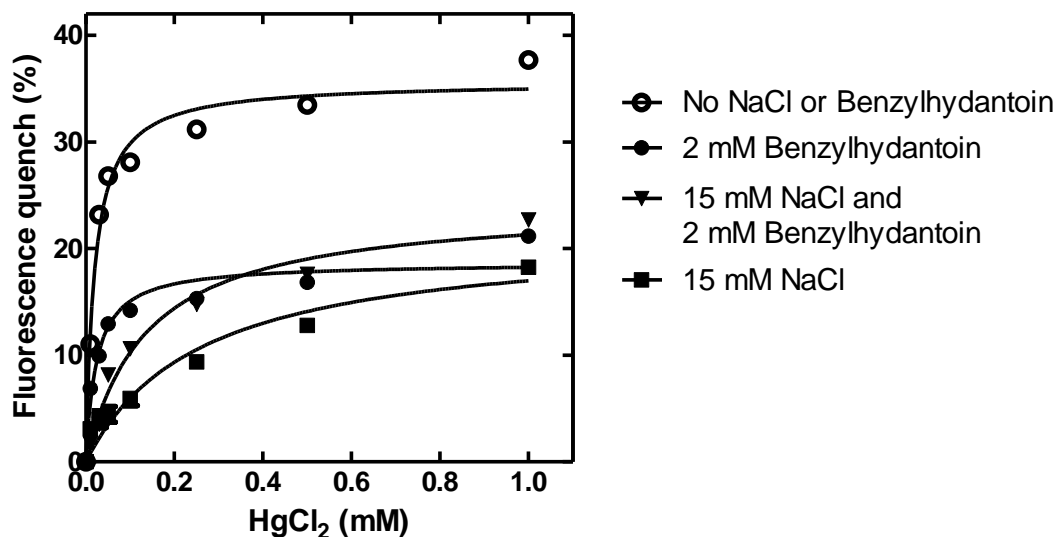


Figure 4.10 Quenching of tryptophan fluorescence in Mhp1 by HgCl₂. Purified Mhp1 was buffer-exchanged into 10 mM Tris-HCl pH 7.6 with 0.05% DDM, and titrated with HgCl₂ (0 – 1.0 mM) (Section 2.6.3.3). The decrease in tryptophan fluorescence at 348 nm was monitored. Results represent duplicates.

4.5 Discussion: the substrate binding studies of Mhp1

Mhp1, as well as other NCS1 family transporters from yeast, fungi and bacteria, were previously thought to be proton-dependent (Suzuki & Henderson, 2006; Pantazopoulou & Diallinas, 2007; Ren *et al.*, 2007). Whole cell transport experiments showed an initial decrease in transport of benzylhydantoin in the presence of 0 – 10 mM sodium, which was followed by an increase in the presence of 10 – 150 mM sodium (Figure 4.2). Sodium dependence in whole cell transport assays may be obscured by the presence of other sodium transport systems in the membrane. Purified Mhp1 (Figure 4.3), which was confirmed to be intact (Figure 4.4), was subjected to reconstitution into liposomes, followed by transport assay, but attempts to measure transport were unsuccessful (data not shown) due to the low solubility and the lipophilic nature of the substrate.

This work showed that binding of benzylhydantoin to Mhp1 is sodium-dependent, with an increased affinity shown for benzylhydantoin when 15 mM NaCl was present (Figure 4.5), and the same was shown for the binding of NaCl in the presence of benzylhydantoin (Figure 4.6). Other cations that were tested showed little stimulatory effect, demonstrated by the modest reduction in fluorescence quench (Figure 4.7). Mhp1 is shown to have a high affinity for sodium over the other cations. The anions tested showed that potassium acetate gave the highest fluorescence quench

(Figure 4.8). The solutions were of variable pHs, and therefore it was difficult to tell if these anions had an effect on Mhp1. The highest fluorescence quench was observed at pH 6.5 (Figure 4.8). Mhp1 was previously thought to be proton dependent; the optimal transport observed at pH 6.6 (Suzuki & Henderson, 2006) - the observed pH in this work is similar. It appears Mhp1, like the MelB sugar-cation symporter (Wilson & Ding, 2001), could have a flexible cation selectivity, being proton- as well as sodium-dependent.

Mercury was used in phasing and therefore its ability to quench fluorescence was tested. A substantial quench by mercury occurred with possible protection of Mhp1 by sodium and benzylhydantoin from mercury (Figure 4.9 and 4.10). It is likely that Hg^{2+} targets the cysteine at position 69, therefore further experiments such as mutation of this residue and repeating the fluorimetry are needed.

Chapter 5

Overexpression and purification of the intact membrane histidine kinases of *Enterococcus faecalis* V583 for activity, structural and ligand binding studies

Chapter 5 – Introduction

This chapter is divided into four sections:

Section 5.1 – Evaluation of the expression, purification and activities of all membrane-bound histidine kinases from *Enterococcus faecalis* V583

Section 5.2 – Trials to identify the signalling ligands for histidine kinase EF3197

Section 5.3 – Structural studies: crystallisation trials of histidine kinase EF1051

Section 5.4 – Characterisation of signal interactions by the quorum sensor EF1820 (FsrC)

This chapter investigates the 16 membrane histidine kinases of the two-component systems of *E. faecalis*, which is a collaborative project between two Principal Investigators, Professor Peter Henderson and Dr Mary Phillips-Jones. Section 5.1 describes the expression, purification and activities of these membrane proteins; the contributions by members of the research team are indicated in Section 5.1.1 (work published in Ma *et al.*, 2008). Sections 5.2 to 5.4 are solely my own work. Section 5.2 describes the signal screening of one of the proteins, a homologue of LytS, EF3197. The work involved testing putative compounds for their possible effects on the autophosphorylation activity of EF3197. Section 5.3 describes the optimisation of expression and crystallisation trials of EF1051. Section 5.4 investigates in detail the quorum sensing kinase EF1820 (FsrC), to characterise the effects of ligand binding and an inhibitor on kinase activities.

5.1. Evaluation of the expression, purification and activities of all membrane-bound histidine kinases from *Enterococcus faecalis* V583

5.1.1. Introduction –the histidine kinases of *Enterococcus faecalis*

The completed genome sequencing of the *E. faecalis* strain V583 led to the identification of 17 two-component systems (TCS) pairs and one orphan response regulator (Hancock & Perego, 2002). All but one histidine kinase (HK) of *E. faecalis* was predicted to be membrane-bound. This study investigates the entire family of intact membrane-bound HKs of *E. faecalis*, evaluating the systematic cloning, expression, purification and activity assays of each of the 16 membrane-bound HKs of *E. faecalis*. In this project, the work for six HKs (EF3197, EF2912, EF1051, EF0373, EF1209 and EF1820) were carried out by myself, nine HKs (EF2219, EF1704, EF3290, EF1261, EF1863, EF0927, EF2298, EF0570 and EF1335) were studied by Dr Hayley Yuille, and work with EF1194 was carried out by Dr Victor Blessie, Miss Nadine Göhring and Miss Zsófia Iglói.

5.1.2. Cloning and expression of histidine kinases of *E. faecalis*

Each membrane-bound HK gene was amplified by PCR (Section 2.3.2) using the specific primer pairs shown in Table 2.4 with *E. faecalis* V583 chromosomal DNA as template. The amplified products were digested with the relevant restriction enzymes and cloned into plasmid pTTQ18-His₆ (Section 2.3.1). The resulting plasmids of the recombinant clones were transformed into *E. coli* BL21(DE3) for expression, purification and activity trials. The clones were confirmed by DNA sequencing (Appendix 7), and the resulting amino acid sequences are shown in Appendix 8.

Since the expression of membrane proteins is significantly lower than that of soluble proteins, amplified expression is therefore required for the production of milligram quantities of proteins. Conditions previously reported for successful production of other bacterial membrane proteins (Ward *et al.*, 2000; Potter *et al.*, 2002; Saidijam *et al.*, 2003; Suzuki & Henderson, 2006) were therefore tested initially. Typically use of LB medium, induction of expression using 0.5 – 1 mM IPTG, and a post-induction temperature of 30°C have proved successful for the overexpression of most of the HKs. For proteins exhibiting low expression, optimisation of culture and expression conditions were subsequently performed, such as testing growth in different media (LB, 2TY and M9 minimal medium), different IPTG concentrations for induction, and different temperature and time of harvesting post-induction with IPTG. The expression of each HK was assessed by analysis of the total *E. coli* membranes by SDS-PAGE and Western blotting to detect each His₆-tagged protein. The results are summarised in Table 5.1.

Most HKs were optimally expressed in LB medium, with only one protein (HK06) expressed at higher levels in 2TY medium and none of the low-expressing proteins exhibited better expression in minimal medium compared with the other media (Table 5.1). The optimisation tests suggested that the most suitable conditions for post-induction temperature varied between 15, 30 and 37°C; suitable IPTG concentrations ranged from 0.2 to 1 mM; and the time of harvesting post-induction was adequate at 3 hours, with only HK08 and HK16 benefitting from overnight incubation at a lower temperature, 15°C (Table 5.1).

The observed molecular masses of all recombinant proteins in the SDS-PAGE and Western blot analysis of the total membranes appeared less than the predicted mass (Table 5.1 and Figure 5.1), which is typical for the migration of membrane proteins on SDS-PAGE due to their hydrophobicity (Ward *et al.*, 2000).

Kinase name	Work carried out by	Predicted number TMs ¹	Expression medium	Post-induction temperature (°C)	[IPTG] (mM)	Harvest time post-induction (hours)	Predicted mass with His ₆ -tag (kDa) ²	Observed mass (kDa) in mixed membranes ³	Proportion of mixed membrane proteins (%) ⁴	Solubilisation detergent	Observed mass (kDa) of intact purified proteins ³
EF2219 (HK01)	HY	2	ND	ND	ND	ND	67.9	ND	ND	NP	NP
EF3197 (HK02)	PM	6	LB	37	0.5	3	66.8	53	12	DDM	55
EF2912 (HK03)	PM	2	LB	37	0.5	3	44.0	38	15	DDM	38
EF1704 (HK04)	HY	2	LB	30	1.0	3	69.3	52	11	DDM	54
EF3290 (HK05)	HY	2	LB	30	1.0	3	46.4	37	14	DDM	38
EF1261 (HK06)	HY	2	2TY	30	0.2	3	57.8	52	8	DDM	54
EF1194 (HK07)	VB, NG and ZI	2	LB	30	0.4	3	71.2	60	11	DDM	60
EF1863 (HK08)	HY	2	LB	15	1.0	16	52.6	42	12	LDAO	45
EF0927 (HK09)	HY	2	LB	30	1.0	3	41.7	31	6.5	DDM	32
EF1051 (HK10)	PM	2	LB	37	0.5	3	60.2	50	16	DDM	50
EF2298 (HK11)	HY	2	LB, 2TY	30	0.5, 1.0	3	52.0	42	10	DDM / UDM	48
EF0570 (HK12)	HY	4	LB, 2TY	30	1.0	3	98.7	38	10	ND	NP
EF0373 (HK13)	PM	2	LB	30	0.5	3	58.1	46	4.5	DDM	44
EF1209 (HK14)	PM	2/3	LB	30	0.5	3	61.1	48	16.5	DDM	48
EF1820 (HK15)	PM	7	LB	37	0.5	3	53.4	35	3.5	DDM	34
EF1335 (HK16)	HY	7	LB	15	1.0	16	52.2	32	ND	DDM	NP

Table 5.1 Expression conditions and solubilisation of the 16 intact histidine kinases of *E. faecalis*. ¹Predicted using TMHMM (Krogh *et al.*, 2001) and/or TopPred (Claros & von Heijne, 1994); ²mass of the His₆-tagged protein predicted using the ExPASy proteomic server (Gasteiger *et al.*, 2003; Gasteiger *et al.*, 2005); ³Western blotting analysis with INDIATM His probe; ⁴determined by scanning densitometry of SDS-PAGE gel shown in Figure 5.1. DDM, n-dodecyl-β-D-maltoside; UDM, n-undecyl-β-D-maltoside; LDAO, N,N-dimethyl dodecylamide oxide; ND, included in the trials but not detected or identifiable; NP, not possible to undertake. HY, Dr Hayley Yuille; PM, Pikyee Ma; VB, Dr Victor Blessie; NG, Miss Nadine Göhring; ZI, Miss Zsófia Iglói. These results have been published in Ma *et al.*, (2008).

In general, it appears that proteins possessing higher numbers of predicted transmembrane regions generally exhibited greater mass differences. For example, EF1820 (HK15) and EF1335 (HK16) (Table 5.1), both possess seven putative transmembrane regions, respectively, and their apparent masses on SDS-PAGE were 35 and 32 kDa, whilst their predicted masses are 53 and 52 kDa, respectively (Table 5.1 and Figure 5.1).

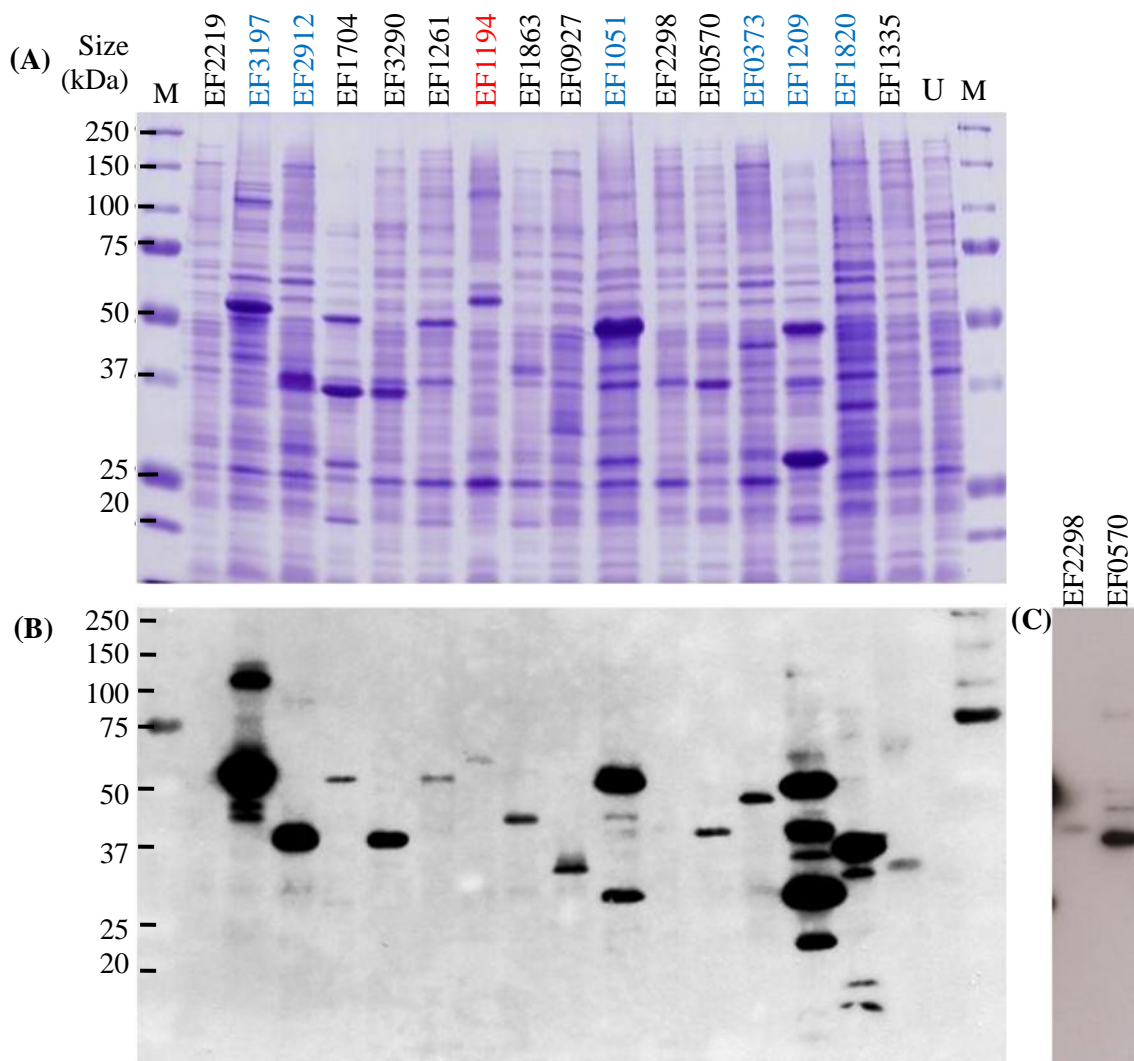


Figure 5.1 Expression of the 16 intact membrane-bound histidine kinases of *E. faecalis* in *E. coli* BL21(DE3). Cells transformed with pTTQ18 harbouring the recombinant gene for each of the intact histidine kinases were cultured as described in Section 2.4.1, and total membrane proteins prepared by water lysis. Samples (15 μ g) were separated and analysed by: (A) SDS-PAGE and visualized using Coomassie blue staining; and (B) and (C), Western blotting to detect the His₆-tagged proteins. Panel C shows a longer exposure required for detection of the successful expression of EF2298 (HK11). M, molecular mass standards; U, total membranes of an uninduced culture of *E. coli* BL21(DE3) pTTQ18(EF2219)-His₆. Total membranes prepared by: myself (blue); Dr Hayley Yuille (black); Miss Zsófia Iglói (red). SDS-PAGE and Western blotting performed by myself and Dr Hayley Yuille, respectively. These results have been published in Ma *et al.* (2008).

Fifteen of the sixteen intact membrane HKs of *E. faecalis* were successfully overexpressed in *E. coli* membranes, with only EF2219 (HK01) not detectable by Western blotting under the test conditions (Figure 5.1). Some degradation was observed for EF1209 (HK14), and use of protease inhibitors in the water lysis membrane preparations of this protein did not prevent such degradation. A large proportion of the EF3197 protein migrated at 53 kDa and some at ~120 kDa on the SDS-PAGE; the higher mass observed for this protein was not seen for the other HKs (Figure 5.1).

5.1.3. Autophosphorylation activities of *E. faecalis* histidine kinases within *E. coli* inner membranes

To test the activities of the cloned histidine kinases, inner membrane vesicles expressing each histidine kinase were prepared as described in Section 2.5.1. Each assay employed 200 µg of inner membrane proteins; the reactions were initiated by addition of radiolabelled ATP, and 20 µg samples were taken at time intervals. The levels of phosphorylated proteins were analysed as described in Section 2.5.5.1. Eleven histidine kinases showed detectable activity in inner membranes (Figure 5.2). No detectable activity was shown in at least three attempts for EF0927 (HK09), EF2298 (HK11), EF0570 (HK12) and EF1335 (HK16), suggesting that these proteins are either inactive or that there is an absence of an activating signal (Figure 5.2). The membrane-localised histidine kinases displayed a range of different kinetics (Figure 5.2). Nine of the kinases (EF3197, EF1740, EF1261, EF1194, EF1863, EF1051, EF0373, EF1209 and EF1820) reached maximum autophosphorylation within the initial 20 minutes, and for EF3197, EF1704, EF1194, EF1863, EF1051, EF0373 and EF1209, this was followed by rapid dephosphorylation (Figure 5.2). The phosphorylation of EF2912 increased steadily throughout the assay time, and did not reach a maximum in the 60 minutes time course (Figure 5.2). EF3290 was atypical, reaching maximum phosphorylation after approximately 30 minutes, followed by slow dephosphorylation, maintaining approximately 60% of its maximum activity by the end of the time period (Figure 5.2).

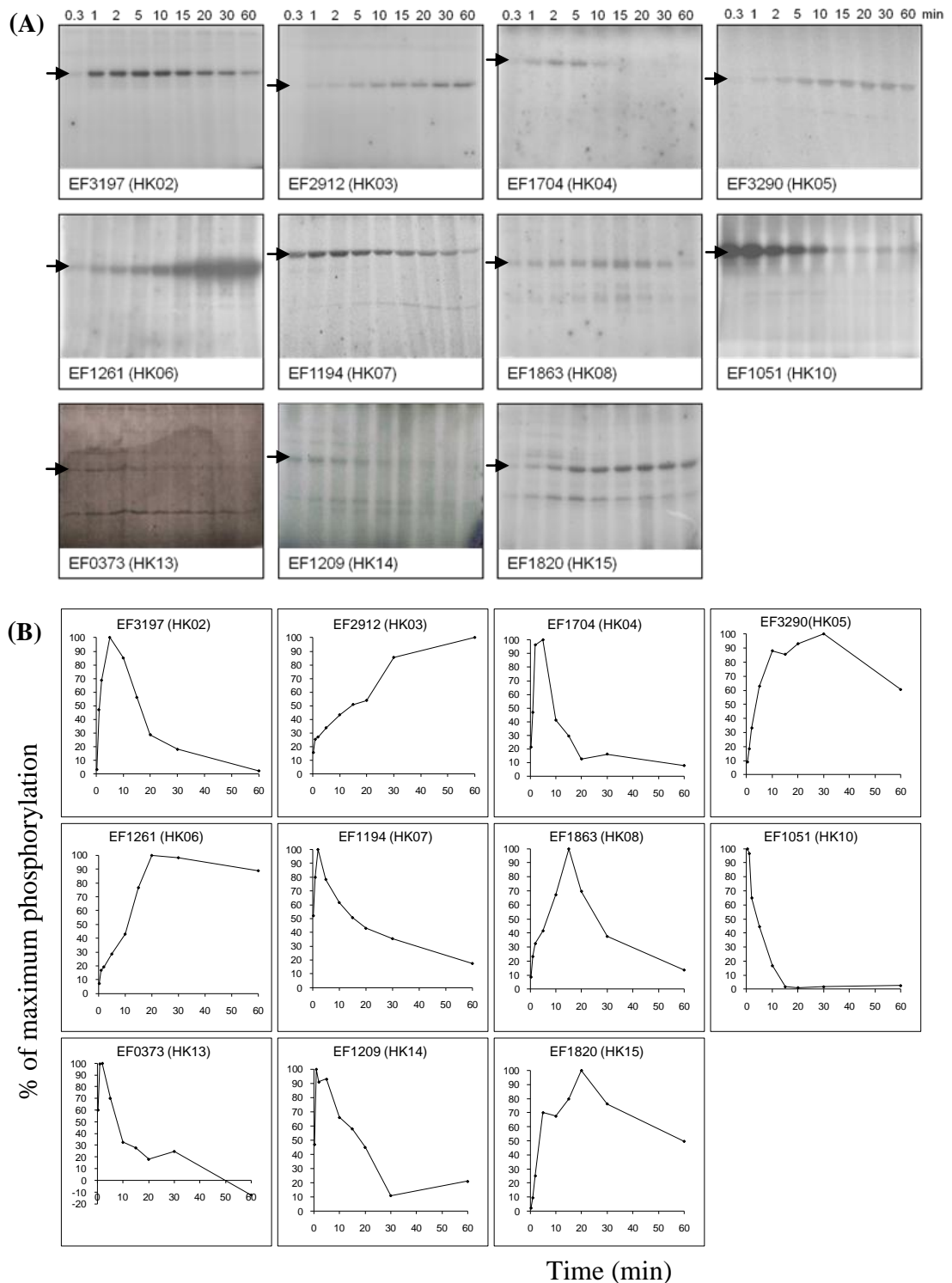


Figure 5.2 Autophosphorylation activities of intact histidine kinases in inner membrane vesicles. Inner membrane proteins (200 μg) were incubated at 24°C and reactions initiated by the addition of radiolabelled ATP in a total reaction volume of 200 μl , as described in Section 2.5.5.1. Samples (20 μl) were removed at the time points indicated. (A) Image and (B) quantitative representation of the data for each radiolabelled phosphorylated protein during phosphorylation. The arrows indicate position of the phosphorylating proteins. Graphs show the pixel counts relative to the maximum value attained during the assays. Assays performed by: myself – EF3197, EF2912, EF1051, EF0373, EF1209 and EF1820; Dr Victor Blessie – EF1194; Dr Hayley Yuille – EF1704, EF3290, EF1261 and EF1863).

5.1.4. Solubilisation and purification of *E. faecalis* histidine kinases

Solubilisation trials were performed as described in Section 2.5.2. Fourteen out of the fifteen histidine kinases were successfully solubilised from the *E. coli* membranes. Most of the histidine kinases were successfully solubilised by DDM (Table 5.1), a widely used detergent for solubilising bacterial membrane proteins (Ward *et al.*, 2000; Saidijam *et al.*, 2003), except for EF1863 (HK08) and EF0570 (HK12). LDAO was used for successful solubilisation of EF1863 (HK08), whilst none of the detergents tested was able to solubilise EF0570 (HK12) from the *E. coli* membranes (Table 5.1).

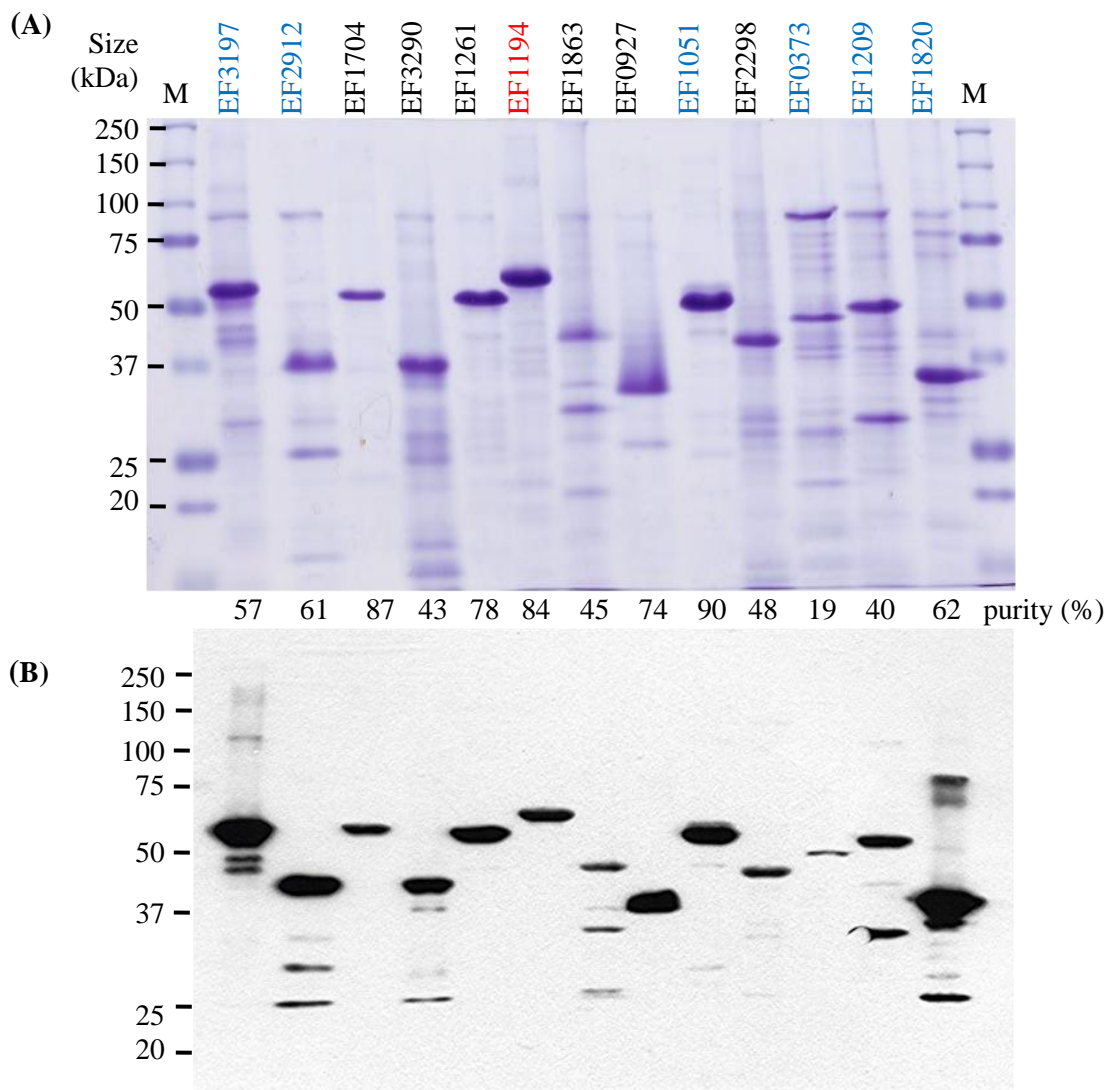


Figure 5.3 One-step purification of the expressed intact membrane histidine kinases of *E. faecalis*. Proteins were solubilised using detergent and purified from the inner membranes by Ni-NTA affinity chromatography, as described in Section 2.5.3.2. Eluted fractions from the purification (10 μ g) were analysed by: (A) separation on SDS-PAGE and staining with Coomassie Blue, and % purity determined by densitometry; and (B) Western blotting analysis using an INDIATM HisProbe (performed by myself). M, molecular mass marker. HKs purified by: myself (blue); Dr Victor Blessie (red); and Dr Hayley Yuille (black).

Following histidine kinase solubilisation from *E. coli* inner membranes, purification of the His₆-tagged proteins was performed by Ni-NTA purification, (Section 2.5.3.2). Thirteen of the fifteen expressed histidine kinases were successfully purified from the *E. coli* inner membranes. Only EF1335 (HK16) could not be eluted from the Ni-NTA column under the test conditions. The one-step purifications of the histidine kinases are shown in Figure 5.3. The purity of the eluted proteins ranged from 19-90 % (Figure 5.3). The purities of EF1704 (HK04), EF1261 (HK06) and EF1051 (HK10) were ranked highest, attaining 87, 78 and 90%, respectively, as determined by scanning densitometry. The lowest purity obtained was 19% for EF0373 (HK13) (Figure 5.3), which could be due to the low level of expression of this protein (Table 5.1). The positions of the purified His₆-tagged proteins on the SDS-PAGE was confirmed by Western blotting using the INDIATM HisProbe-HRP, and their observed masses agreed with their obtained masses in mixed membranes (Table 5.1, Figures 5.1 and 5.3).

Western blotting of the purified histidine kinases suggested possible degradation of some of the proteins, as proteins of lower masses were observed in gels for the eluted samples of EF3197 (HK02), EF2912 (HK03), EF3290 (HK05), EF1863 (HK08), EF1209 (HK14) and EF1820 (HK15), in addition to the undegraded intact form.

5.1.5. Integrity of purified histidine kinases determined by Western blotting, N-terminal sequencing and/or mass spectrometry

The 13 purified proteins were verified to be intact and undegraded by N-terminal sequencing and/or electrospray ionisation mass spectrometry (ESI-MS), together with Western blotting that shows the His₆-tagged C-terminus is retained. N-terminal sequence and/or MS data were obtained for all but three proteins: EF3197 (HK02) for which no N-terminal or MS data could be obtained; EF0570 (HK12), which was not solubilised from the membranes; and EF1335 (HK16), which was not successfully eluted from the Ni-NTA columns (Table 5.2). Both EF3197 (HK02) and EF1820 (HK15) appear to be N-terminally blocked by possibly a formyl group, and initial N-terminal sequencing gave very high backgrounds, suggesting random degradation of protein. These proteins were therefore treated with acid-methanol (Section 2.6.5), prior to N-terminal sequencing. The EF1820 protein was chemically unblocked and N-terminal sequence successfully obtained, whereas the treatment failed for the EF3197 protein.

Kinases name	N-terminal sequence of purified protein	Electrospray ionisation mass spectrometry of purified protein
EF2219 (HK01)	No expression	No expression
EF3197 (HK02)	Possible formylation of protein preventing Edman-degradation of protein	N/T
EF2912 (HK03)	MNSHMTDRIS	N/T
EF1704 (HK04)	MNSHMKKRL	N/T
EF3290 (HK05)	MNSHMLVKP	N/T
EF1261 (HK06)	MNSHMK	N/T
EF1194 (HK07)	MNSKKV	71235 Da
EF1863 (HK08)	MNSKCLKIFP	N/T
EF0927 (HK09)	MNSHMTILKY	N/T
EF1051 (HK10)	MNSHMKRT	60189 Da
EF2298 (HK11)	MNSHME	N/T
EF0570 (HK12)	Not solubilised from membrane	Not solubilised from membrane
EF0373 (HK13)	MNSHMTIKR	N/T
EF1209 (HK14)	MNSHMKRGGK	N/T
EF1820 (HK15)	MNSHMILSLL	N/T
EF1335 (HK16)	Not purifiable	Not purifiable

Table 5.2 N-terminal sequencing and ESI-MS of histidine kinases of *E. faecalis*. Purified proteins separated by SDS-PAGE were electroblotted onto PVDF-membrane and stained with Coomassie Blue. The stained protein was subjected to N-terminal sequence (Section 2.6.4) by Edman degradation (Edman, 1950), performed by Dr Jeff Keen. ESI-MS was performed as described in Section 2.6.2 and samples analysed by Miss Lynsey Jones. N/T, Not yet tested for ESI-MS. Samples for N-terminal sequencing and ESI-MS were prepared by: myself – EF3197, EF2912, EF1051, EF0373, EF1209 and EF1820; Dr Victor Blessie – EF1194; Dr Hayley Yuille – EF1704, EF3290, EF1261, EF1863, EF0927 and EF2298.

5.1.6. Autophosphorylation of purified *E. faecalis* histidine kinases

Twelve out of thirteen histidine kinases retained autophosphorylation activity after purification from *E. coli* inner membrane (Figure 5.4). Two of these proteins, EF2298 (HK11) and EF1209 (HK14) were also active, but the very weak activity observed is not shown. Of the eleven histidine kinases that were active in the *E. coli* inner membranes (Figure 5.2), all but one, EF1863 (HK08), remained active post-purification. EF1863 (HK08) was highly aggregated after extraction from the inner membranes and 0.1% LDAO was added to the assays to reduce this aggregation, so the

loss of detectable activity observed may be attributable to the LDAO. In this work, even low concentration of detergents, such as 0.05% DDM in the autophosphorylation assays were observed to be inhibitory for the activities of some purified histidine kinases, and therefore detergents were often not added post-solubilisation unless it was necessary to prevent the aggregation of proteins.

It appears that EF0927 (HK09), which previously produced no detectable activity in the inner membranes, was active following purification (Figure 5.2). EF2298 (HK11) VanS_B had a low activity level and showed no detectable increase in phosphorylation upon addition of a 10-fold molar excess of vancomycin, a putative signal for this kinase. This suggests that this antibiotic is not an activating signal for this kinase under the test conditions, or that the recombinant protein is not able to respond to its signal.

Only two histidine kinases, EF2912 (HK03) and EF1261 (HK06), showed similar kinetics in both membranes and in the purified form (Figure 5.2 and 5.4). The kinetics of most of the purified histidine kinases, except for EF1704 (HK04), EF3290 (HK05) and EF1261 (HK06), exhibited increased autophosphorylation during the sixty minutes time-course (Figure 5.4). For purified EF3197 (HK02), EF1704 (HK04), EF3290 (HK05), EF1194 (HK07), EF1051 (HK10), EF0373 (HK13) and EF1820 (HK15), autophosphorylation activities reached maximum levels later in the time-course, with little desphosphorylation compared to that observed when membrane-bound (Figure 5.2 and 5.4). Although the different protein concentrations present may affect kinetics, these data suggest that autophosphorylation rates are faster when the kinases are bound to the membrane than when extracted or in their purified form (Figure 5.2 and 5.4). As the kinases are expressed heterologously in *E. coli*, the inner membrane vesicles prepared and used for the assay may contain *E. coli* phosphatases that could also contribute to the decrease in phosphorylation of the kinases in the membrane-bound kinases assay; also there could be *E. coli* components promoting the rapid phosphorylation of the membrane-bound histidine kinases (Figure 5.2).

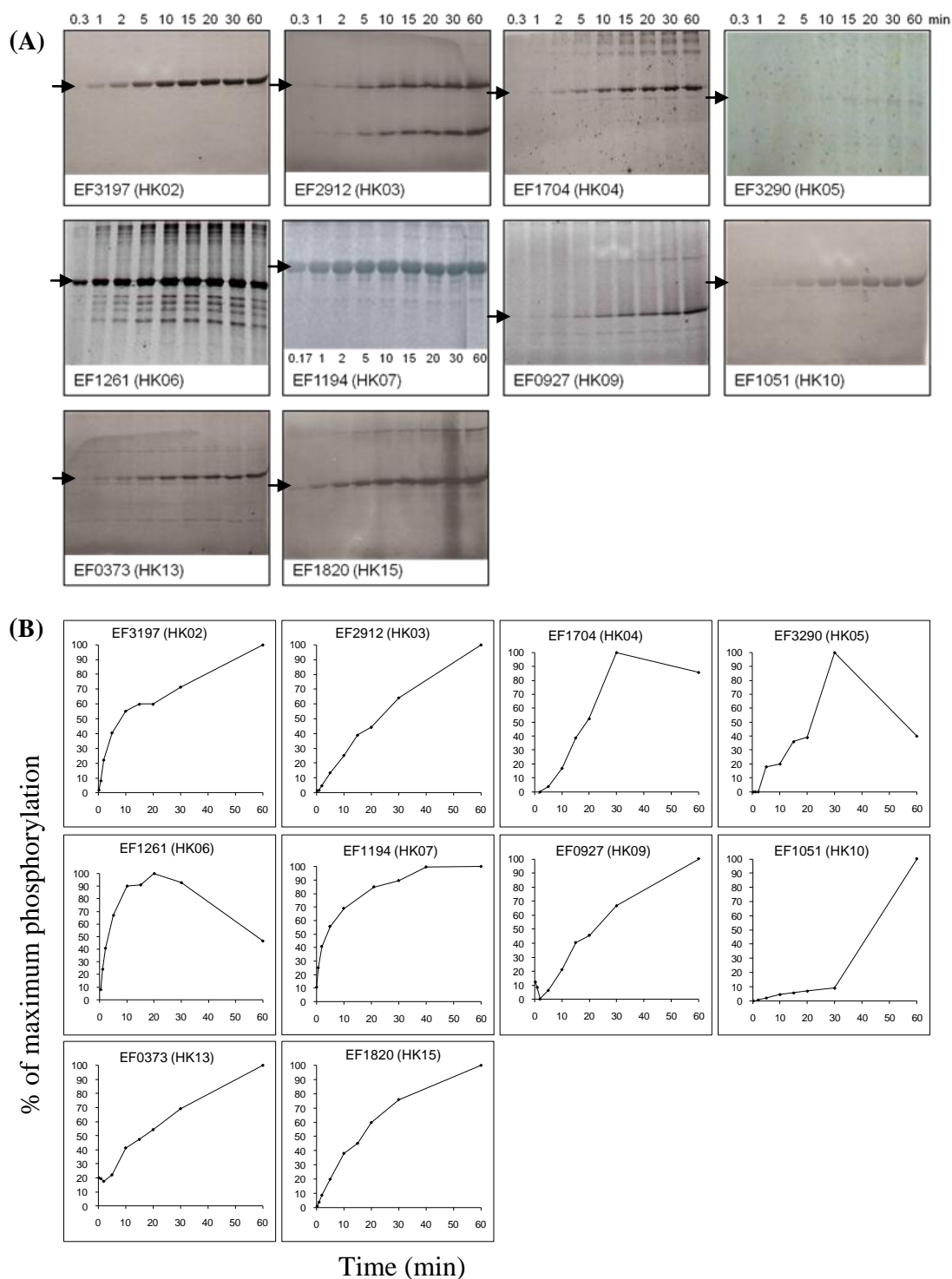


Figure 5.4 Autophosphorylation activities of purified intact histidine kinases. Purified proteins (600 – 1000 pmoles) were added to reaction buffer, and autophosphorylation initiated using radiolabelled ATP in a total reaction volume of 150 μ l (Section 2.5.5.2). Samples (15 μ l) were removed at the time points indicated. (A) Image and (B) quantitative representation of the data for each radiolabelled phosphorylated protein during autophosphorylation. The arrows indicate position of the phosphorylating proteins. Graphs show the pixel counts relative to the maximum value attained during the assays. Assays performed by: myself – EF3197, EF2912, EF1051, EF0373 and EF1820; Dr Mary Phillips-Jones – EF1194; Dr Hayley Yuille – EF1704, EF3290, EF1261 and EF0927.

5.1.7. Discussion: expression, purification and activities of the entire family of intact membrane histidine kinases from *E. faecalis*

Production of membrane proteins has previously been one of the bottlenecks in determination of structure and activities, and studies of most of the previously studied histidine kinases have been confined to only the soluble domain or truncated forms of the kinase, owing to the difficulties of producing proteins that possess their transmembrane domains. The present study evaluated the heterologous expression, production and activities of sixteen full-length membrane histidine kinases of *E. faecalis*. The genes for each histidine kinase were systemically amplified and cloned into plasmid pTTQ18-His₆, and all but one (EF2219 – HK01) were successfully expressed in *E. coli* membranes (Figure 5.1). Plasmid pTTQ18 has been shown to be successful for production of many membrane proteins (Ward *et al.*, 2000; Saidijam *et al.*, 2003; Bettaney, 2008), and in a comparison of expression data for 37 membrane proteins, out of eight vectors tested, pTTQ18 showed the greatest number of targets expressed (Surade *et al.*, 2006).

The expression trials for the sixteen histidine kinases showed variable expression levels, 3.5 – 16.5% of total membrane proteins (Table 5.1 and Figure 5.1). In this study the choice of host strain, expression medium, induction temperature, concentration of inducer and post-induction time were based on previous experience in the laboratory for expression of transporters and histidine kinases (Ward *et al.*, 2002; Potter *et al.*, 2002; Saidijam *et al.*, 2003). Generally, LB medium at induction temperatures of 30 or 37°C with 0.5 or 1 mM IPTG and 3 hours post-induction time were successful for the overexpression of most of the *E. faecalis* histidine kinases (Table 5.1). Here, high success rates for the solubilisation and purification of the histidine kinases were also achieved. Fourteen out of fifteen were solubilised successfully from *E. coli* membranes and thirteen were successfully purified. Most were successfully solubilised using DDM and the purity of the purified kinases ranged from 19 – 90% in these one-step experiments (Table 5.1 and Figure 5.3).

There are different mechanisms of stimulus perception by histidine kinases, either by periplasmic, transmembrane or cytoplasmic-sensing, as described in the Introduction (Figure 1.4.5). Most previous studies of histidine kinases employed only the cytoplasmic domains that constituted soluble forms lacking their sensing domains. One example, includes a genomic study of the two-component systems of *E. coli* (Yamamoto *et al.*, 2005), which consequently was unable to characterise the sensing

mechanisms. But our work here demonstrates successful production of full-length histidine kinases that generally retain their autophosphorylation activities, either within inner membranes or after purification. Unlike the truncated soluble versions, the full-length histidine kinases are suitable for signal identification. The activity studies showed that eleven of the fifteen membrane-bound proteins and twelve of the thirteen purified proteins were active in autophosphorylation assays (Figures 5.2 and 5.4), and that a remarkable thirteen of the fifteen expressed kinases were active in *E. coli* membranes and/or as purified proteins.

5.2. Trials to identify the signalling ligands for histidine kinase EF3197

5.2.1. Introduction

5.2.1.1. Histidine kinase EF3197: a putative LytS homologue in *E. faecalis*. The histidine kinase EF3197 of *E. faecalis* is 55% identical (85% similar) to *Staphylococcus aureus* LytS and is therefore likely to be the orthologue of the functionally characterised LytSR two-component regulatory system. The *S. aureus* LytSR system was first identified in 1996 (Brunskill & Bayles, 1996a), and is involved in controlling the rate of autolysis during cell wall biosynthesis by affecting the intrinsic murein hydrolase activity of the cell (Brunskill & Bayles, 1996b). Downstream from the *lytSR* genes is a bicistronic operon, *lrgAB*, containing genes that modulate murein hydrolase activity and penicillin tolerance (Groicher et al., 2000). Upon signal perception by LytS, the histidine kinase presumably interacts with its partner response regulator, LytR, which then activates the transcription of genes under its control i.e. *lrgAB* (Figure 5.5).

In bacteria, the LytSR system is associated with the molecular control of bacterial death and lysis. The self-destructive end stage of the bacterial life cycle is termed 'autolysis' and involves the bacteriolytic enzymes referred to as murein or peptidoglycan hydrolases, or autolysins. Purified autolysins were able to dissolve isolated bacterial cell walls, digesting the peptidoglycan layer (Nomura & Hosoda, 1956). Murein hydrolases are widely diverse, a unique family of enzymes that specifically cleave structural components of peptidoglycan, shown to be important in a number of biological processes, such as cell growth and division, daughter cell separation and peptidoglycan recycling and turnover (Perkins, 1980; Shockman *et al.*, 1983; Ward & Williamson, 1985; Holtje *et al.*, 1991; Archibald *et al.*, 1993; Shockman *et al.*, 1994).

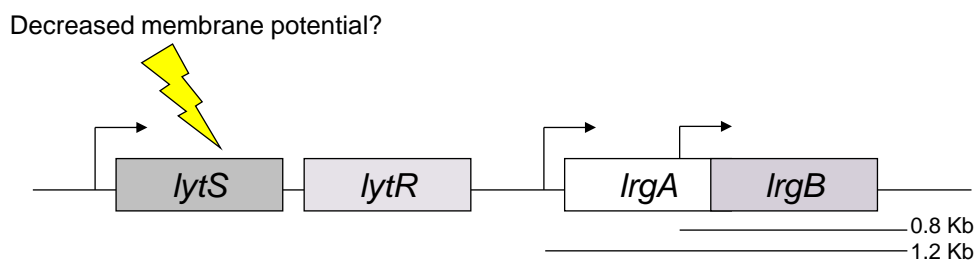


Figure 5.5 LytSR two-component regulatory system of *S. aureus*. The LytSR system is hypothesised to sense decrease in membrane potential and respond by inducing *lrgAB* transcription (Patton *et al.*, 2006). The overlapping *lrgA* and *lrgB* transcripts are indicated by the black bars.

5.2.1.2. Bacterial cell lysis. The best-characterised example of regulated bacterial cell lysis comes from studies of the control of the lytic cycle during bacteriophage infection, in which a ‘holin’ and an ‘endolysin’ (a murein hydrolase) are involved (Young, 1992; Wang *et al.*, 2000). There are two proposed mechanisms for holin-controlled murein hydrolase activity in cell lysis (Figure 5.6).

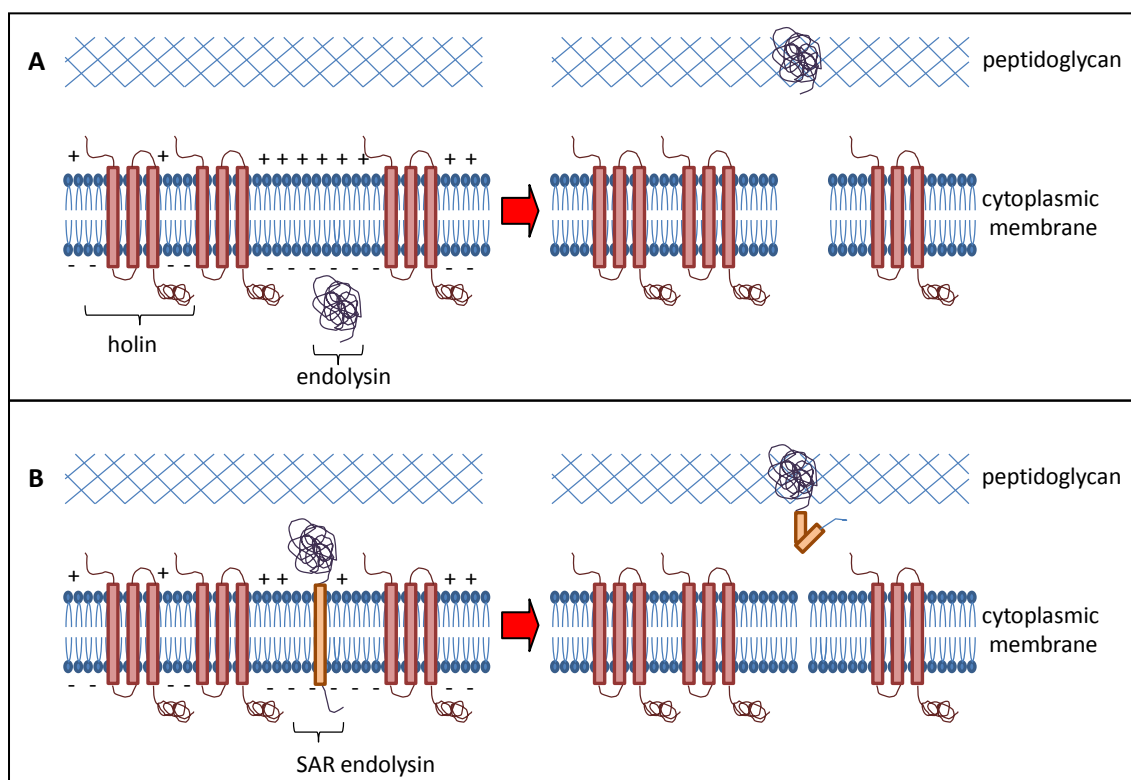


Figure 5.6 Model of the holin and endolysin function. (A) Endolysin (murein hydrolase), such as that produced by bacteriophage lambda, accumulates in an active form in the cytoplasm. Holin forms rafts by helix-helix packing and is activated by localised depolarisation causing changes in holin conformation, leading to asymmetric disruption of the helix packing. The membrane becomes leaky, followed by release of endolysin to the peptidoglycan. (B) SAR-endolysin, such as that produced by bacteriophage P1, is attached to the cytoplasmic membrane in an inactive form by the SAR-domain (orange bar). Upon holin activation the membrane becomes de-energised and the endolysin is released. Adapted from Rice & Bayles (2008).

The first mechanism, utilised by lambda- and T4-like bacteriophage, involves the control of murein hydrolase transport across the membrane. Murein hydrolase typically accumulates harmlessly in the cytoplasm, lacking signal peptides, until holin allows passage across the membrane and thus access to its substrate, peptidoglycan (Figure 5.6). Although the mechanism by which the holins mediate the passage of murein hydrolases is uncertain, it is thought to involve their oligomerisation in the membrane causing membrane permeabilisation and allowing passive passage of murein hydrolase into the periplasm (Gründling *et al.*, 2000; Young *et al.*, 2000; Wang *et al.*, 2003). The second mechanism, utilised by P1 bacteriophage, involves murein hydrolases containing signal-arrest-release (SAR) domains (Xu *et al.*, 2004; Xu *et al.*, 2005). Similar to signal sequences, the SAR domain targets the murein hydrolase to the Sec machinery, but anchors the protein in the outer face of the membrane in an inactive form until holin releases it (Figure 5.6).

5.2.1.3. The *lrgAB* and *cidABC* genes. In *S. aureus*, *lrgA* and *lrgB*, putatively encoding anti-holin-like protein and murein hydrolase regulatory protein respectively, are regulated by the LytSR two-component regulatory system (Brunskill & Bayles, 1996a). Deletion of the *lrgAB* operon caused a decrease in penicillin tolerance for cells that were nearing stationary phase, whereas overexpression of this operon caused increased penicillin tolerance for cells in the early exponential phase, consistent with the observation that *lrgAB* was only minimally expressed during the early exponential phase of growth (Groicher *et al.*, 2000). Changes in membrane potential affecting uptake of antibiotics, such as gentamicin, were observed approximately thirty years ago (Mates *et al.*, 1982). Although the mechanism affecting uptake is unknown, the *lrgAB* genes could putatively be involved by producing holin lesions and thus allowing or preventing antibiotic uptake.

Study of the *lrg* operon later led to the identification of the *cid* operon, a homologue of *lrgAB*, designated *cidA* and *cidB* (23% and 31% amino acid identity with LrgA and LrgB, respectively). CidAB also regulate cell death and lysis but mutants lacking these proteins exhibited phenotypes that were the reverse of those of LrgAB knockouts (Rice *et al.*, 2003); mutants of *cidA* showed decreased murein hydrolase activity and increased tolerance to antibiotics (Rice *et al.*, 2003; Patton *et al.*, 2005; Rice *et al.*, 2005). Expression of the *cid* and *lrg* operons is regulated by LytSR (Brunskill *et al.*, 1996) and CidR (Yang *et al.*, 2005) proteins, which function in response to changes in membrane potential (Patton *et al.*, 2006) and acetic acid

produced during glucose metabolism, respectively (Rice *et al.*, 2005). The accumulation of acetate or acetic acid byproducts in culture media could result from the expression of the *cidC* gene, which encodes a pyruvate oxidase (Patton *et al.*, 2005). The *cidB* and *lrgB* genes encode homologous hydrophobic proteins and the exact function is unknown, but hydrophobicity analyses of these proteins reveal that they are unlikely to encode murein hydrolases. The *cidA* gene encodes a putative holin protein, an effector of murein hydrolase activity and cell lysis, and the *lrgA* gene encodes a putative anti-holin that inhibits these processes; both of which are homologous hydrophobic proteins and are believed to be analogues to bacteriophage-encoded holins and anti-holins respectively (Groicher *et al.*, 2000; Rice *et al.*, 2003).

Recent studies revealed that *cid* and *lrg* operons that regulate death of bacterial cells are important for biofilm development (Bayles, 2007; Rice *et al.*, 2007; Rice & Bayles, 2008), by providing a source of extracellular genomic DNA (eDNA), which is an essential matrix molecule produced by many bacterial species during biofilm development (Whitchurch *et al.*, 2002; Allesen-Holm *et al.*, 2006; Spoering & Gilmore, 2006; Rice *et al.*, 2008). The *cidA* gene product was a positive effector of cell lysis and DNA release in biofilm (Rice *et al.*, 2007), and the *lrg* operon was an inhibitor of cell lysis and inhibition of DNA release in the biofilm (Mann *et al.*, 2009).

5.2.1.4. The histidine kinase EF3197 of *E. faecalis*. Genome sequencing of *E. faecalis* and identification of the seventeen two-component systems led to discovery of a putative LytSR system of *E. faecalis* (EF3197-EF3196), named the LytST two-component regulatory system (Figure 5.7). The histidine kinase EF3197 is 55% and 52% identical to the LytS proteins of *S. aureus* and *B. anthracis*, respectively. EF3196, EF3194 and EF3193 are also highly similar to the *Bacillus* species LytR response regulator, murein hydrolase regulator, and anti-holin-like protein, respectively. A recent study of a *lytS* knockout mutant generated in a clinical *S. aureus* isolate (UAMS-1) showed that LytSR is required for *lrgAB* expression, and that this mutant formed a more adherent biofilm compared to wild-type and complemented strains (Sharma-Kuinkel *et al.*, 2009), suggesting LytS is a negative effector of biofilm formation. The histidine kinase EF3197 was therefore further characterised in this project for its possible function in regulating cell lysis and biofilm development.

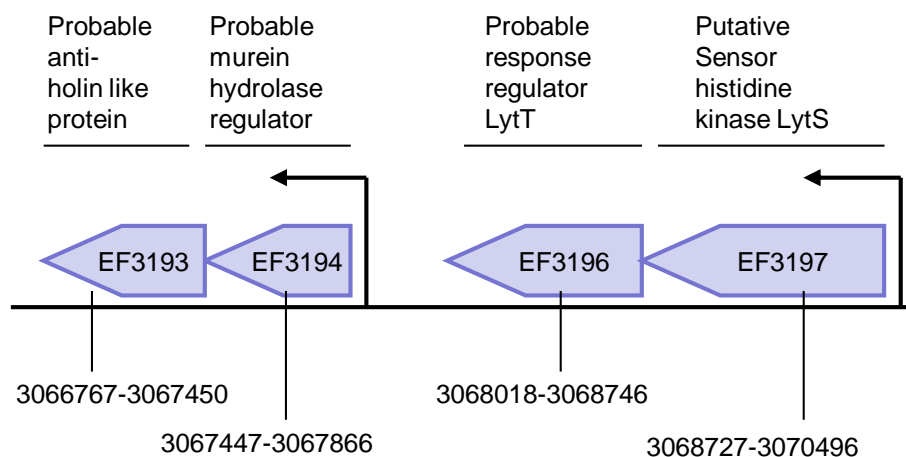


Figure 5.7 Map of the LytST two-component regulatory system in *E. faecalis*. The bent arrows denote promoter sites (\curvearrowright). The blue arrows represent genes (\blacktriangleleft), and the number below corresponds to the genome position of the gene. This diagram is based on the genome map in the KEGG database (Kanehisa, 2002).

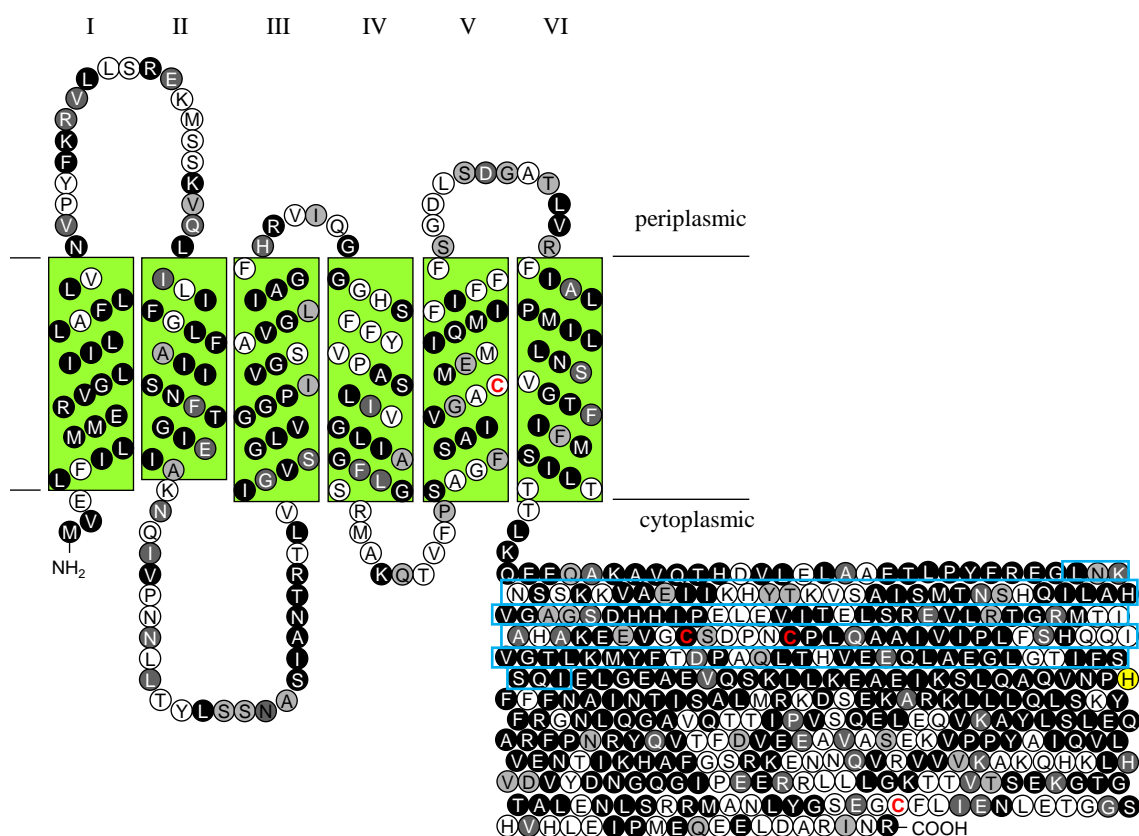


Figure 5.8 Topology model of *E. faecalis* histidine kinase EF3197. Topology prediction by TMHMM (Krogh *et al.*, 2001). The predicted helices are numbered from the N- to C-terminus and are shown in roman numerals. The phosphorylatable histidine is highlighted in yellow, and cysteine residues in red. The residues boxed in blue comprise a putative GAF domain. Conserved residues are highlighted in grey and black according to the alignment in Appendix 9.

EF3197 of *E. faecalis* is classified as a Group I kinase, as determined by the conservation of amino acids flanking the phosphorylatable histidine residue (Hancock & Perego, 2002). It possesses six putative transmembrane helices, followed by a GAF domain presumed to bind cGMP, and a core catalytic domain (Figure 5.8). Alignment of the EF3197 primary sequence with the LytS protein sequences of *S. aureus* and *Bacillus* strains revealed that these proteins are highly similar with many conserved sequence (Appendix 9).

EF3197 clusters most closely to LytS of *S. aureus* on the phylogenetic tree, and the LytS proteins of *Bacillus* species also clustered very closely (Figure 5.9). Included in the phylogenetic tree analyses are redox-sensing kinases; EF3197 was shown not to be closely related to these proteins (Figure 5.9). The SMART web-based tool showed that EF3197 and LytS of other bacteria have a conserved GAF domain, and that the sizes and positions of the conserved domains in these proteins are highly similar (Figure 5.10). The redox sensors were also compared because EF3197 was shown to respond to redox (Section 5.2.7).

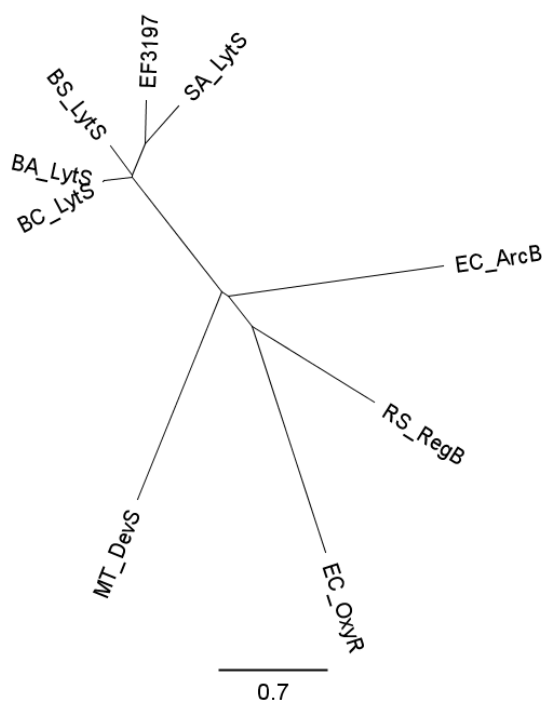


Figure 5.9 Phylogenetic tree of EF3197, LytS and redox sensors. Phylogenetic tree was created by Geneious Pro 4.6.5 software using amino acid sequences of EF3197 and the LytS proteins. LytS of: *Bacillus subtilis* (BS_LytS); *Bacillus cereus* (BC_LytS); *Bacillus anthracis* (BA_LytS); *Staphylococcus aureus* (SA_LytS). ArcB of *E. coli* (EC_ArcB). RegB of *Rhodobacter sphaeroides* (RS_RegB). OxyR of *E. coli* (EC_OxyR). DevS of *Mycobacterium tuberculosis* (MT_DevS).

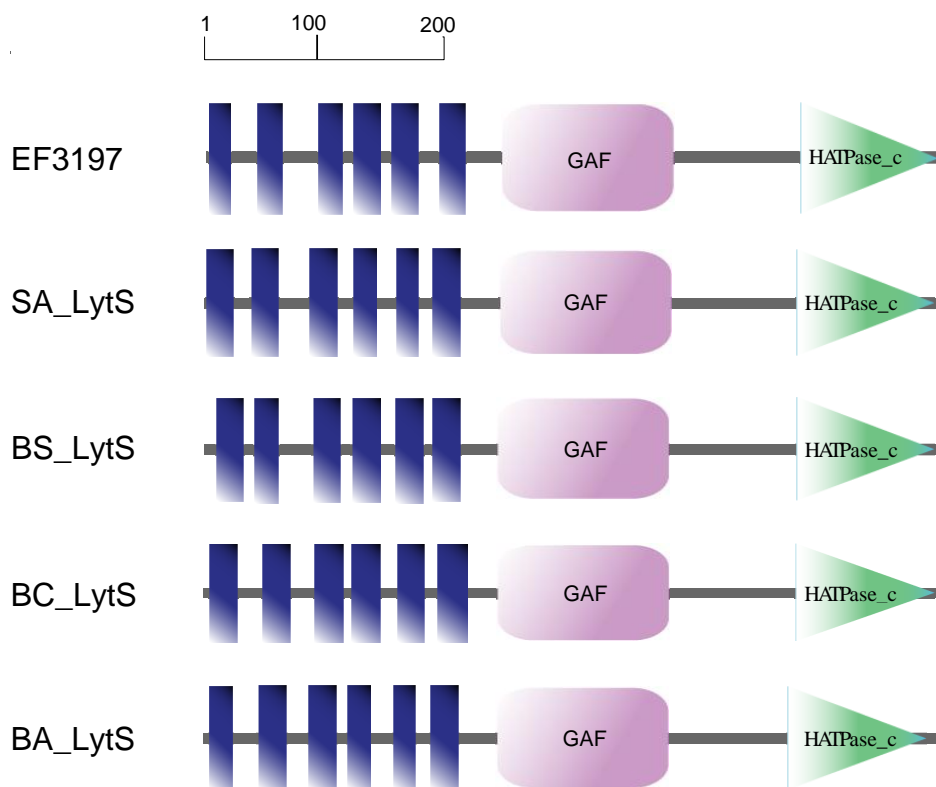


Figure 5.10 Domain analysis of EF3197 and LytS of *S. aureus* and *Bacillus*. The figure is based on the graphical output of the SMART web-based tool (Schultz *et al.*, 1998; Letunic *et al.*, 2008). The scale bar is in amino acids. Blue vertical bars represent putative transmembrane segments (■). Sizes and positions of conserved domains are indicated by the labeled symbols. EF3197 (putative LytS of *E. faecalis*), SA_LytS (LytS of *S. aureus*), BS_LytS (LytS of *B. subtilis*), BC_LytS (LytS of *B. cereus*), BA_LytS (LytS of *B. anthracis*).

5.2.2. Cloning, expression and purification of EF3197

The EF3197 gene was amplified from the chromosomal DNA of *E. faecalis* V583 strain by PCR and cloned into plasmid pTTQ18-His₆, followed by expression trials. The results are shown in Section 5.1.2 of this chapter. Histidine kinase EF3197 was heterologously expressed in *E. coli* membranes (Figure 5.1) to a level of 12% of total membrane proteins, migrating at 53 kDa in SDS-PAGE (Table 5.1). Expression of EF3197 in LB, 2TY and M9 minimal media suggests that the protein is present at 12, 9 and 5% of total *E. coli* membranes, respectively (Figure 5.11), as determined by scanning densitometry; therefore LB medium was used for subsequent expression of EF3197.

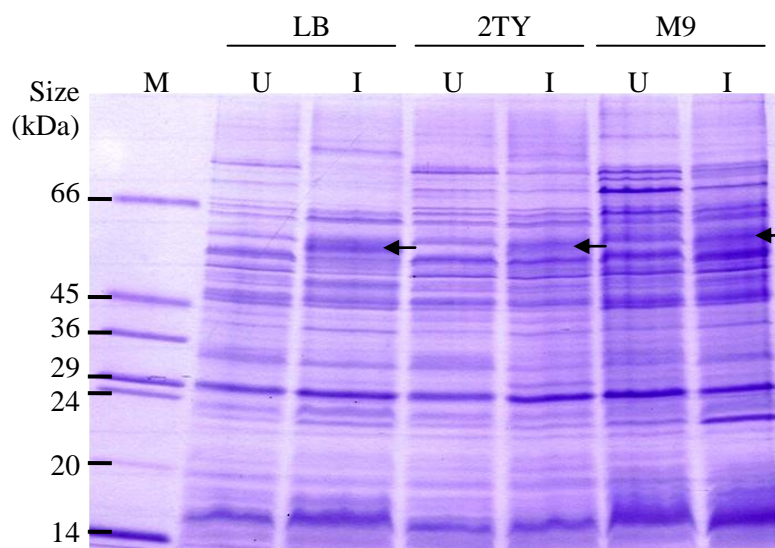


Figure 5.11 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF3197)-His₆. BL21(DE3) cells transformed with pTTQ18(EF3197)-His₆ were grown in LB, 2TY or minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction, followed by total *E. coli* membranes preparation and analysis by SDS-PAGE. M (marker), U (uninduced) and I (induced). Arrow indicates position of EF3197 deduced from previous SDS-PAGE and Western blots (Figure 5.1).

A six litre culture of *E. coli* BL21(DE3) cells expressing His₆-tagged EF3197 was grown aerobically at 37°C (Section 2.4.1). The harvested cells were lysed by explosive decompression and membranes separated on sucrose density gradients (Section 2.5.1). The membrane fractions obtained were analysed by SDS-PAGE, and showed the presence of EF3197 mostly in the inner membrane fraction (Figure 5.12). EF3197 was expressed at approximately 14, 4 and <1% in the prepared inner, mixed and outer membranes respectively (Figure 5.12). The inner membrane proteins were subjected to solubilisation and purification using 0.05% DDM added to all buffers post-solubilisation, as described in Section 2.5.3.2. Fractions from the purifications were analysed by SDS-PAGE (Figure 5.13).

EF3197 was successfully solubilised from inner membranes (Figure 5.13), and separated from a protein band migrating at a slightly lower molecular weight in the insoluble fractions (Figure 5.13). Following analysis by scanning densitometry, the purity of EF3197 was 85% when 0.05% DDM was included in buffers post-solubilisation (Figure 5.13), compared to 57% when no DDM was added (Figure 5.3). Addition of DDM to post-solubilisation buffers improved the purity of EF3197.

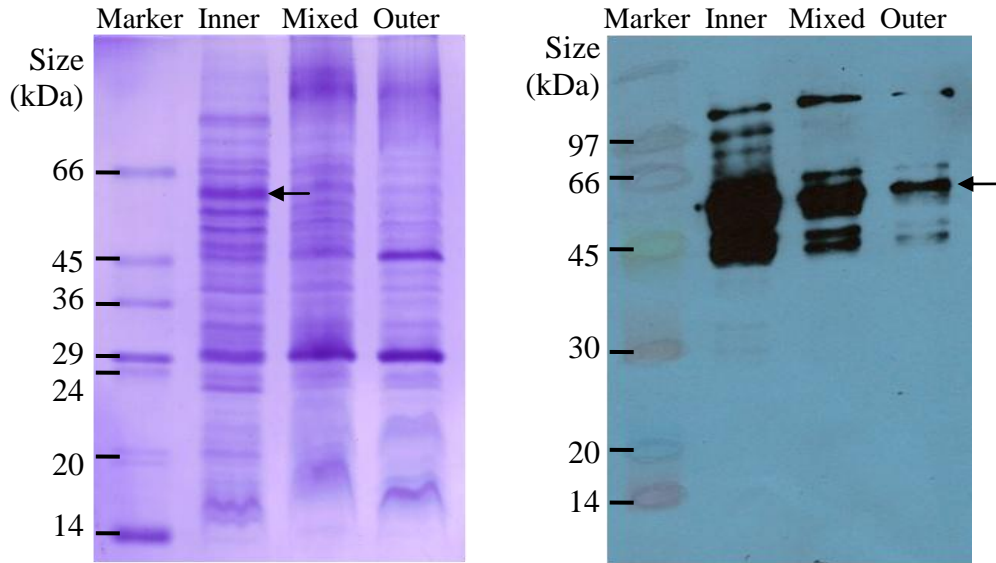


Figure 5.12 SDS-PAGE of membrane fractions of *E. coli* BL21(DE3) cells expressing EF3197. Total *E. coli* membranes were prepared from *E. coli* BL21(DE3) pTTQ18(EF3197)-His₆ grown on LB medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes were separated on 25-55% sucrose gradients and membrane fractions extracted from the gradients were analysed by SDS-PAGE. The arrow indicates position of EF3197.

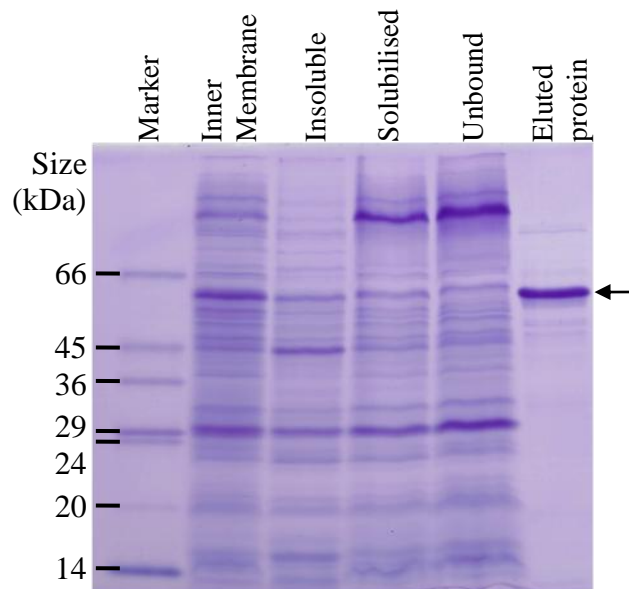


Figure 5.13 SDS-PAGE of fractions obtained from the purification of *E. faecalis* EF3197. EF3197 was purified from the inner membranes of *E. coli* BL21(DE3) pTTQ18(EF3197)-His₆ cells. The inner membranes proteins were solubilised in 1% DDM. Fractions (15 µg) and final eluted protein fraction (5 µg) were analysed by SDS-PAGE. The arrow indicates the position of EF3197.

Although increased purity could be achieved through optimisation trials using different detergent levels, the phosphorylation activity of this histidine kinase was reduced when detergents were included. This was first shown to be true for VicK (Phillips-Jones, personal communication) and later for other histidine kinases tested. One example is shown in Figure 5.31. This histidine kinase and all other kinases were therefore purified according to the procedure in Section 2.5.3.2 in which no detergent was included in post-solubilisation buffers, unless stated otherwise.

5.2.3. Circular dichroism of purified EF3197

The secondary structure of EF3197 was examined by circular dichroism (Section 2.6.1). The negative indentation at 208 and 222 nm, and a positive signal at ~190 nm, are indicative of a predominantly α -helical content (Figure 5.14), agreeing with the prediction by TMHMM that this protein is highly hydrophobic, and confirming the retention of the secondary structure of EF3197 after purification.

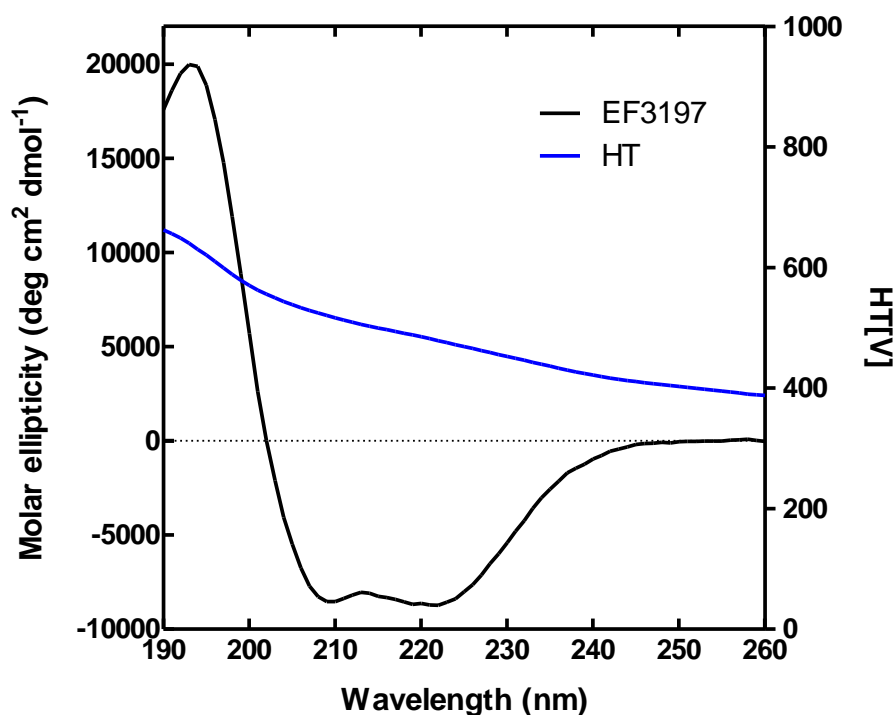


Figure 5.14 Circular dichroism spectrum of purified EF3197. Purified EF3197 (0.05mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of EF3197 was performed using Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample were analysed in Hellma quartz-glass cell of 1 mm path length. Spectrum was recorded with 1 nm step resolution at a scan rate of 10 nm/min. Response time was set at 1 second with a sensitivity of 20 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of ten scans, from which the buffer contribution was subtracted. The blue line represents the voltage applied to the photomultiplier.

5.2.4. Autophosphorylation activities of EF3197

Different phosphorylation kinetics were observed for EF3197 when bound in *E. coli* inner membranes compared with the purified version (Figure 5.2 and 5.4). Using inner membrane preparations, EF3197 exhibited rapid autophosphorylation up to the first minute, followed by a slow desphosphorylation after 10 minutes (Figure 5.2). In the purified form, EF3197 showed continuing phosphorylation and maximum phosphorylation was not reached even in sixty minutes (Figure 5.4). As indicated earlier in this chapter, the membrane vesicles containing *E. coli* phosphatase may give rise to desphosphorylation of kinases. However, as EF3197 is highly identical to the histidine kinase LytS of *S. aureus* that is involved in cell wall metabolism, it is possible there are other components in the prepared membrane vesicles that could give rise to the dephosphorylation effect, for example cell wall derivatives.

5.2.5. Autophosphorylation of purified EF3197 is not affected by penicillin

Previous studies indicated that a *S. aureus lrgAB* mutant produced increased level of extracellular murein hydrolase activity compared to the wild-type strain, and complementation of the *lrgAB* defect restored the wild-type phenotype, indicating that these genes conferred negative control on extracellular murein hydrolase activity (Groicher *et al.*, 2000). Furthermore, expression of the *lrgAB* operon inhibited penicillin-induced killing of *S. aureus*, independently of cell lysis (Groicher *et al.*, 2000). Since the LytSR two-component system of *S. aureus* regulates the *lrgAB* operon, and EF3197 is highly homologous to LytS of *S. aureus*, experiments were therefore performed to determine whether penicillin exerted any effect on the autophosphorylation activities of EF3197.

Autophosphorylation assays were performed in the presence of a range of penicillin concentrations (0 – 10 mM) (Section 2.5.5.3). It was found that the presence of the antibiotic exerted no detectable effect on the phosphorylation of purified EF3197 (Figure 5.15). However, it is possible that penicillin exerts an effect on dephosphorylation rate of EF3197, rather than autophosphorylation rate, but this was not investigated in this study.

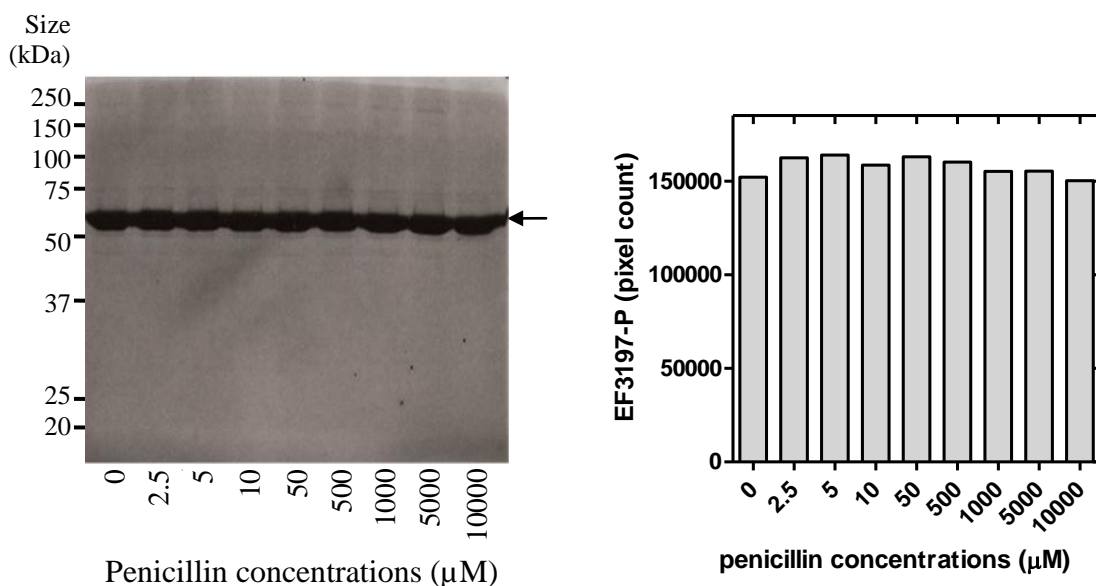


Figure 5.15 Autophosphorylation of purified EF3197 in response to the presence and absence of penicillin. Each reaction employed 80 pmoles (5.3 μM) of purified EF3197 in a final reaction volume of 15 μl. Reactions were initiated by the addition of 50 μM radiolabelled [γ - 33 P]ATP and were allowed to proceed for 40 minutes in the presence of the indicated penicillin concentrations (Section 2.5.5.3). Autoradiographic film (left) and graphical representation (right) of phosphorylated EF3197 are shown.

5.2.6. Signal testing of EF3197

The molecular signal for EF3197 is unknown and therefore a range of candidate signals were screened using purified protein in *in vitro* autophosphorylation assays. This includes testing monovalent cations, temperature and sodium fumarate, which were previously shown to affect the rate of cell wall autolysis (Cheung & Freese, 1985; Madiraju *et al.*, 1987; Tobin *et al.*, 1994); hence Na⁺, K⁺, Mg²⁺ and Fe²⁺ (NaCl, K₂HPO₄/KH₂PO₄, MgCl₂, Fe₂SO₄) were tested. Also, cGMP was tested since EF3197 was predicted to have a putative GAF domain (Figure 5.8 and 5.10). Glucose and acetic acid were tested as they were shown to affect CidR which stimulated *cid* and *lrg* transcription (Rice *et al.*, 2003; Patton *et al.*, 2005; Rice *et al.*, 2005).

In the signal testing assays, 80 pmoles of purified EF3197 (5.3 μM) were added per 15 μl reaction in the presence of each candidate signal and assayed as described in Section 2.5.5.3. The proportion of phosphorylated EF3197 was determined by scanning densitometry and represented graphically (Figure 5.16). Results indicated that Fe₂SO₄, acetic acid and temperature at 55°C greatly reduced the proportion of phosphorylated EF3197 (EF3197-P) to 3, 2 and 7% respectively (Figure 5.16), suggesting that EF3197 is sensitive to acidic conditions and high temperature stress conditions.

Phosphohistidine residues are known to be acid labile (Attwood *et al.*, 2007) and therefore this may also be responsible for the loss of activity. However Fe^{2+} needs to be retested at an appropriate pH. Meanwhile, the inactivity of the kinase at high temperature could be caused by denaturation of the protein and therefore a circular dichroism spectrum or thermal melt of this kinase could be performed to test if the protein is unfolded at the tested temperature. MgCl_2 , cGMP and sucrose caused a reduction in the proportion of EF3197-P (84, 72 and 80% respectively); whereas DTT, NaCl, erythromycin (dissolved in 100% ethanol), $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, glucose and sodium fumarate caused an increase in proportion of EF3197-P (Figure 5.16). All candidate signals tested showed possible effects on the phosphorylation of EF3197, though without further repeat experiments no statistical analyses can be performed and therefore no firm conclusions can yet be drawn.

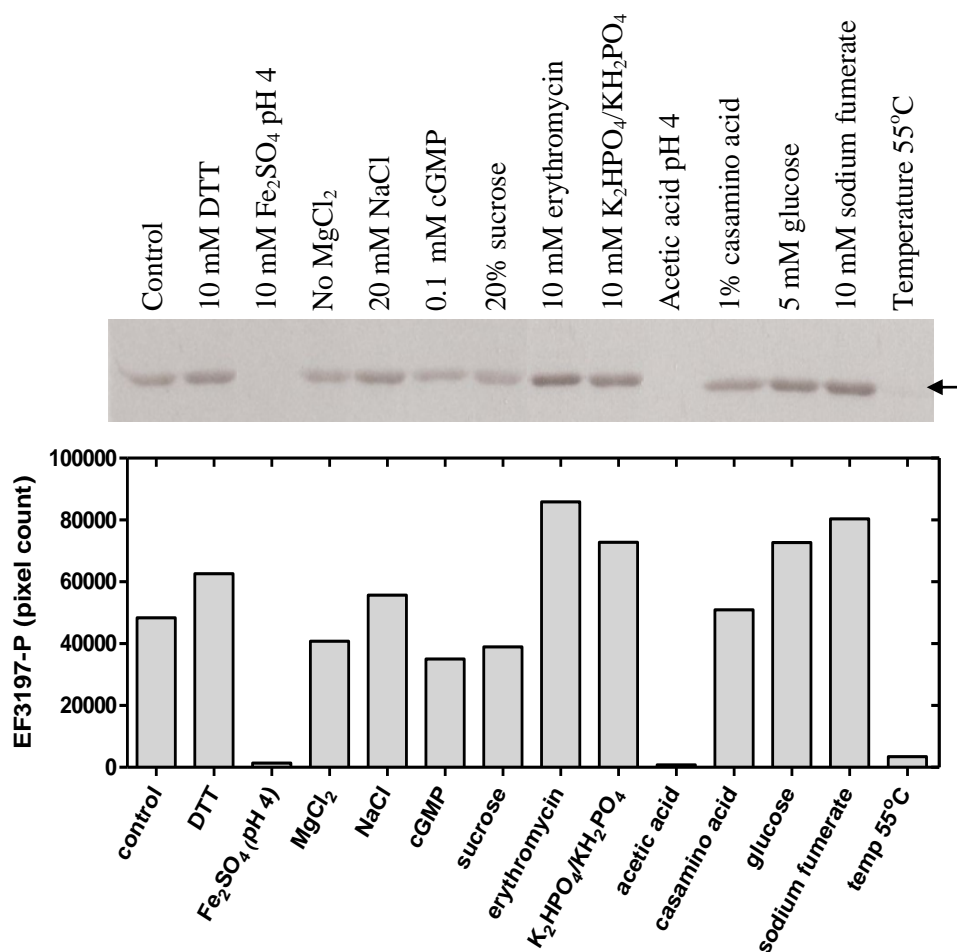


Figure 5.16 *In vitro* signal screening for EF3197. Purified EF3197 (80 pmoles or 5.3 μM) was incubated in the presence of 50 mM Tris-HCl pH 7.6, 10 mM MgCl_2 and 50 mM KCl and in the presence and absence of candidate signals for 20 minutes in a final reaction volume of 15 μl . Reactions were initiated and were allowed to proceed for 15 minutes (Section 2.5.5.3). Autoradiographic film (top) and graphical representation (bottom) of phosphorylated EF3197. Control (no signal added).

5.2.7. Changes in redox potential affect EF3197 autophosphorylation

To determine whether reducing agents exert any effect on the autophosphorylation activity of EF3197, assays were prepared using 80 pmoles of purified EF3197 (5.3 μ M) per 15 μ l reaction in the presence of a range of reducing agent concentrations (Section 2.5.5.3). Initiation of reactions was through addition of ATP. The results indicated that the proportion of EF3197-P was increased by 21% in the presence of 1 mM DTT; and with further increases in DTT concentrations (5 – 20 mM), a gradual decline in the proportion of EF3197-P was observed (Figure 5.17). When DTT concentration was increased to 50 mM, the proportion of EF3197-P was greatly reduced to 12% of the maximum (Figure 5.17). Further increase in DTT concentrations, 75 and 100 mM, showed little further effect on the proportion of EF3197-P (Figure 5.17).

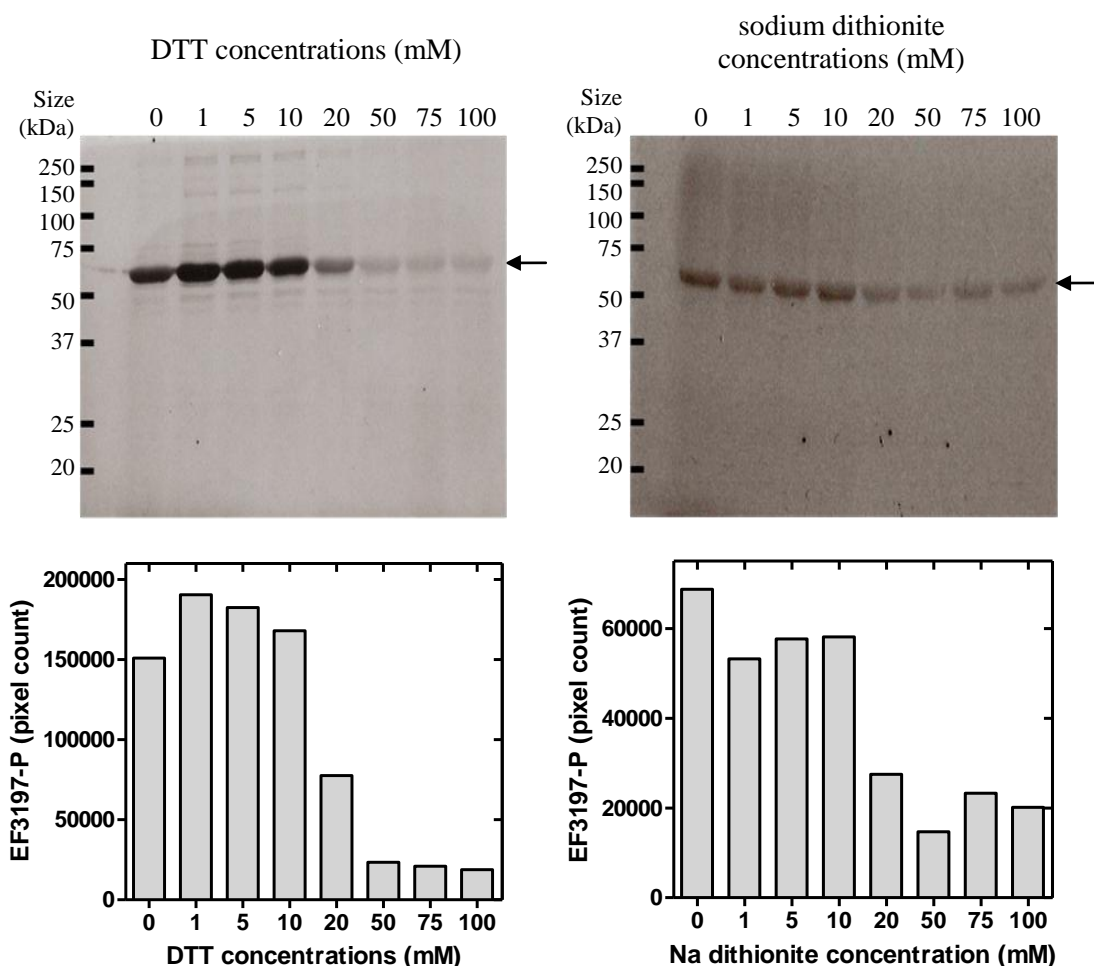


Figure 5.17 Effect of reducing agents on autophosphorylation activity of EF3197 *in vitro*. DTT or sodium dithionite (0 – 100 mM) with 80 pmoles (5.3 μ M) purified EF3197 were added to a final reaction volume of 15 μ l. Reactions were initiated and were allowed to proceed for 40 minutes (Section 2.5.5.3). Autoradiographic film (top) and graphical representation (bottom) of phosphorylated EF3197.

A similar effect was observed using an alternative reducing agent, sodium dithionite. Increasing concentrations of dithionite caused a decrease in levels of EF3197-P (Figure 5.17). The proportion of EF3197-P was maximal in the absence of sodium dithionite (0 mM), and the greatest change was observed in the presence of between 10 and 20 mM sodium dithionite, which decreased levels from 85% to 40% of maximum respectively (Figure 5.17). Between 50 and 100 mM sodium dithionite, the proportion of EF3197-P was lowest, 21 – 34% of the maximum EF3197-P (Figure 5.17).

5.2.8. Discussion: towards identification of a signalling ligand for EF3197

Prior to this study, the function of EF3197 was unknown, and therefore my *in silico* analyses, such as topology prediction and searches for putative domains, found that the two-component system EF3197-EF3196 is highly similar to the LytSR of *S. aureus*; the histidine kinase EF3197 possess 55% identity to LytS (Brunskill & Bayles, 1996a; Hancock & Perego, 2002), suggesting a possible role for the enterococcal system in cell wall integrity, mediated through the *lrgAB* operon that regulates murein hydrolase activity of the cells. However, Hancock and Perego found no change in autolysis rates in a mutant of response regulator EF3196, and therefore the role was unconfirmed, but alterations in uptake of antibiotics such as erythromycin were observed in the mutant, suggesting some involvement in cell wall-related permeability function (Hancock & Perego, 2004b).

Recently, the LytSR system of *B. anthracis* was identified and it was shown to have an important role in the control of cell death and lysis, as for the LytSR system in *S. aureus*, and to positively regulate *lrgAB* expression. Transcription of *lrgAB* was induced by dissipation of the proton motive force in a *lytSR*-dependent manner (Chandramohan *et al.*, 2009). The enterococcal two-component system EF3197-EF3196 which is highly identical to these systems (Appendix 9) is therefore likely to have the same important function.

Successful production of active EF3197 (Figure 5.2 and 5.4) allowed investigation of the signal for this protein. The candidate signals tested showed effects on the phosphorylation of EF3197 (Figure 5.16) and these need to be further investigated in future studies; for example, a range of concentrations of the candidate signal should be tested to determine the extent of any effect on EF3197 phosphorylation. Here, EF3197 was shown not to respond to increasing concentrations of penicillin (0 – 10 mM) (Figure 5.15). Perhaps penicillin has an indirect effect. Penicillin is known to

bind to the DD-transpeptidase that links peptidoglycan molecules, thereby weakening the cell wall via the peptidoglycan layer (Yocum *et al.*, 1979), which may in turn induce phosphorylation of EF3197 that activates the anti-holin like protein and murein hydrolase regulator and therefore prevents cell lysis, and leads to penicillin tolerance. The penicillin test should therefore be repeated with EF3197 bound in the inner membrane.

A reducing environment generated by DTT or sodium dithionite, *in vitro*, caused a decreased in the proportion of phosphorylated protein (Figure 5.17), suggesting a role of EF3197 in redox-sensing.

A sequence alignment of EF3197 with LytS of *S. aureus* and *Bacillus* species showed the presence of two conserved cysteine residues in a putative GAF domain (Appendix 9). Cysteine residues play important roles in the biochemistry of many proteins because of their redox properties, and ability to coordinate metal ions. Cysteine is therefore a common amino acid involved in forming key catalytic components of enzymes (Barford, 2004). In the past decade, cysteines have been recognised in reversible thiol modifications in response to reactive oxygen or nitrogen oxide species, as well as having important roles in gene transcription, metabolism and signal transduction cascades in both prokaryotic and eukaryotic cells (Paget & Buttner, 2003; Brandes *et al.*, 2009). Some bacterial histidine kinases, RegB and ArcB, were also found to have redox-active cysteine residues involved in regulating the environmental redox potential (Swem *et al.*, 2003; Malpica *et al.*, 2004). It appears that redox-active cysteines are widespread in bacteria, for example the electron transporter, DsbD (Katzen & Beckwith, 2003), and the thioredoxin-like protein, CcmG (Edeling *et al.*, 2004), both possess redox-active cysteine centres that play an important role in their function.

The nature of the signal that modulates activity of EF3197 was not identified further due to the time constraints of this study but there are further tests which could be performed and are mentioned below. As EF3197 was shown to respond to reducing agents, DTT and sodium dithionite (Figure 3.13), the conserved cysteine residues could play a role in regulation of protein activity. Additionally, to test whether EF3197 is a true redox-sensing protein, further tests such as use of hydrogen peroxide to determine whether the reaction is reversible in the presence of oxidising agents, and mutation of the cysteine residues will be important for determining how these residues contribute to its function. Since DTT, a metal chelating agent, had a greater effect on

phosphorylation of EF3197 than sodium dithionite, this might suggest metal is a cofactor for redox control *in vitro*, and therefore should be tested for metal dependence; for example, effects of Cu^{2+} and Fe^{2+} on the protein.

Bacteria are able to employ various mechanisms to sense redox and oxygen (Bauer *et al.*, 1999; Green & Paget, 2004). The conserved cysteine residues of the LytS proteins (Appendix 9) could have important roles in redox-sensing or otherwise could be involved in oligomerisation of the protein into the active or inactive form, just like the global regulator RegB of *R. capsulatus*, which under oxidising conditions resulted in formation of a metal-dependent intermolecular disulphide bond that converted the kinase from an active dimer to an inactive tetramer state (Swem *et al.*, 2003). Most histidine kinases have only two transmembrane helices (Stock *et al.*, 1989), but analysis of the amino acid sequence of LytS revealed six potential transmembrane domains (Figure 5.10). These multiple domains could be ideally suited for sensing a membrane-associated signal such as $\Delta\Psi$. Therefore, the effects of $\Delta\Psi$ -dissipating agents such as valinomycin (Ahmed & Booth, 1983; Rose & Jenkins, 2007) on membrane expressed EF3197 activity could also be tested, as well as producing truncated forms of EF3197 to determine the position of the sensing domain.

EF3197 might respond to signals of cell energy status. Another redox sensor, RegB of *R. capsulatus*, was shown to respond to oxidised ubiquinone resulting in decreased phosphorylation activity (Swen *et al.*, 2006). Sensor histidine kinase PrrB of *R. sphaeroides*, which also responds to DTT (Potter *et al.*, 2006) is known to be connected to electron flow through the terminal *cbb*₃-type cytochrome *c* oxidase of the electron transport chain (Oh & Kaplan *et al.*, 2000). Both these redox sensors appear to have connection with components of the respiratory systems. Thus, it is possible that EF3197 also senses such signals, or alternatively the energy status of the cell, and its phosphorylation activity in turn regulates of the *lrg* genes thereby contributing to a role in cell lysis or survival.

GAF and PAS domains are ubiquitous cytoplasmic signalling units in prokaryotic and eukaryotic organisms (Martinez *et al.*, 2002; Galperin, 2004), and are capable of binding small molecules, such as flavin, adenine and guanine, and heme as well as detecting changes in redox and light (Zhulin *et al.*, 1997; Sardiwal *et al.*, 2005; Ioanoviciu *et al.*, 2007; Sousa *et al.*, 2007). These domains are unrelated in amino acid sequence, but share similarities in their structural topologies (Ho *et al.*, 2000). The role of the conserved cysteine residues found in the putative GAF domain of EF3197 (Figure

5.8 and Appendix 9) needs to be investigated further. Some sensory kinases, such as DosT and DevS of *Mycobacterium tuberculosis*, involved in direct and specific oxygen sensing, contain two GAF domains, one of which has a heme prosthetic group and binding of oxygen leads to inhibited kinases activity (Sousa *et al.*, 2007; Podust *et al.*, 2008). Therefore, it is possible that the highly conserved GAF domain in the LytS proteins could have an important role in signal sensing, especially the conserved cysteine residues that is positioned in the GAF domain.

The exact molecular mechanism of the *lrg* and *cid*-regulated pathways remains unclear, but it is presumed that when the cell is under stress, for example the presence of antibiotics, holin is postulated to oligomerise within the bacterial membrane, creating lesions that disrupt the proton motive force, which then cause loss in cell viability. These lesions create a path for the murein hydrolases to access cell wall substrates, or perhaps even cause the release of eDNA for biofilm formation. In unstressed respiring cells, antiholin binds to holin and prevents break down of the cell wall, whilst constitutive activity of holin gives rise to cell wall metabolism necessary for cell growth and renewal.

A model of *lrg*- and *cid*- regulated cell wall metabolism, which summarises material in the literature, is presented in Figure 5.18. The cysteine residues are also shown. A decrease in membrane potential, proton motive force or changes in redox potential (oxidising conditions) is sensed by histidine kinase LytS, which then undergoes a phosphorylation cascade leading to transcription of the *lrg* genes that encode anti-holin and murein hydrolase regulatory proteins, and resulting in prevention of cell lysis (Rice & Bayles, 2008). Alternatively, glucose metabolism in the cell leads to accumulation of acetate, causing activation of CidR (Patton *et al.*, 2006). Activated CidR binds to the *cidABC* and *lrgAB* promoters and increases the transcription of these operons. The *cid* operon encodes for holin and a murein hydrolase regulatory protein that induce cell wall metabolism.

Overall, the results of this work provide an insight into signals that affect autophosphorylation of EF3197 *in vitro*, and highlights the potential roles of the conserved cysteine residues amongst LytS homologues. Continued studies of the histidine kinase EF3197, together with LytS of other bacteria, including signal testing could provide further details on the nature of signals sensed by these proteins. Alongside this, crystallisation studies could be performed to elucidate the structure, to provide additional information relating to structure and function.

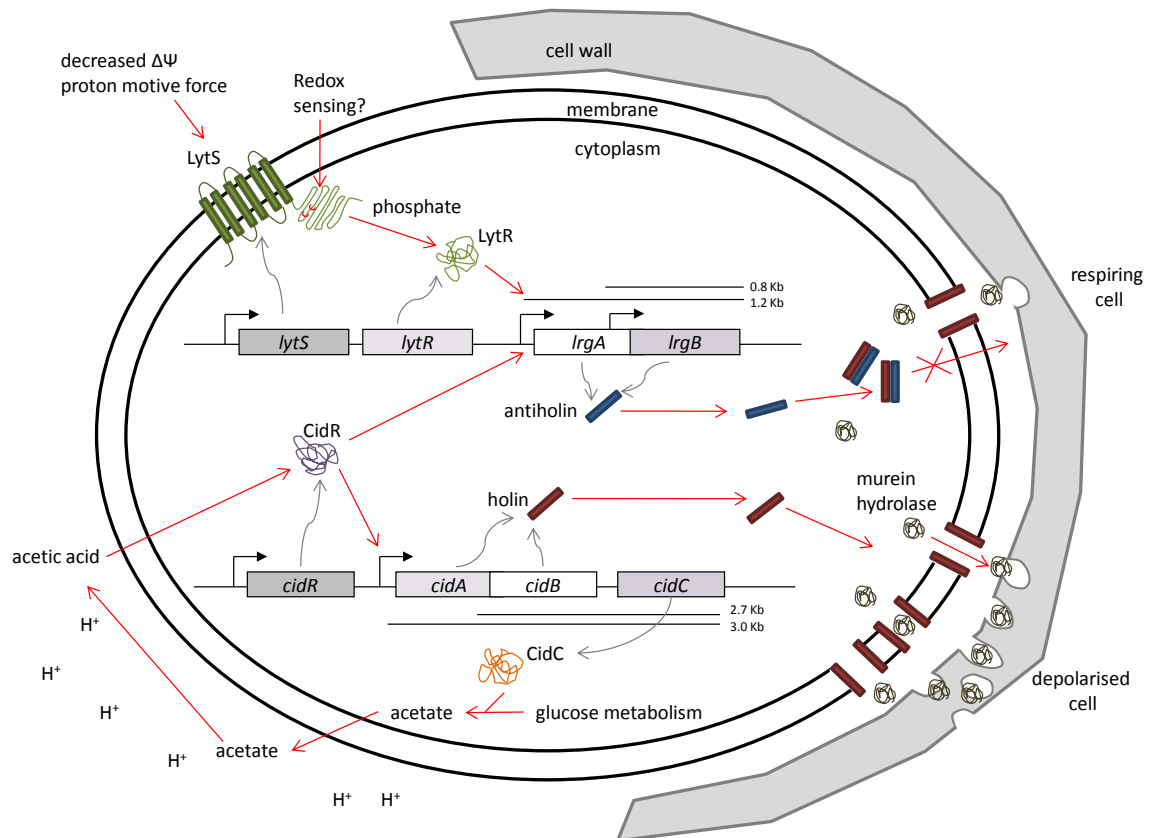


Figure 5.18 A model of the *lrg*- and *cid*- regulated cell wall metabolism. The *cid* genes encode for effectors of murein hydrolase activity and cell lysis, and the *lrg* genes encode for proteins that inhibit the process.

5.3. Structural studies: crystallisation trials of histidine kinase EF1051

5.3.1. Introduction

5.3.1.1. A similarity search of histidine kinase EF1051 (EtaS). A similarity search of histidine kinase EF1051 using FASTA3 (Pearson, 2000) indicates this protein is most similar to putative sensory proteins of the enterococcal species, sharing 69.7 – 100% identity and 90.4 – 100% similarity, suggesting that this protein is highly conserved within the enterococcal species. Next, EF1051 is most similar to proteins of the *Lactobacillus* and *Streptococcus* species, having 35.8 – 46.6% identity and 71.7 – 79.8% similarity. The enterococcal EF1051 shows strong similarity to the LisK of *Listeria monocytogenes* (41.4% identity and 72.2 % similarity) (Cotter *et al.*, 1999), CsrS of *S. agalactiae* (35.8% identity and 71.7%) (Lamy *et al.*, 2004), CsrS of *S. pyrogenes* (35.7% identity and 71.1% similarity) (Levin & Wessels, 1998; Federle *et al.*, 1999) and ArlS of *S. aureus* (32.1% identity and 70.7% similarity) and these proteins are involved in virulence (see below).

An alignment of these proteins with EF1051 (Appendix 10) showed highly conserved residues in the phosphorylatable histidine kinase region and in the most of the cytoplasmic domain, whereas the transmembrane and periplasmic region are more variable (Appendix 10). Construction of a phylogenetic tree reveals that EF1051 clusters most closely to the LisK of *L. monocytogenes* (Figure 5.19), and SMART analysis of these proteins suggests that the size and position of the domains are highly similar (Figure 5.20). All proteins have two putative transmembrane regions, a HAMP (histidine kinases, adenylyl cyclises, methyl binding proteins, phosphatases) domain, a HisKA (phosphoacceptor) domain, and a histidine kinase-like ATPase domain (Figure 5.20).

The LisRK system of *L. monocytogenes* plays a significant role in the virulence potential of this food-borne pathogen. In addition to the contribution of this system in the tolerance of ethanol, pH, hydrogen peroxide stresses and osmolarity (Cotter *et al.*, 1999; Kallipolitis *et al.*, 2001; Sleator & Hill, 2005), LisRK was shown to be involved in the ability of the cell to tolerate antimicrobial agents used in food and in medicine, such as lantibiotic nisin and the cephalosporin family of antibiotics (Cotter *et al.*, 2002). It was also demonstrated, using a mouse model, that a deletion in the LisK histidine kinase component resulted in a 10-fold reduction in virulence compared to the parent wild-type strain (Cotter *et al.*, 1999), confirming the importance of the LisRK system in listerial virulence.

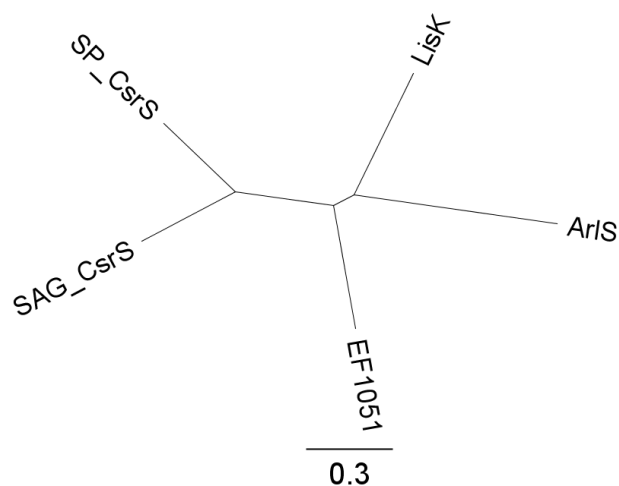


Figure 5.19 Phylogenetic tree of EF1051, LisK, CsrS and ArlS. The phylogenetic tree was created by Geneious Pro 4.6.5 software using amino acid sequences of EF1051, LisK, CsrS and ArlS. EF1051 (histidine kinase of *E. faecalis*); LisK (LisK of *L. monocytogenes*), SAG_CsrS (CsrS of *Streptococcus agalactiae*), SP_CsrC (CsrC of *S. pyogenes*), ArlS (ArlS of *S. aureus*).

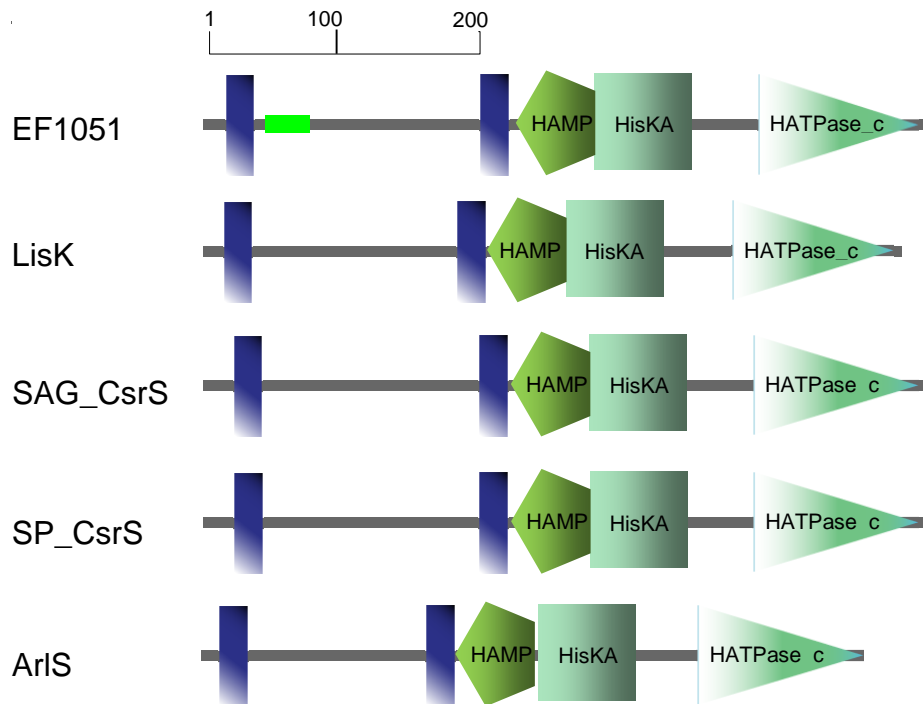


Figure 5.20 Domain analysis of EF1051 and its homologues. The figure is based on the graphical output of the SMART web-based tool (Schultz *et al.*, 1998; Letunic *et al.*, 2008). The scale bar is in amino acids. Blue vertical bars represent putative transmembrane segments (■). Green horizontal bar represents coil regions (■). Sizes and positions of conserved domains are indicated by the labeled symbols. LisK (LisK of *L. monocytogenes*), SAG_CsrS (CsrS of *Streptococcus agalactiae*), SP_CsrS (CsrC of *S. pyogenes*), ArlS (ArlS of *S. aureus*).

The CsrRS system was found to regulate hyaluronic acid capsule synthesis in group A *Streptococcus* (Levin & Wessels, 1998), and is therefore also considered to be a virulence factor. An in-frame deletion mutation within response regulator *csrR*, resulted in sixfold increase in capsule production, producing mucoid colonies, and a corresponding increase in transcription of the *has* operon, which contains the essential genes for hyaluronic acid synthesis, and indicated that CsrR as a negative regulator of hyaluronic acid capsule synthesis and virulence (Levin & Wessel, 1998).

5.3.1.2. The EF1051-EF1050 system of *E. faecalis*. Using both a mouse peritonitis model and peritoneal macrophages, the EF1051-EF1050 system (Figure 5.21) has been shown to be involved in the stress response and virulence of *E. faecalis* and was named EtaSR (Teng *et al.*, 2002; Muller *et al.*, 2008). An *etaR* (EF1050) disruption mutant was shown to be more sensitive to low pH than the parent strain, OG1RF, during logarithmic growth, similar to the phenotype of the *lisK* deletion mutant (Cotter *et al.*, 1999), and it also exhibited altered tolerance to high temperature (Teng *et al.*, 2002). In

common with the homologous CsrRS system, the EF1051-EF1050 system may also negatively regulate capsule synthesis, since an *E. faecalis etaR* mutant, when recovered from the mouse spleen, was mucoid on plates, while the wild-type OG1RF was nonmucoid (Levin & Wessels, 1998; Federle *et al.*, 1999; Heath *et al.*, 1999; Miller *et al.*, 2001; Teng *et al.*, 2002). This phenotype was only observed once and disappeared quickly *in vitro*, suggesting that some specific conditions *in vivo* may trigger the phenotype.

The histidine kinase EF1051 gene is located downstream of the gene encoding the putative response regulator EF1050. Located upstream is an open reading frame encoding a possible 6-phosphogluconate dehydrogenase (EF1049) which is transcribed in the same orientation (Figure 5.21). Interestingly, upstream of listerial *lisR* there is also an open reading frame encoding a 6-phosphogluconate homolog, suggesting that *lisRK* and EF1051-EF1050 have similar gene organisation at their loci and that the dehydrogenase is linked in some way to the role of the two-component system. EF1051 of *E. faecalis* is classified as a group IIIA kinase based on the conservation of amino acids surrounding the phosphorylatable histidine residue (Fabret *et al.*, 1999), having two putative transmembrane helices and a long periplasmic loop (Figure 5.22). Like most histidine kinases, there is a large cytoplasmic domain in the C-terminus consisting of the catalytic domain (Figure 5.22). Surprisingly, there are no cysteine residues in EF1051 (Figure 5.22), suggesting that this protein is not capable of forming disulphide bridges.

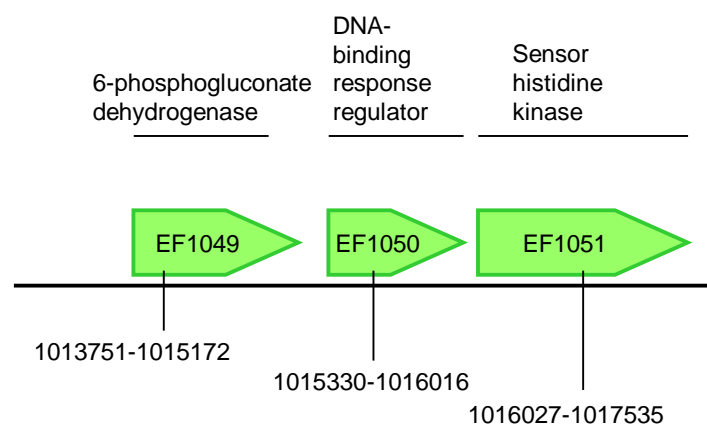


Figure 5.21 Map of the EF1051-EF1050 two-component regulatory system in *E. faecalis*. The green arrow represents a gene (▶), and the number below corresponds to the genome position of the gene. This diagram is drawn based on the genome map in KEGG database (Kanehisa, 2002).

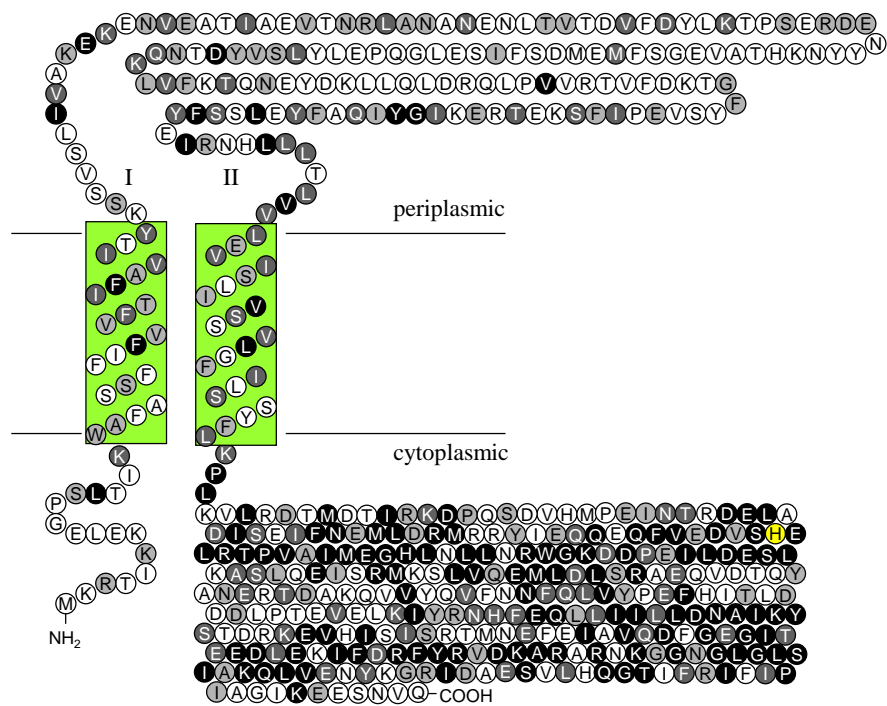


Figure 5.22 Topology model of *E. faecalis* histidine kinase EF1051. Topology prediction by TMHMM (Krogh *et al.*, 2001). The predicted helices are numbered from the N- to C-terminus and are shown in roman numerals. The phosphorylatable histidine is highlighted in yellow. Conserved residues are highlighted in grey and black according to the alignment in Appendix 10.

Membrane proteins are generally difficult to crystallise due to their hydrophobic nature. Detergents are needed for their extraction from the membrane, and once extracted, certain conditions are needed to keep the protein from forming aggregates. Increased contact points on proteins will enhance interaction and stacking of the protein in crystal formation. EF1051, with only two transmembrane helices, and a large periplasmic and cytoplasmic domain therefore may be considered to constitute a promising protein for crystallisation success.

This section describes the preparation of the EF1051 protein for crystallisation trials; including testing the activity, integrity, purity and monodispersity of the purified protein.

5.3.2. Cloning, expression and purification of EF1051

The EF1051 gene of *E. faecalis* V583 was successfully amplified and cloned into plasmid pTTQ18-His₆, followed by expression trials (Section 5.1.2). EF1051 was expressed to 16% of *E. coli* membrane proteins and migrates at 50 kDa in SDS-PAGE gels (Table 5.1 and Figure 5.1). To increase the expression of this protein, *E. coli*

BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆ were cultured and trials were performed to optimise expression conditions. Expression in different growth media and using different the IPTG concentrations were trialled and the results are summarised in Table 5.1. Growth of IPTG-induced cells transformed with pTTQ18(EF1051)-His₆ was most rapid in 2TY medium, then in LB medium and finally in the minimal M9 medium, attaining absorbances (at 680 nm) of 2.464, 1.735 and 1.677 respectively at the time of harvesting. Scanning densitometry of protein gels revealed that EF1051 constituted 13, 18 and 12% respectively, of total *E. coli* membrane proteins (Figure 5.23 and 5.24), and therefore that some minor improvement over the 16% obtained initially were achieved. Although the overall cell density attained was highest when cells were grown in 2TY medium, expression of EF1051 was shown to be highest in LB medium, and taking into account the cost effectiveness of the medium, LB was therefore used for subsequent optimisation tests.

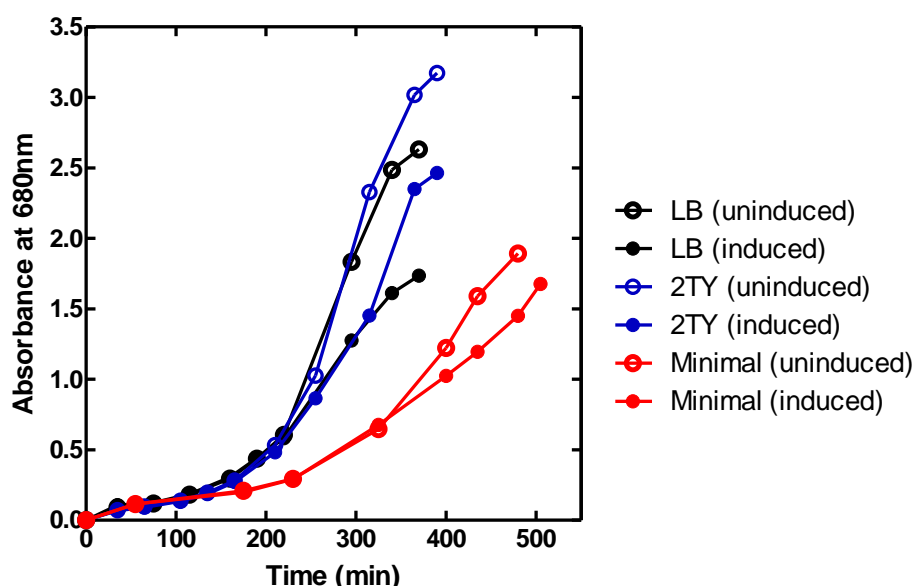


Figure 5.23 Growth of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆ in LB, 2TY and minimal media. *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆ were grown at 37°C in LB, 2TY or minimal media supplemented with 20 mM glycerol. When the absorbance (at 680 nm) of cultures reached 0.4-0.6, cells were induced with 0.5 mM IPTG and harvested 3 hours post-induction.

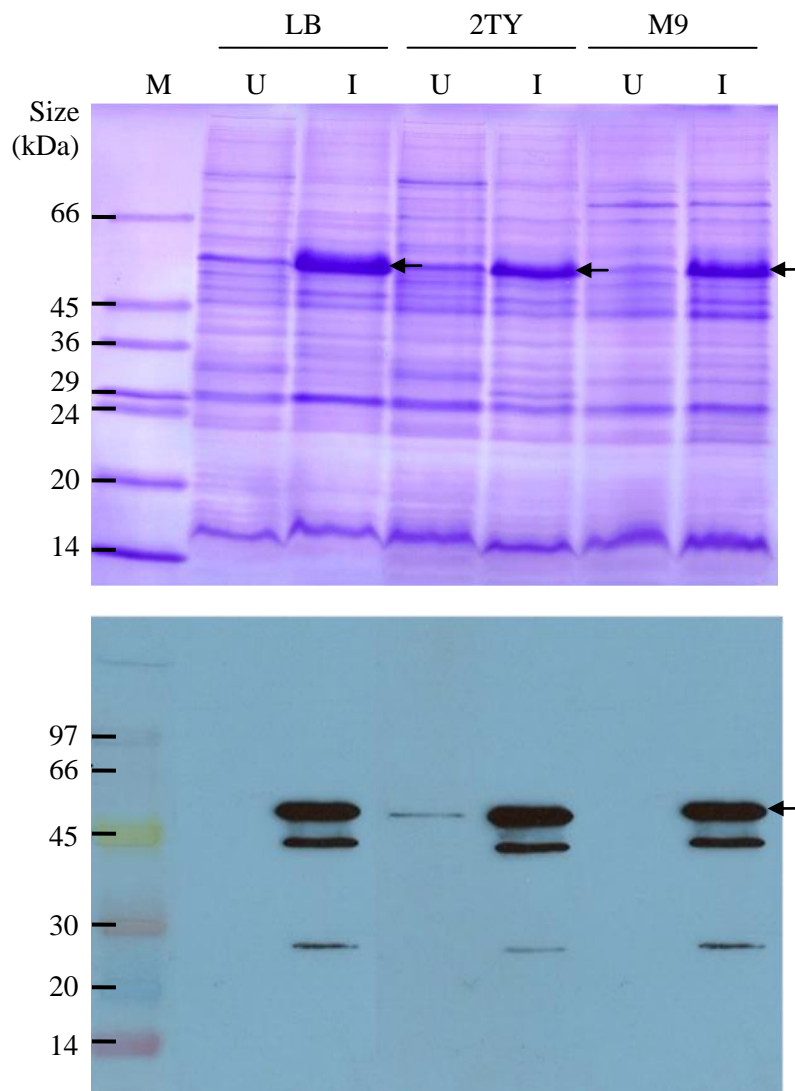


Figure 5.24 Expression of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆. BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆ were grown in LB, 2TY or minimal media supplemented with 20 mM glycerol. Cells were induced with 0.5 mM IPTG when $A_{680\text{nm}}$ reached 0.4 – 0.6 and harvested 3 hours post-induction. Total *E. coli* membranes were prepared and proteins analysed by SDS-PAGE. M (marker), U (uninduced) and I (induced). Arrow indicates position of EF1051.

Addition of 0.10 – 1.00 mM IPTG to the cultures attenuated cell growth of BL21(DE3) transformed with pTTQ18(EF1051)-His₆ (Figure 5.25). Scanning densitometry indicated that induction with 0.5 mM IPTG was optimal for expression of EF1051 under the test conditions (Figure 5.25). The cells were still in the exponential phase of growth at time of harvesting, therefore longer induction times should be tested to obtain more cells expressing the EF1051 protein.

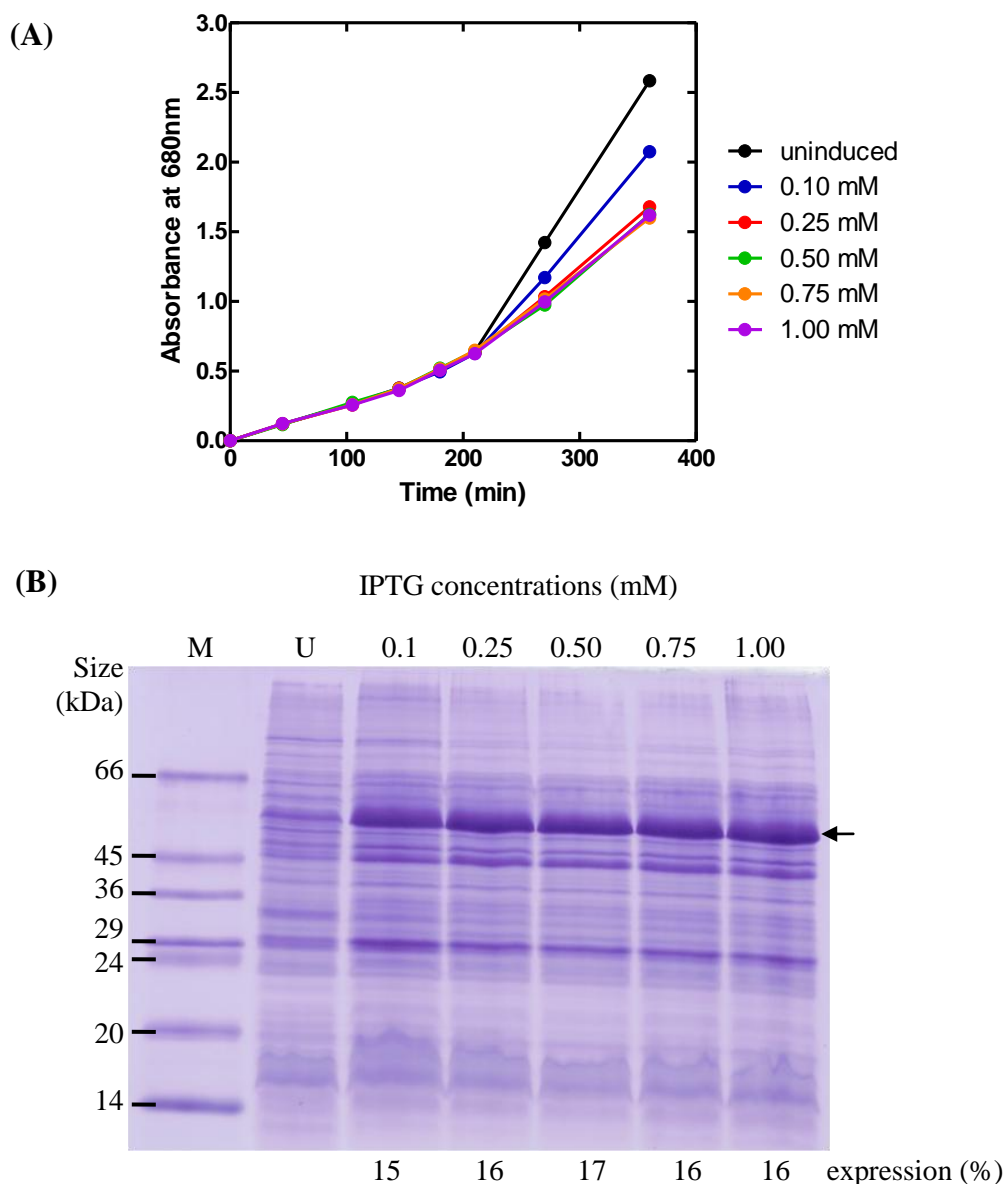


Figure 5.25 Growth and expression of *E. faecalis* EF1051 using a range of IPTG concentrations in LB medium. *E. coli* BL21(DE3) pTTQ18(EF1051)-His₆ cells were cultured at 37°C in LB in the presence of 100 µg/ml carbenicillin and 20 mM glycerol, and samples were induced with a range of IPTG concentrations (0 – 1 mM) when A_{680nm} reached 0.4 – 0.6, and harvested 3 hours post-induction. (A) Growth curves. Total membranes were prepared from the harvested cells and analysed by (B) SDS-PAGE. Arrow indicates position of EF1051.

A six litre culture of *E. coli* BL21(DE3) cells expressing His₆-tagged EF1051 was grown aerobically at 37°C (Section 2.4.1). Harvested cells were lysed by explosive decompression and membranes separated on sucrose density gradients (Section 2.5.1). The proteins of the membrane fractions obtained were analysed by SDS-PAGE. The results showed that the EF1051 protein was enriched in the inner membrane fraction

(Figure 5.26). EF1051 is present at approximately 18, 10 and <3% in the prepared inner, mixed and outer membranes, respectively (Figure 5.26).

The inner membranes expressing EF1051 were subjected to solubilisation trials (Section 2.5.2). The percentage of solubilised EF1051 was determined by calculating the intensities of EF1051 on the SDS-PAGE, in the inner membrane and in the soluble fractions, and representing the value for solubilised protein as the percentage of the total membrane proteins. All detergents tested were successful for solubilising 56 – 95% of EF1051, and n-octyl pentaoxy-ethylene, DDM and HECAMEG were the most effective (Figure 5.27).

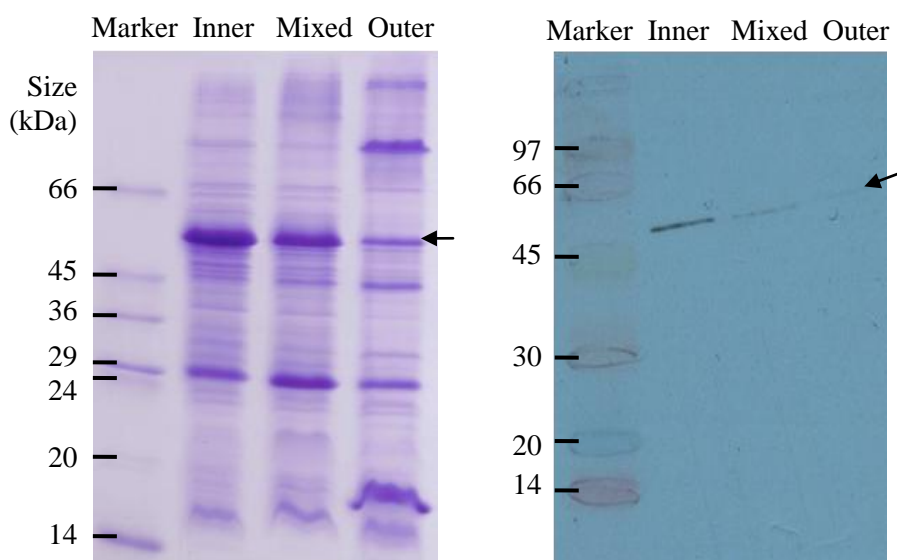


Figure 5.26 SDS-PAGE and Western blot of membrane fractions of *E. coli* BL21(DE3) cells expressing EF1051. Total *E. coli* membranes were prepared from *E. coli* BL21(DE3) pTTQ18(EF1051)-His₆ grown on LB medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. The membranes were separated on 25-55% sucrose gradients (Section 2.5.1) and membrane fractions extracted from the gradients were analysed by SDS-PAGE and Western blotting. The arrow indicates position of EF1051.

Inner membranes expressing EF1051 were therefore solubilised in 1% (w/v) DDM and purification of His₆-tagged EF1051 was performed by Ni-NTA affinity chromatography (Section 2.5.3.2). Fractions from the purification were analysed by SDS-PAGE, which revealed that the majority of the protein was successfully solubilised by DDM (solubilised fraction) and successfully bound to the Ni-NTA resin (Figure 5.28). EF1051 was purified to >90% purity, as determined by scanning densitometry, and was shown to migrate at a position corresponding to ~50 kDa on SDS-PAGE gels (Figure 5.28).

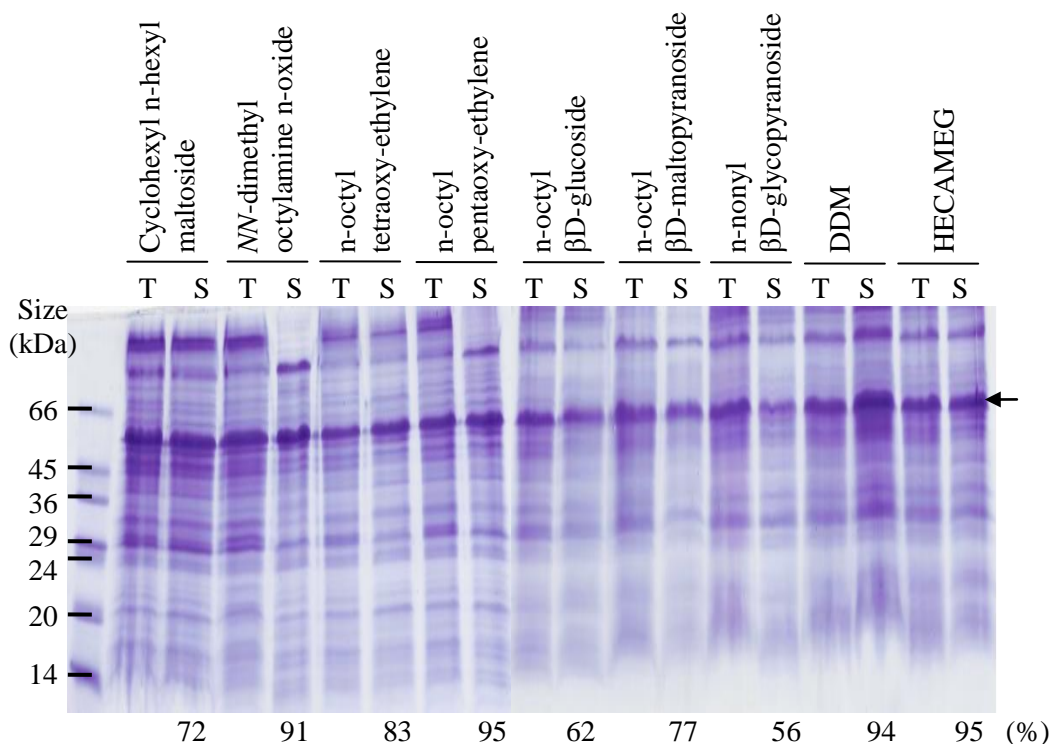


Figure 5.27 SDS-PAGE of total and solubilised inner membranes of *E. coli* BL21(DE3) cells expressing EF1051. Total *E. coli* inner membranes transformed with pTTQ18(EF1051)-His₆ were tested for solubilisation in nine detergents. Membranes were solubilised in 1% detergent for 1 hour and subjected to centrifugation to remove the insoluble fraction (Section 2.5.2). Total inner membranes (T) and soluble fractions (S) were then analysed by SDS-PAGE. Protein molecular weight markers (M). Arrow indicates position of EF1051.

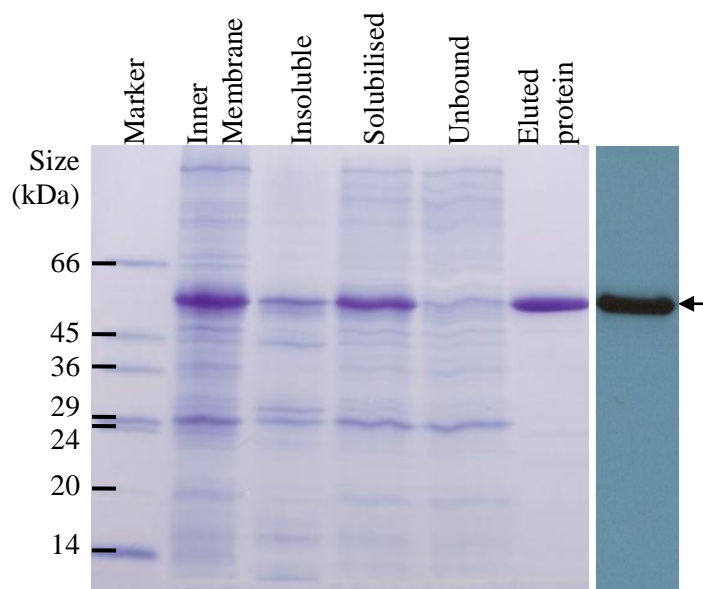


Figure 5.28 SDS-PAGE and Western blot of fractions from purification of *E. faecalis* EF1051. EF1051 was purified from inner membrane of *E. coli* BL21(DE3) cells transformed with pTTQ18(EF1051)-His₆. The inner membranes were solubilised in 1% DDM. Fractions (15 µg), and eluted fraction (5 µg) were analysed by SDS-PAGE. The panel on the right is a Western blot of the purified protein.

5.3.3. Size exclusion chromatography of EF1051

To check the monodispersity of purified EF1051, the eluted protein was buffer-exchanged into FPLC buffer, followed by size exclusion chromatography to separate the proteins on basis of size (Section 2.5.3.3). The protein was shown to elute in two separate elution volumes, indicated by the two symmetrical peaks in the chromatogram (Figure 5.29). The majority of proteins were eluted in the first volume of higher molecular weight (~617 kDa) material and this fraction could constitute aggregates of EF1051. The second volume of lower molecular weight species (~123 kDa) may comprise dimers of EF1051 (Figure 5.29).

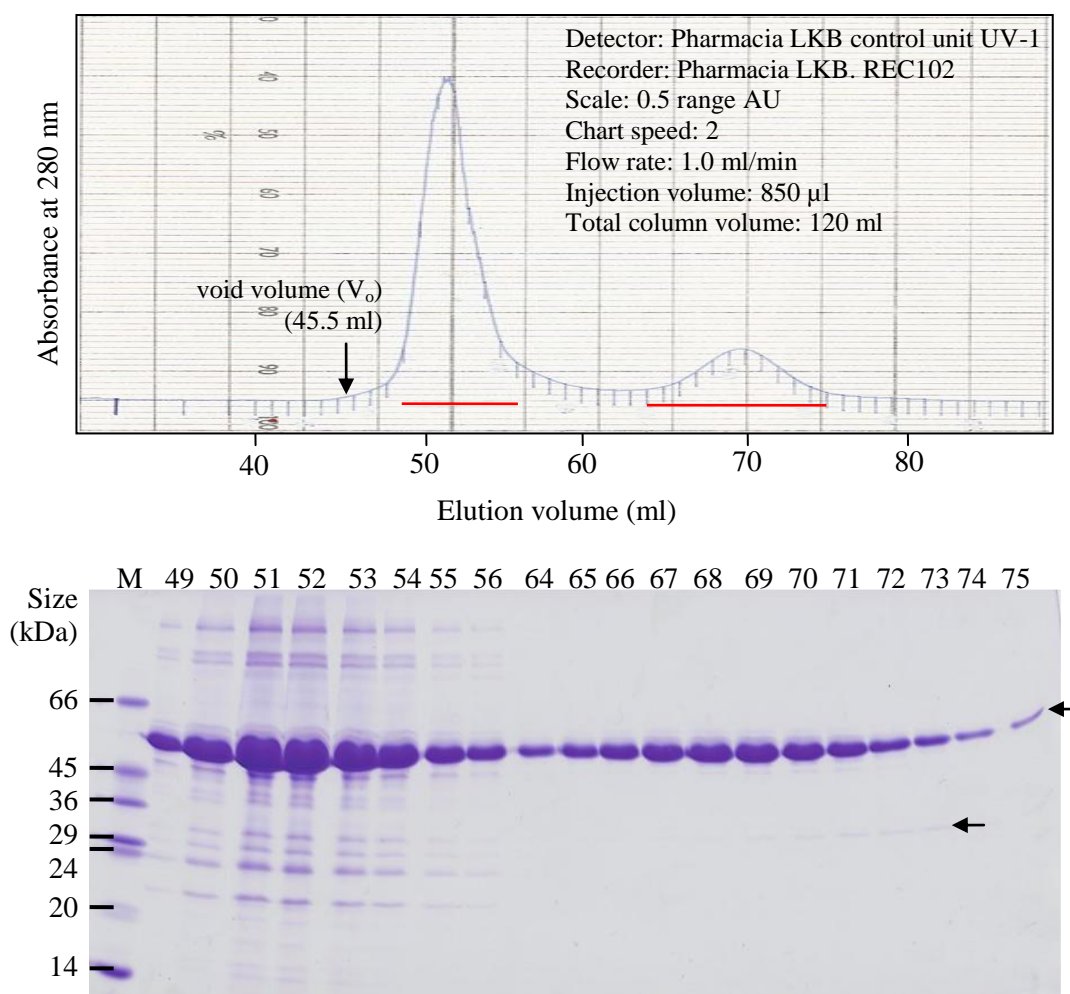


Figure 5.29 Size exclusion chromatography of purified EF1051. The His₆-tagged EF1051 protein purified by Ni-NTA was buffer exchanged into FPLC buffer (25 mM Tris-HCl pH8.0, 100 mM NaCl, 5% glycerol and 0.05% DDM) prior to gel filtration by size exclusion chromatography (Section 2.5.3.3). EF1051 (10 mg) was loaded onto the S-200 column equilibrated and eluted in FPLC buffer. Chromatogram of separation of purified EF1051 (top). Proteins and detected by UV absorbance (280 nm) were analysis by SDS-PAGE (bottom) – analysed fractions are indicated by the red line on chromatogram. Protein molecular weight marker (M). The arrow indicates position of EF1051 on the SDS-PAGE gel.

The molecular weights of the eluted volumes were determined by the calibration curve shown in Figure 5.30. Analysis by SDS-PAGE revealed that the EF1051 protein (~50 kDa) in the first peak contained many contaminants, and that the second peak showed highly pure EF1051 proteins (Figure 5.29). A very faint band (~32 kDa) on SDS-PAGE is also co-eluted with the pure protein and could indicate possible degradation of the protein (Figure 5.29). The fractions containing the pure proteins was concentrated and subjected to crystallisation trials. The results are shown in Section 5.3.6.

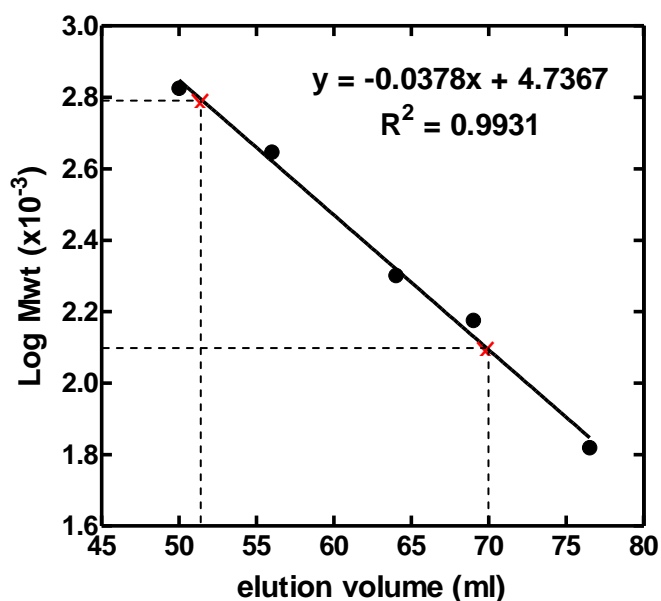


Figure 5.30 Calibration of the S200 column. The graph was constructed using protein standards (blue dextran, void volume; tryglobulin, 669 kDa; apofemitin, 443 kDa; β -amilase, 200 kDa; alcohol dehydrogenase, 150 kDa ; albumin, 76.5 kDa). Red crosses indicate the positions of the elution volumes of monodispersed EF1051 proteins.

5.3.4. Autophosphorylation activities of EF1051

Different kinetics were observed for the autophosphorylation of EF1051 bound in inner membranes and its purified form (Section 5.1). In the inner membrane, EF1051 exhibited rapid autophosphorylation initially, followed by rapid desphosphorylation after 1 minute, and a steady level being reached after 15 minutes and to the end of the reaction (Figure 5.2). In the purified form, EF1051 showed slower rates of phosphorylation and maximum phosphorylation was not reached within the sixty minute time-course (Figure 5.4). The rapid dephosphorylation of EF1051 observed when the protein is membrane-bound may occur as a result of the activities of *E. coli*

phosphatases present, or it could arise as an intrinsic property of the protein in this environment.

Prior to obtaining active EF1051 as described in the previous section and as shown in Figure 5.2 and 5.4, initial attempts to produce active protein were unsuccessful. However, when DDM was excluded from all post-solubilisation buffers during protein purification steps, then active protein resulted. The exclusion of detergent at key stages in the protein purification process was routinely adopted for the histidine kinases following initial observations that this was a requirement for purification of active VicK (Phillips-Jones, personal communication). Therefore, in the subsequent purifications of EF1051, DDM was not added to the protein post-solubilisation, and two buffers were tested, including Tris and HEPES. The result gave rise to active proteins, though the protein in Tris buffer appears to be less active and partially degraded (Figure 5.31).

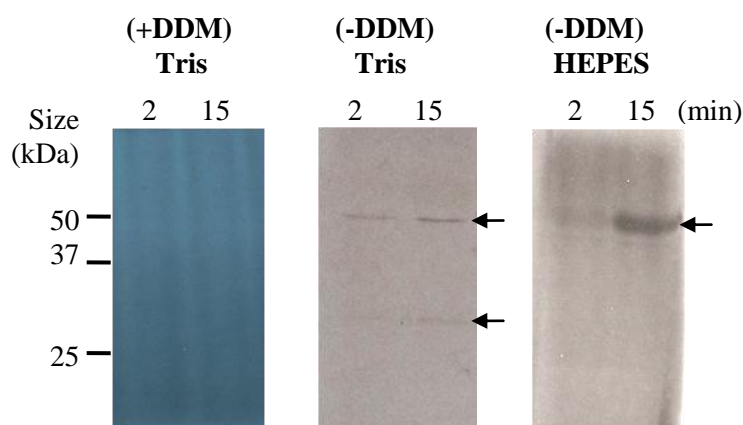


Figure 5.31 Effect of buffer composition and detergent on autophosphorylation activities of purified EF1051. EF1051 purified with DDM (+DDM) or without DDM (-DDM) addition post-solubilisation was buffer-exchanged into storage buffer (Section 2.5.5.2). The protein (160 pmoles) was then added to reaction buffer containing 10 mM Tris-HCl or HEPES-Na pH7.6 in a final reaction volume of 30 μ l and autophosphorylation initiated using radiolabelled ATP (Section 2.5.5.2). Samples volumes (15 μ l) were removed at the time points indicated. The autoradiographic film is shown. The arrows indicate position of the phosphorylated EF1051 protein.

5.3.5. Mass spectrometry and circular dichroism spectroscopy studies of purified EF1051

Mass spectrometry (Section 2.6.2) was performed to confirm that the purified EF1051 protein was intact. The results are shown in Figure 5.32, and these indicate that the measured mass of the protein was 60189 Da. This represents a mass difference of 32 Da compared to the calculated mass of 60157 Da (Figure 5.32).

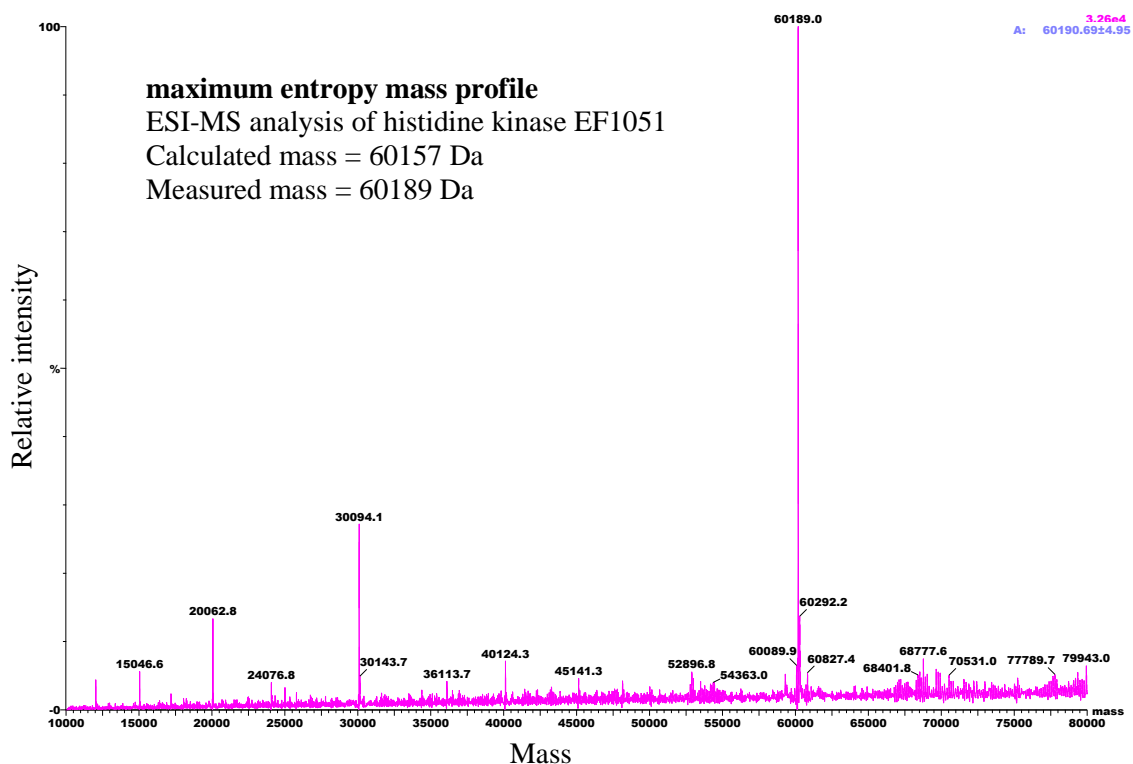
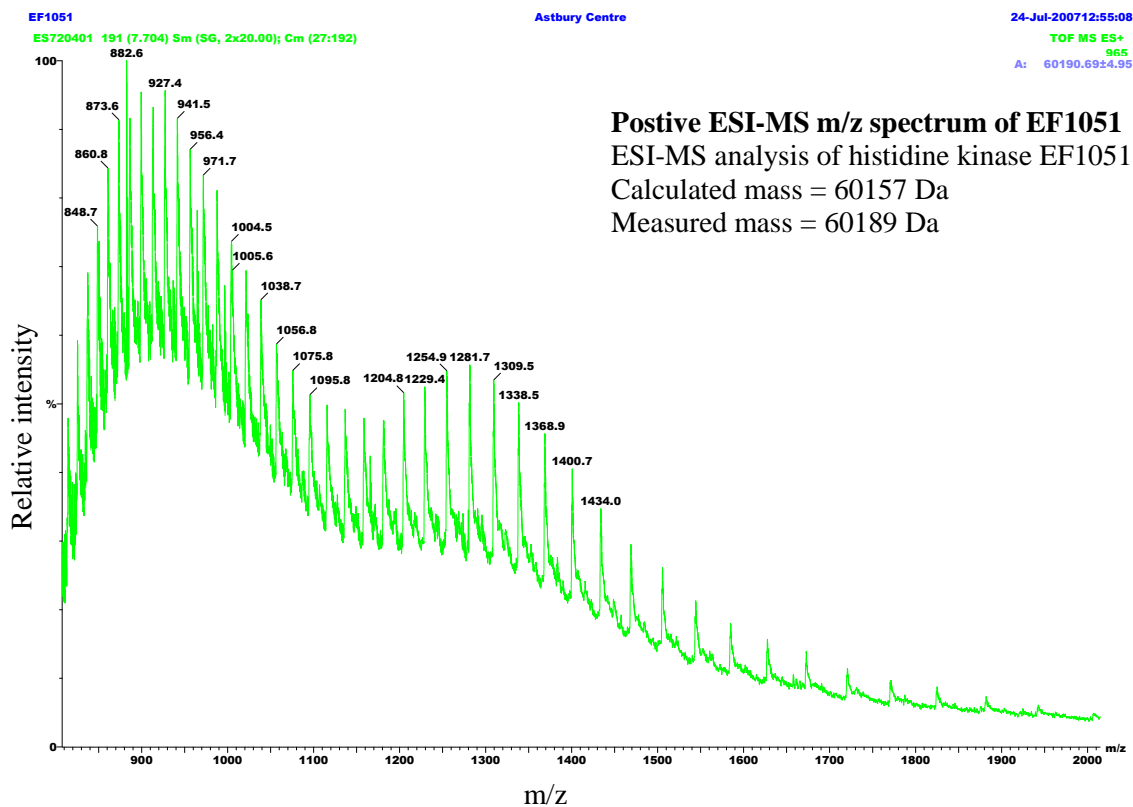


Figure 5.32 Electropray ionisation mass spectrometry of EF1051. Purified protein (5 mg/ml) in 10 mM HEPES pH 7.9 and 5% glycerol was prepared as described in Section 2.6.2. Mass spectrometry analysis was performed by Miss Lynsey Jones.

This indicates possible N-terminal formylation of the protein, since formic acid was added to dissolve the protein prior to size exclusion chromatography for mass spectrometry (Section 2.6.2). Thus, it was concluded that the protein is intact. The other three peaks in the maximum entropy mass profile, of sizes 30094.1, 20062.8 and 15046.6, which are 1/2, 1/3 and 1/4 of the main peak are artefacts from the maximum entropy algorithm (Figure 5.32).

The secondary structure integrity of purified EF1051 was examined by circular dichroism spectroscopy (Section 2.6.1). The negative indentation at 208 and 222 nm, and the positive signal at ~190 nm obtained in this study are indicative of α -helical content (Figure 5.33), suggesting that the protein is correctly folded.

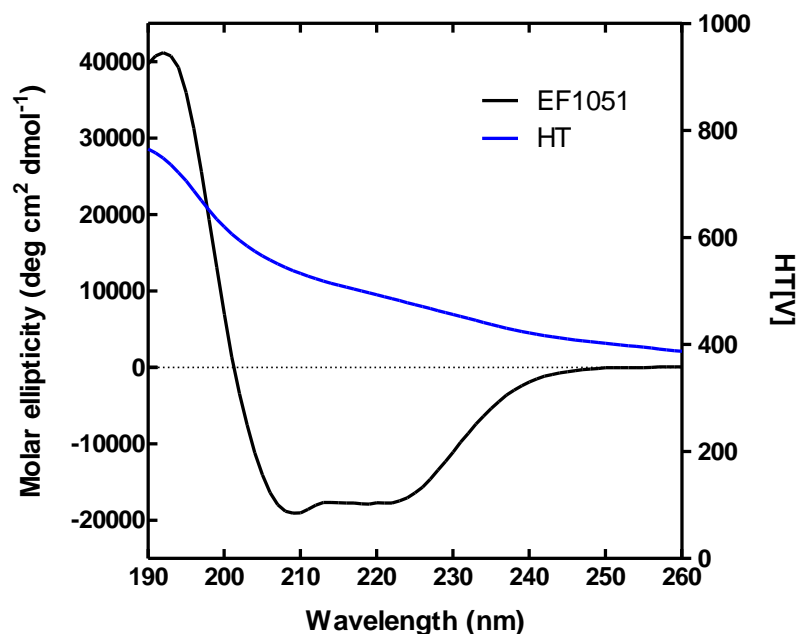


Figure 5.33 Circular dichroism spectrum of purified EF1051. Purified EF1051 (0.05mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of EF1051 was performed using a Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample was analysed in a Hellma quartz-glass cell of 1 mm path length. The spectrum was recorded with 1 nm step resolution at a scan rate of 10 nm/min. Response time was set at 1 second with a sensitivity of 20 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of ten scans, from which the buffer contribution was subtracted. The blue line represents the voltage applied to the photomultiplier.

5.3.6. Crystallisation trials of EF1051

The pure homogeneous EF1051 protein prepared as described above (Section 5.3.3) was subjected to crystallisation trials by the vapour diffusion method (Section

2.6.2). Microscopic images of crystals obtained from the crystallisation trials were taken and are shown in Figure 5.34.

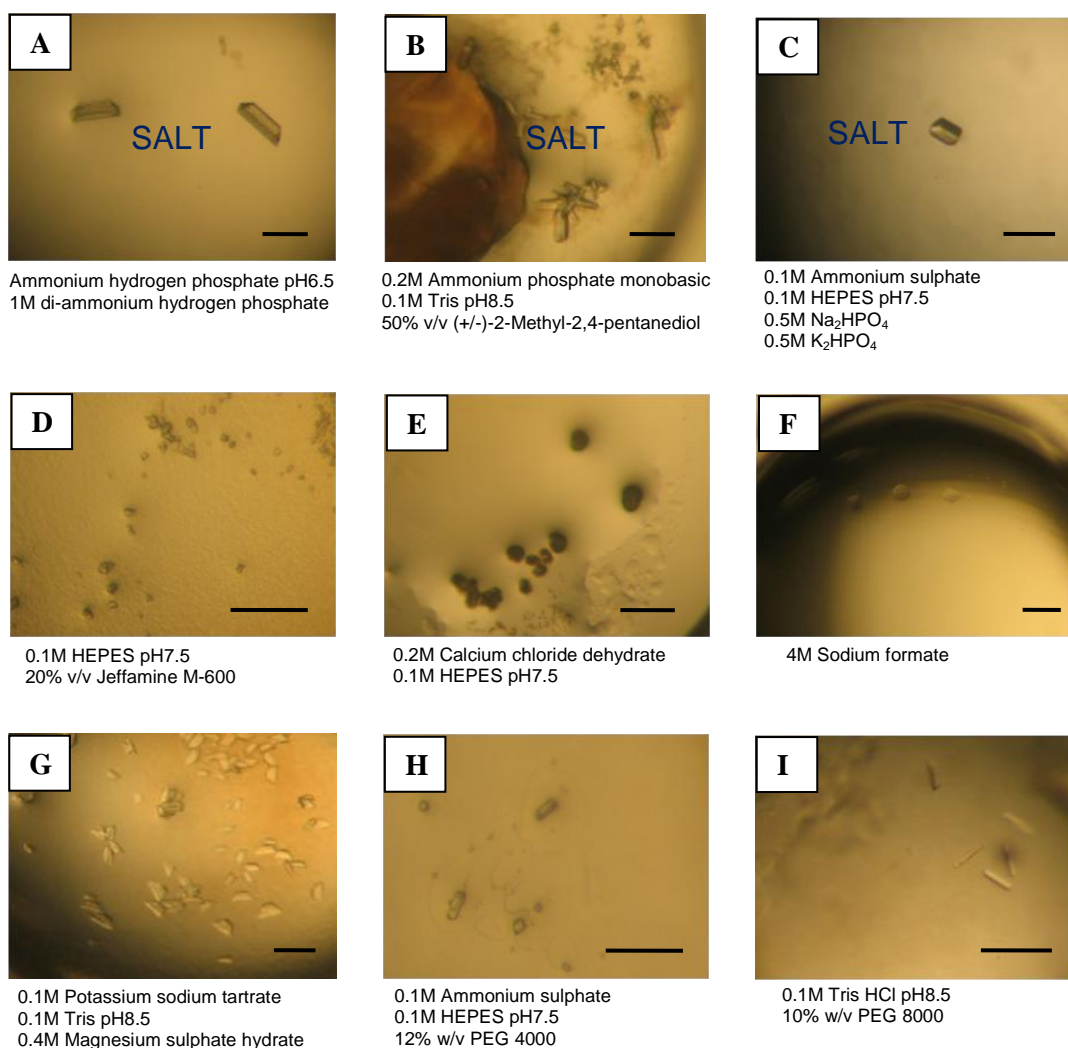


Figure 5.34 Microscopic images of crystals from the crystallisation trials of EF1051. Purified homogeneous EF1051 proteins (~15mg/ml) were subjected to crystallisation trials by vapour diffusion (Section 2.6.6). Each droplet (1 μ l) consists of 0.5 μ l protein and 0.5 μ l screen buffer in each well of the 96 wells plate was monitored and visible images of conditions containing crystals were photographed. Screen buffer conditions are shown below the image. Bar represents 150 μ m.

Crystals of different shape and size were formed in the crystallisation trials (Figure 5.34). There were large crystals (Images A, B and C), micro-crystals (Image D), thin needle clusters or sea urchins (Image E), 2D crystals (Image F), rhombus-shaped crystals (Image G), irregular crystals, (Image H) and rod crystals (Image I) (Figure 5.34). One of these conditions also gave rise to heavily precipitated proteins which is likely to be denatured proteins (Image B), and in some cases there was evidence of phase separation (Image E and H) (Figure 5.34). Precipitation occurs in a state of

supersaturation and crystals can only grow from supersaturated states. Phase separation (the tendency for detergent to separate from the aqueous solution) can inhibit crystal formation, due to high detergent concentrations (Newby *et al.*, 2009).

To determine whether the crystals grown obtained in these trials were composed of protein or salt, the larger sized crystals (Images A, B and C of Figure 5.34) were removed for further analysis. They were firstly placed in the mother liquor (the buffer that the crystal was grown in) containing 10% glycerol, and flash frozen in liquid nitrogen. The crystals were then subjected to x-rays at the Radiation Synchrotron (with the assistance from Dr Gerda Szakonyi). The diffraction patterns obtained revealed that these crystals were composed of salts (Figure 5.34). The remaining crystals (Images D – I of Figure 5.34) are yet to be tested.

5.3.7. Discussion: structural studies of EF1051

The EF1051 protein of *E. faecalis* is highly similar to the LisK, CsrS and ArlS of other bacteria (Section 5.3.1), and is involved in stress response and virulence of *E. faecalis* (Teng *et al.*, 2002; Muller *et al.*, 2008). This work optimised the expression conditions for EF1051 in *E. coli* membranes and describes a successful purification method for the intact protein, and prepared the protein for subsequent crystallisation trials. Purity, homogeneity and stability of membrane proteins, as well as individual properties, such as surface charge and isoelectric point of the protein, are the most important parameters for crystallisation. It is therefore important to optimise the quality of the purified protein prior to crystallisation studies, and this often includes testing the purity of the protein, checking the CD spectrum of the protein to confirm that it has secondary structure, ensuring the mixture exists in solution as a single oligomeric species (monodispersity), e.g. monomer or dimer, and therefore ensuring that the protein is not degrading with time into a heterogeneous mixture.

The EF1051 protein was firstly subjected to an *in silico* analysis to identify its putative function by a similarity search using FASTA3 (Pearson, 2000) for homologues with known functions (Section 5.3.1). In agreement with previous reports, analysis here reveals the high similarity of the EF1051 protein with LisK of *Listeria monocytogenes* and with CsrS of *Streptococcus agalactiae* (Cotter *et al.*, 1999; Lamy *et al.*, 2004).

Primary sequence analysis of EF1051 revealed an absence of cysteine residues, and prediction by TMHMM showed that it possess two predicted transmembrane segments and a large soluble periplasmic and cytoplasmic domain (Figure 5.20 and

5.22). In this work, the EF1051 protein was overexpressed successfully in *E. coli* membranes and it was shown to retain activity in both inner membrane preparations as well as in its purified form (Section 5.1). The expression of EF1051 was optimised using different media and a range of IPTG concentrations (Figure 5.23 to 5.25), and these optimised conditions were subsequently used for routine production of purified EF1051.

In this study, choice of detergents for solubilisation of EF1051 from *E. coli* membranes was also examined. All detergents screened showed successful solubilisation of EF1051 from inner membranes, suggesting that this protein is easily solubilised under these conditions. Subsequent purification steps yielded protein purities of >90% (Figure 5.28). Mass spectrometry and CD spectral data confirmed that purified EF1051 was intact and that its secondary structure was correctly folded (Figure 5.32 and 5.33).

The size exclusion chromatography of EF1051 produced two groups of monodispersed proteins, one of higher molecular weight which could be composed of aggregated proteins and the other of lower molecular weight that could be dimers of EF1051 (Figure 5.29). Results indicated that most of the EF1051 protein was forming higher molecular weight complexes and that a small volume of pure homogeneous proteins was achieved. The impure fraction could also be subjected to crystallisation trials, and some buffer optimisation might prove beneficial for increased yields of pure homogenous protein. The pure EF1051 homogeneous fractions were concentrated, and subjected to crystallisation trials (Section 5.3.6), which produced crystals of different shapes and sizes (Figure 5.34). Three of the nine conditions tested produced crystals that were later confirmed as salt crystals, and the remaining are yet to be tested. Unfortunately, at the time these trials were set up (Feb 2007) there was no UV-fluorescence microscope in the lab to test whether these crystals were protein or salt, and therefore the trial needs to be repeated in future crystallisation studies of this protein. The UV-fluorescence microscopic technique is based on the UV-fluorescence of the amino acid tryptophan, which fluoresces at 360 nm when excited with UV-light at 280 nm, therefore easily distinguishing the protein crystal from salt (Judge *et al.*, 2005).

It was previously shown in an analysis of 9596 unique protein crystals, that there is a relationship between calculated pI and reported pH at which proteins were crystallised. The analysis showed that proteins with calculated pIs of between 5 and 7

turn out to be the group exhibiting the highest frequency of crystallisation successes, and that most of the proteins crystallised at pHs between pH 7 and 8 (Kantardjieff & Rupp, 2004). EF1051 has a calculated pI of 5.22 that lies in the range indicated, and could possibly indicate that this membrane protein stands a good chance of crystallisation. However, it needs to be further tested.

Monodispersed protein particles are the building blocks of stable crystal lattices. Membrane proteins that are hydrophobic in their nature tend to be subjected to non-specific aggregation over time. The irregular association of the protein particles in solution reduces the possibility of the regular association essential for crystal lattice formation. In this case, the choice of detergent is one of the key factors to reduce formation of non-specific aggregates. Large micelle-forming detergents such as DDM and C₁₂E₉ are more likely to stabilise membrane proteins in solution, but often prevent the protein-protein interactions necessary for crystal lattice formation due to the large size of the micelle (Keyes *et al.*, 2003). Smaller micelle-forming detergents, such as n-octyl- β -D-glucoside and n-octyl- β -D-maltoside, allow more protein-protein interactions but could lead to undesirable aggregation of proteins (Keyes *et al.*, 2003). The DDM which was shown to be successful for solubilisation of most of the histidine kinases forms a large detergent micelle around the protein; but since it has a low critical micelle concentration (CMC), it enables buffer exchange of the protein into other detergents with smaller micelle-size that are more suitable for subsequent crystallisation trials.

Not all types of detergents are suitable for solubilising membrane proteins, as they can inactivate the protein (Le Maire *et al.*, 2000). This study revealed that the presence of detergents in the post-solubilisation stages of protein purification did appear to be inhibitory to the activities of both EF1051 and EF2912, both of which possess just two transmembrane helices. Other histidine kinases appear to be more resistant to the effects of detergent, including those with more transmembrane segments such as EF3197. Transphosphorylation between monomers is required for activity of histidine kinases and therefore active kinases must occur as dimers both *in vivo* and *in vitro*. Thus it is possible that the inhibitory effects seen here and with other histidine kinases are attributable to an inhibition of the essential dimerisation events.

Solving the structure of membrane proteins has always been hampered by their low expression level and therefore producing sufficient pure proteins for structural study is one of the main barriers. This work has overcome one of the bottlenecks in structural study of membrane protein; highly pure milligram quantity of

monodispersed EF1051 protein were obtained (Figure 5.28 and 5.29), which is sufficient for protein crystallisation studies. The next barrier for determining the structure of EF1051 is crystallisation. However, there are many more tests that could be performed to further test the stability of this protein, such as described in Postis *et al.* (2008). After all, this study opened up more paths for further investigations, and there are more components, such as buffer, detergents and temperature, that could be tested to further improve the stability and homogeneity of this protein.

5.4. Characterisation of signal interactions by the quorum sensor EF1820 (FsrC)

5.4.1. Introduction

5.4.1.1. Quorum sensing in bacteria. Quorum sensing was discovered and described over thirty years ago in two bioluminescent marine bacterial species, *Vibrio fischeri* and *Vibrio harveyi*, one of which colonises the light organ of the Hawaiian Bobtail Squid *Euprymna scolopes* and the other is a free-living marine bacterium (Nealson & Hastings, 1979). Loss of luminescence was observed in these bacteria growing at low cell densities, and the enzymes responsible for light production; in both species, were found to be encoded by the luciferase structural operon *luxCDABE* (Engebrecht & Silverman, 1984; Miyamoto *et al.*, 1988). The exchange of chemical signals between biological cells is a fundamental requirement for synchronisation of the activities of large groups of cells and was previously assumed to be a trait characteristic of eukaryotes. It was only in the past two decades that cell-to-cell communication has been identified as of importance to bacterial cells and has become known as ‘quorum sensing’. Quorum sensing involves the production, release and detection of, and response to small hormone-like molecules termed autoinducers. It is a process that allows the bacteria to monitor their environment for cell density and to alter gene expression within populations in response to changes in the number and/or species present in a community (Waters & Bassler, 2005). Bacterial quorum sensing controls a wide variety of physiological activities such as symbiosis, virulence, competence, conjugation, antibiotic production, motility, sporulation, and biofilm formation (Magnuson *et al.*, 1994; Ji *et al.*, 1995; Pestova *et al.*, 1996; Swift *et al.*, 1996; Gray, 1997; Hardman *et al.*, 1998; Hellingwerf *et al.*, 1998; Miller & Bassler, 2001; Nakayama *et al.*, 2001; Waters & Bassler, 2005).

5.4.1.2. Comparison of quorum sensing in Gram-negative and Gram-positive bacteria. There are several notable similarities and differences in the quorum sensing systems of Gram-negative and Gram-positive bacteria. The main similarity of course is the overall function of these systems amongst Gram-positive and Gram-negative bacteria regarding cell-cell communication, which is presumably a fundamental requirement of both bacterial groups. The main difference is the nature of the signalling molecule. Generally, Gram-negative bacteria use fatty acids derivatives, ‘acylated homoserine lactones’ as autoinducers and many possess LuxIR-type proteins (Fuqua & Greenberg, 1998; Eberl, 1999; Parsek *et al.*, 1999; Fuqua *et al.*, 2001; Whitehead *et al.*, 2001; Manefield & Turner, 2002), whereas Gram-positive bacteria use oligopeptides termed ‘peptide pheromones’ together with a two-component system to mediate quorum sensing (Nes *et al.*, 1996; Dunny & Leonard, 1997; Kleerebezem *et al.*, 1997; Podbielski & Kreikemeyer, 2004; Shapakov, 2009).

One of the first and best-studied quorum sensing systems in a Gram-negative bacterium includes the LuxIR circuit of *V. fischeri* (Figure 5.35). There are five luciferase structural genes (*luxCDABE*) and two regulatory genes (*luxR* and *luxI*) involved in the control of light emission, and these genes are arranged in two adjacent but divergently transcribed units. The LuxI protein synthesises the autoinducer, *N*-(3-oxohexanoyl)-homoserine lactone, which increases in concentration both intracellularly and extracellularly when the cell-population density increases. This autoinducer is able to diffuse freely in and out of the cell, and when a critical concentration is reached, it binds to the LuxR protein forming the LuxR-autoinducer complex. This complex acts on two promoter sites. Binding to the *luxICDABE* promoter results in an exponential increase in autoinducer synthesis and light production, while binding to the *luxR* promoter represses the transcription of *luxR* which compensates for, and is in balance with the positive regulation of the *luxICDABE* promoter.

In contrast to the widespread use of LuxR-type proteins as autoinducers in Gram-negative bacteria, Gram-positive bacteria employ peptide signals together with two-component signal transduction systems for detection of the peptide autoinducer (Hoch & Silhavy, 1995; Kleerebezem *et al.*, 1997; Lazazzera & Grossman, 1998; Bassler, 1999; Podbielski & Kreikemeyer, 2004; Ng & Bassler, 2009). For example, one of the best studied examples of a quorum sensing system in Gram-positive bacteria is the *agr* (accessory gene regulator) system in staphylococci (Dunman *et al.*, 2001; Kornblum *et al.*, 2003).

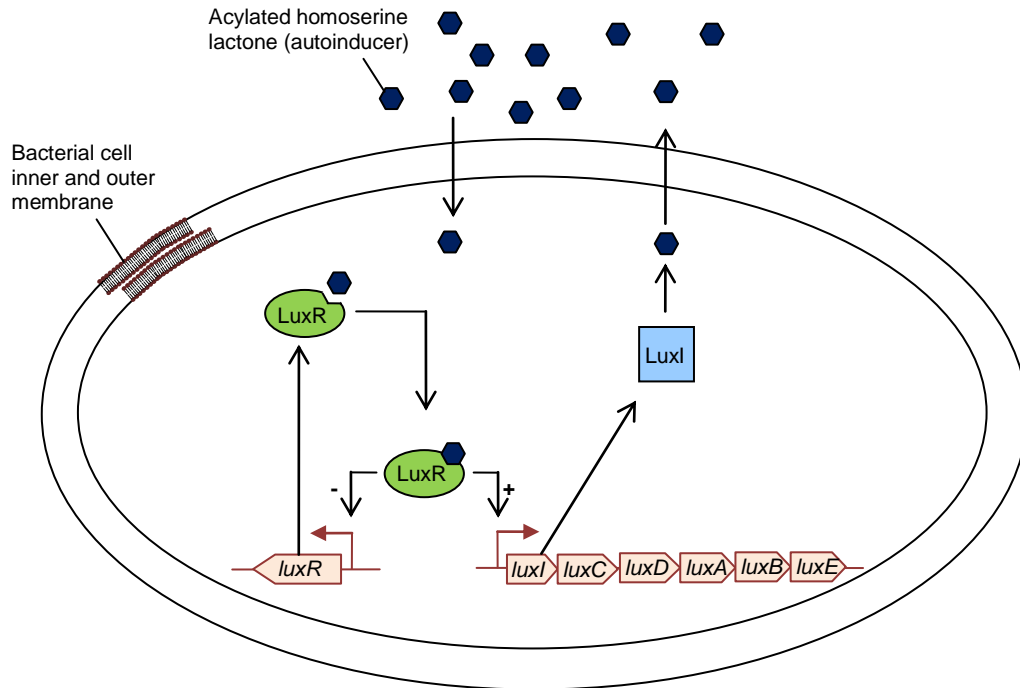


Figure 5.35 The LuxIR quorum sensing circuit in Gram-negative bacteria. Schematic diagram of a general model for ‘acylated homoserine lactone’-mediated quorum sensing. This LuxIR-type quorum sensing which is utilised by Gram-negative proteobacteria. The regulon is not drawn to scale. Diagram adapted from Miller & Bassler (2001).

In pathogenic staphylococci, the *agr* system (Novick, 2003; Lyon & Novick, 2004) is reported to have a critical role in biofilm formation (Vuong *et al.*, 2000; Shompole, *et al.*, 2003; Vuong *et al.*, 2003). The *agr* gene clusters are conserved throughout the genus *Staphylococcus* and they are divided into more than four types or specificity groups according to the mutual inhibition by the peptides of the *agr* response in heterologous pairs (Ji *et al.*, 1997; Jarraud *et al.*, 2000; Geisinger *et al.*, 2008). Interestingly, the staphylococcal species have hypervariable regions in the *agr* locus and this probably gives rise to the different peptides sensed by these species (Dufour *et al.*, 2002). The advance in genome sequencing projects led to the identification of *agr*-like systems in other pathogenic bacteria (Wuster & Babu, 2008), including the *agr* system of *Listeria monocytogenes* (Autret *et al.*, 2003) and the *fsr* system in *Enterococcus faecalis* (Hancock & Perego, 2004a; Nakayama *et al.*, 2001a; Nakayama *et al.*, 2006), in which it was shown that these systems are important for virulence in these species. In addition, *agr*-like systems are also present in non-pathogenic commensal organisms, such as the *com* system in *Bacillus subtilis* (Comella & Grossman, 2005), and the *lam* system in *Lactobacillus plantarum* (Kleerebezem *et al.*, 2003; Sturme *et al.*, 2005; Fujii

et al., 2008). Recently, the *lam* system of *L. plantarum*, which naturally inhabits the human gastrointestinal tract (Ahrne *et al.*, 1998), was shown to be involved in the regulation of adherence, cell morphology, cell viability, and biofilm formation (Storme *et al.*, 2005; Fujii *et al.*, 2008). It appears that *agr* systems in bacteria are important for cell survival, biofilm formation and hence in virulence for many adherent pathogenic species.

5.4.1.3. The Fsr quorum sensing pathway of *E. faecalis*. The *agr*-like operon in *E. faecalis* is named *fsr* and comprises four genes (*fsrA*, *fsrB*, *fsrD*, and *fsrC*) together with the two virulence genes (*gelE* and *sprE*) (Figure 5.36) that are involved in biofilm formation (Hancock & Perego, 2004a; Pillai *et al.*, 2004). Mutants in *fsr* were shown to be less virulent in *Caenorhabditis elegans* killing, mouse peritonitis and rabbit endophthalmitis models (Qin *et al.*, 2000; Mylonakis *et al.*, 2002; Sifri *et al.*, 2002), indicating the importance of the *fsr* system in enterococcal virulence.

In common with the *agr* system of staphylococci, the *fsr* two-component system is involved in quorum sensing. With increasing cell population density in *E. faecalis*, a peptide pheromone named ‘gelatinase biosynthesis-activating pheromone’ (GBAP) consisting of eleven amino acids (Figure 5.43), is secreted from the cell by a putative dedicated GBAP transporter (FsrB). The initially synthesised GBAP propeptide encoded by *fsrD* is inactive and presumably modifications such as processing and cyclisation are associated with secretion by FsrB that was found to function as a cysteine protease-like processing enzyme, to produce the active cyclic-GBAP peptide (Nakayama *et al.*, 2001a; Nakayama *et al.*, 2006). A model depicting the mechanism of quorum sensing in enterococci is shown in Figure 5.36. The histidine kinase FsrC detects the secreted GBAP peptide and undergoes autophosphorylation of a conserved histidine, followed by phosphotransfer to a conserved aspartate on its cognate response regulator FsrA. Phosphorylated FsrA binds to the *gelE-sprE* and *fsrBCD* promoters and activates transcription of these genes (Su *et al.*, 1991; Qin *et al.*, 2000; Qin *et al.*, 2001; Nakayama *et al.*, 2001a; Nakayama *et al.*, 2001b; Bourgogne *et al.*, 2006). The histidine kinase FsrC (previously referred to as EF1820 in Section 5.1) was therefore further studied in this project to initiate studies of the molecular details of FsrC-GBAP binding and to determine whether the *in vitro* approach developed by Potter *et al.* (2002) and Ma *et al.* (2008) is suitable for identifying signals of intact histidine kinases.

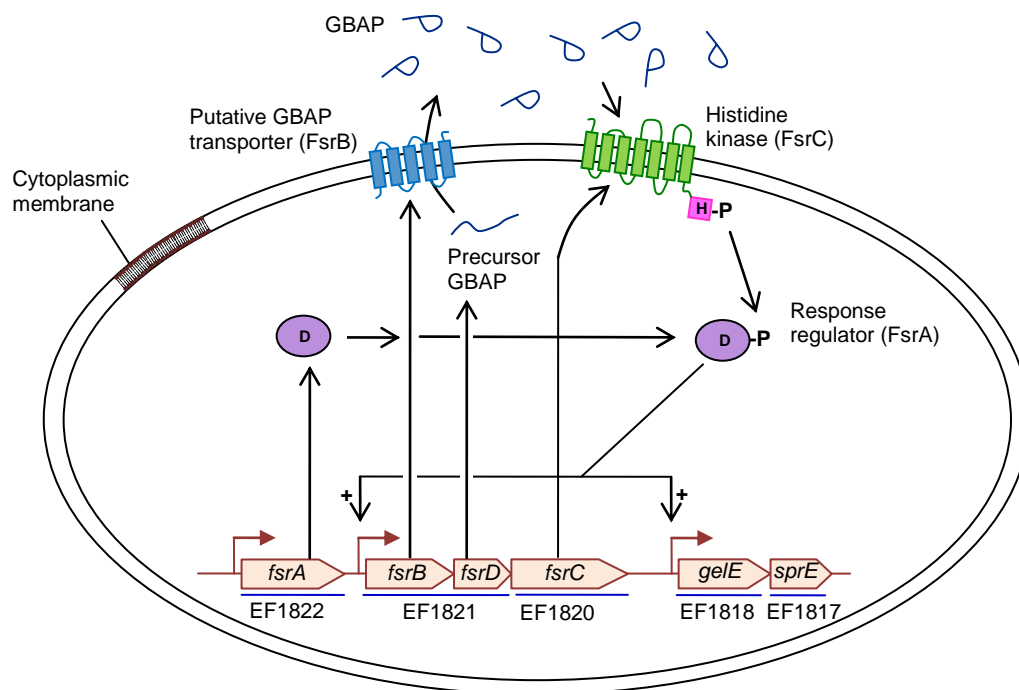


Figure 5.36 Quorum sensing regulation in enterococci. Schematic diagram of a general model for peptide-mediated quorum sensing in enterococci is shown. The regulon is not drawn to scale. Diagram adapted from Podbielski & Kreikemeyer (2004), and Nakayama *et al.* (2006).

5.4.2. The EF1820 protein (FsrC) of *E. faecalis*

The FsrCA (EF1820-EF1822) proteins of *E. faecalis* belong to the Agr family of two-component regulatory systems according to the KEGG database (Kanehisa *et al.*, 2010). The Agr family are characterised by possession of five to eight transmembrane segments in their N-terminal domain, connected to a C-terminal domain containing catalytic features of histidine kinases. FsrC is grouped with the Agr kinases, and is predicted to possess seven putative transmembrane segments and a large soluble cytoplasmic catalytic domain (Figure 5.37). The amino acid sequences of FsrC proteins are highly conserved amongst enterococcal species (25.6 – 100% identity). *Bacillus* species exhibit less similarity with FsrC, for example Agr-like kinase BCE3043 exhibited 25.8% identity and 68.2% similarity to FsrC, whilst AgrC of *S. aureus* exhibited 24.5% identity and 61.8% similarity. An alignment of FsrC with some other Agr kinases (Appendix 11) reveal conserved leucine, isoleucine and valine residues in the N-terminal transmembrane domains, which could be involved in autoinducer binding, but otherwise the N-terminal/transmembrane regions are variable and this is perhaps consistent with the different autoinducers bound by these bacterial sensory proteins. The cytoplasmic domain was also shown to be highly conserved around the phosphorylatable histidine region (Appendix 11).

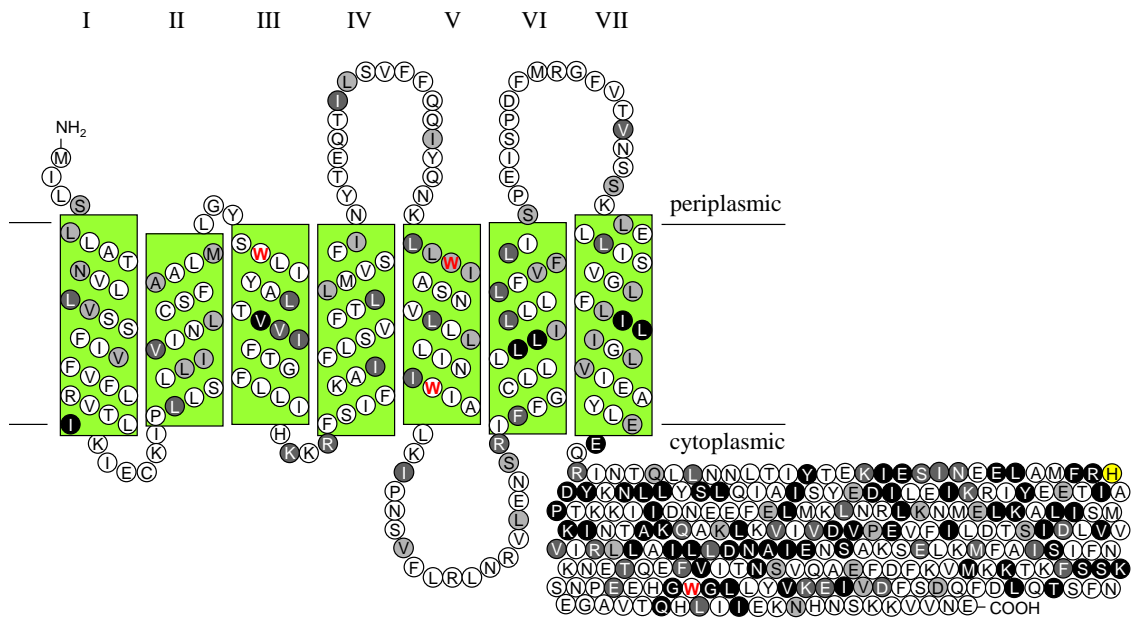


Figure 5.37 Topology model of *E. faecalis* histidine kinase FsrC. Topology prediction by TMHMM (Krogh *et al.*, 2001). The predicted transmembrane segments are numbered from the N- to C-terminus and are shown in roman numerals. The phosphorylatable histidine is highlighted in yellow, and tryptophan in red. Conserved residues are highlighted in grey and black according to the alignment in Appendix 11.

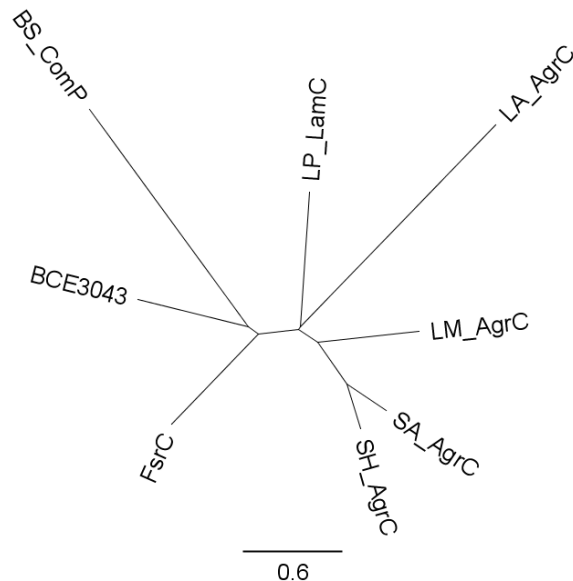


Figure 5.38 Phylogenetic tree of FsrC and Agr family kinases. The tree was created by Geneious Pro 4.6.5 software using amino acid sequences of FsrC (EF1820 of *E. faecalis*), BCE3043 (putative AgrC of *B. cereus*), BS_Comp (CompP of *B. subtilis*), LP_LamC (LamC of *L. plantarum*), LA_AgrC (AgrC of *L. acidophilus*), LM_AgrC (AgrC of *L. monocytogenes*), SA_AgrC (AgrC of *S. aureus*), SH_AgrC (AgrC of *S. haemolyticus*).

Construction of a phylogenetic tree reveals that the Agr kinase proteins of various bacteria do not cluster closely together, except for kinases derived from the same genus, e.g. *S. aureus* and *S. haemolyticus*, suggesting all Agr kinases are at least genus-specific, which is consistent with the different autoinducing peptides that they sense (Figure 5.38). Analysis by SMART web-based tool showed that organisation of the predicted domains, including the transmembrane domains, amongst these Agr kinases is similar (Figure 5.39), though the number of transmembrane regions varies – FsrC has seven putative transmembrane segments compared to other Agr kinases which have six (Figure 5.39).

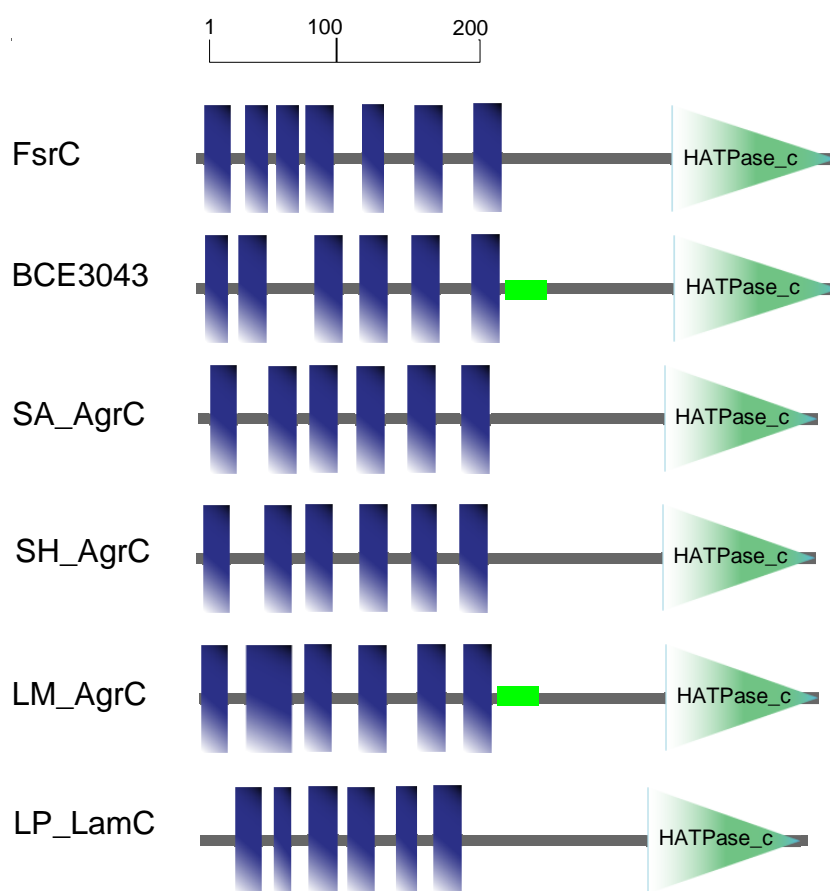


Figure 5.39 Domain analysis of FsrC and its homologues. The figure is based on the graphical output of the SMART web-based tool (Schultz *et al.*, 1998; Letunic *et al.*, 2008). The scale bar is in amino acids. Blue vertical bars represent putative transmembrane segments (■). Green horizontal bar represents coil regions (■). Sizes and positions of conserved domains are indicated by the labelled symbols. FsrC (EF1820 of *E. faecalis*); BCE3043 (putative AgrC of *Bacillus cereus*), SA_AgrC (AgrC of *S. aureus*), SH_AgrC (AgrC of *S. haemolyticus*), LM_AgrC (AgrC of *L. monocytogenes*), LP_LamC (LamC of *L. plantarum*).

5.4.3. Cloning, expression and purification of FsrC

The *fsrC* gene of *E. faecalis* V583 was amplified by PCR and cloned into plasmid pTTQ18-His₆, followed by expression trials which revealed that the intact protein was successfully expressed to 3.5% of *E. coli* total membrane proteins (Table 5.1 and Figure 5.1). Six litre cultures of *E. coli* BL21(DE3) cells transformed with pTTQ18(FsrC)-His₆ were prepared as described in Methods, the cells were harvested by centrifugation and lysed by explosive decompression. The total membrane fraction was isolated by sucrose density gradients (Section 2.4.1 and 2.5.1) and analysed by SDS-PAGE and Western blotting to reveal the presence of intact FsrC in the both inner and outer membrane fractions (Figure 5.40) (constituting approximately 4, 3 and 3% of inner, mixed and outer membrane protein fractions, respectively), indicating poor separation by the sucrose gradients for this protein and therefore perhaps an unessential step in the purification protocol. The inner membranes were subjected to solubilisation trials (Section 2.5.2) which revealed that *NN*-dimethyloctylamine *N*-oxide was a poor detergent for solubilisation of FsrC (Figure 5.40). Thus, it is possible that most of the protein was not successfully solubilised and therefore lost in the pellet during the centrifugation step. DDM and LDAO were optimal for solubilising FsrC (Figure 5.41) and consequently DDM was used for subsequent purifications.

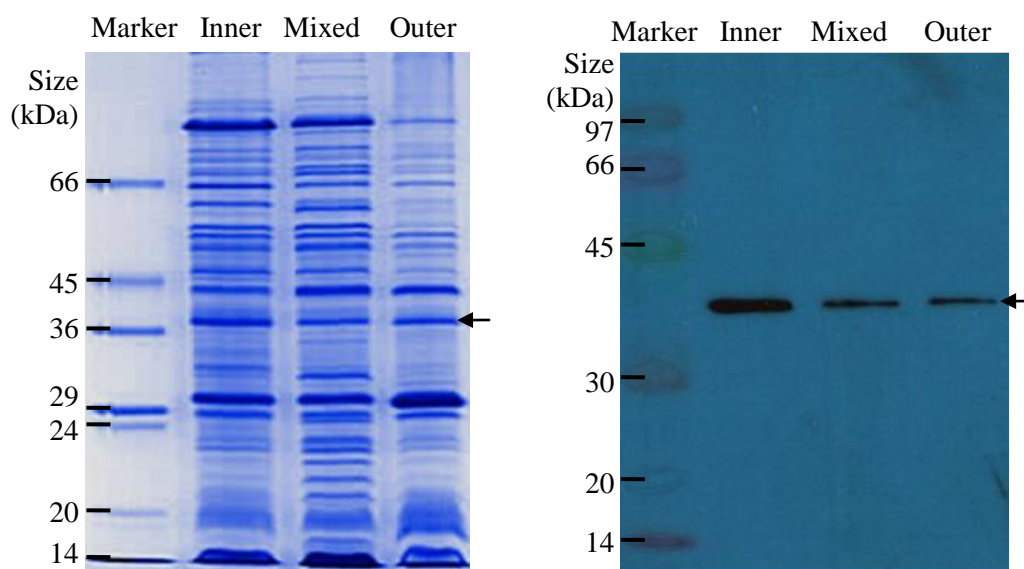


Figure 5.40 SDS-PAGE and Western blot of membrane fractions of *E. coli* BL21(DE3) cells expressing FsrC. Total *E. coli* membranes were prepared from *E. coli* BL21(DE3) pTTQ18(*fsrC*)-His₆ grown on LB medium containing 20 mM glycerol and 100 µg/ml carbenicillin in the presence of 0.5 mM IPTG and harvested 3 hours post-induction. Total membranes were separated on 25-55% sucrose gradients and membrane fractions were extracted from the gradients and analysed by SDS-PAGE and Western blotting. The arrow indicates the position of FsrC (EF1820).

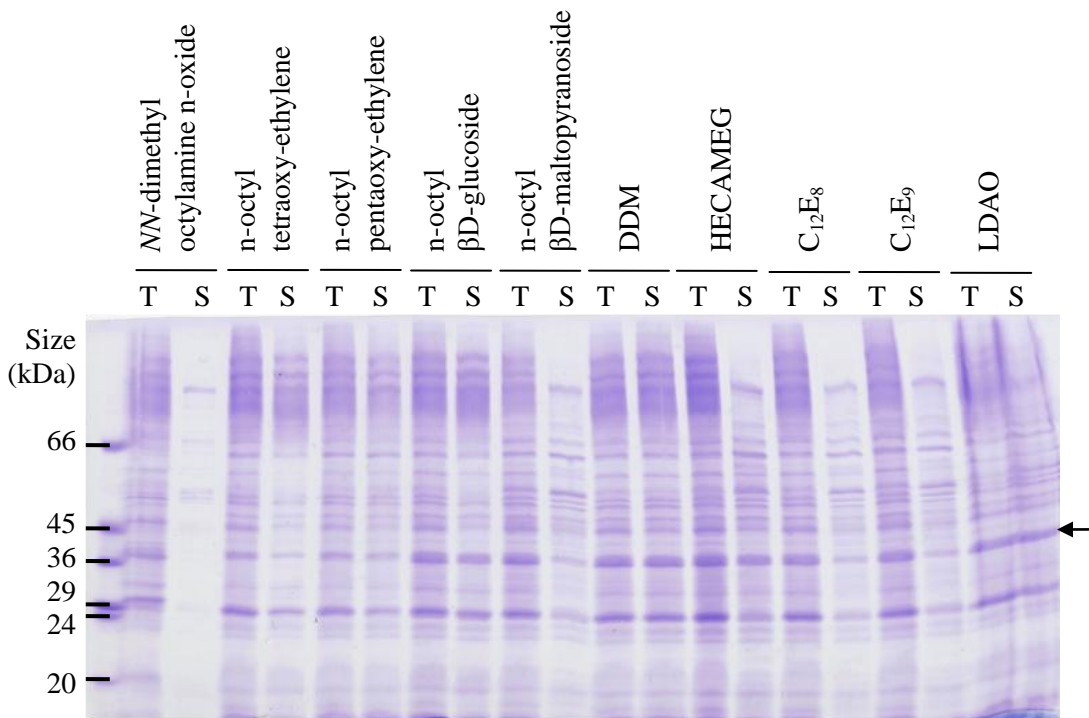


Figure 5.41 SDS-PAGE of total and solubilised inner membranes of *E. coli* BL21(DE3) cells expressing FsrC. Total *E. coli* inner membranes transformed with pTTQ18(*fsrC*)-His₆ were tested for solubilisation in ten detergents. Membranes were solubilised in 1% (w/v) detergent for 1 hour and subjected to centrifugation to remove the insoluble fractions (Section 2.5.2). Total inner membranes (T) and soluble fractions (S) were then analysed by SDS-PAGE. Protein molecular weight markers (M). Arrow indicates position of FsrC.

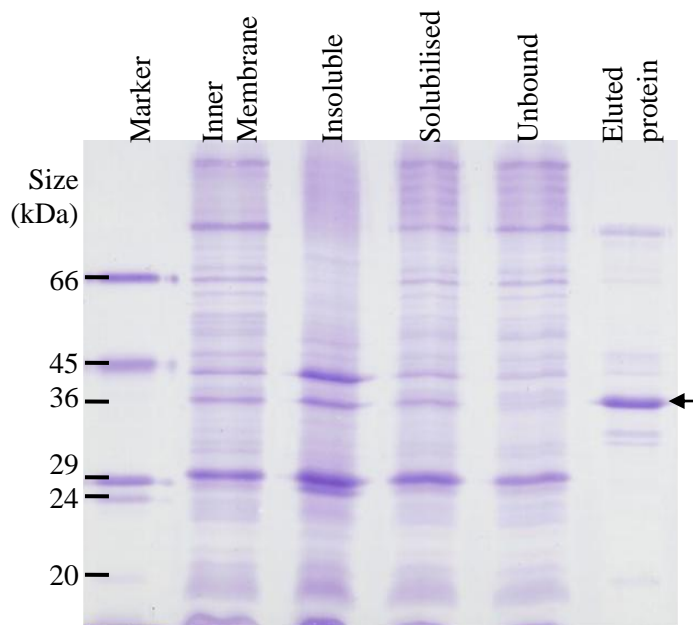


Figure 5.42 SDS-PAGE of fractions isolated during the purification of *E. faecalis* FsrC. FsrC (EF1820) was purified from the inner membranes of *E. coli* BL21(DE3) cells transformed with pTTQ18(*fsrC*)-His₆. The inner membrane proteins were solubilised in 1% (w/v) DDM. Fractions (15 µg), and eluted fraction (5 µg) from the purification are shown. Arrow indicates the position of FsrC.

Fractions resulting from the subsequent purification protocols (Section 2.5.3.2) for FsrC were analysed by SDS-PAGE and confirmed that FsrC was successfully extracted from inner membranes (Figure 5.42). The purified FsrC was confirmed intact by N-terminal sequencing together with the Western blotting shown in Table 5.2 and Figure 5.3.

5.4.4. Autophosphorylation activities of FsrC

The phosphorylation kinetics of FsrC bound within inner membranes are different when compared with those obtained using the purified protein (Figures 5.2 and 5.4). In inner membranes, FsrC exhibited a rapid initial autophosphorylation rate, followed by a gradual dephosphorylation after 20 minutes (Figure 5.2). In the purified form, FsrC showed a steady and slower rate of increase in phosphorylation and maximum phosphorylation was not reached until sixty (Figure 5.4). However, this difference was observed for several of the intact histidine kinases, suggesting that this may be a general feature (Section 5.2 and 5.4).

5.4.5. GBAP specifically increases autophosphorylation of FsrC in *in vitro* autophosphorylation assays

In previous studies, the FsrCA two-component system was shown to regulate virulence genes such as *gelE* and *sprE* (Qin *et al.*, 2000; Qin *et al.*, 2001; Hancock & Perego, 2004; Pillai *et al.*, 2004; Bourgogne *et al.*, 2006), and also that GBAP (Figure 5.43) was the activating signal of the Fsr system (Nakayama *et al.*, 2001; Nakayama *et al.*, 2006).

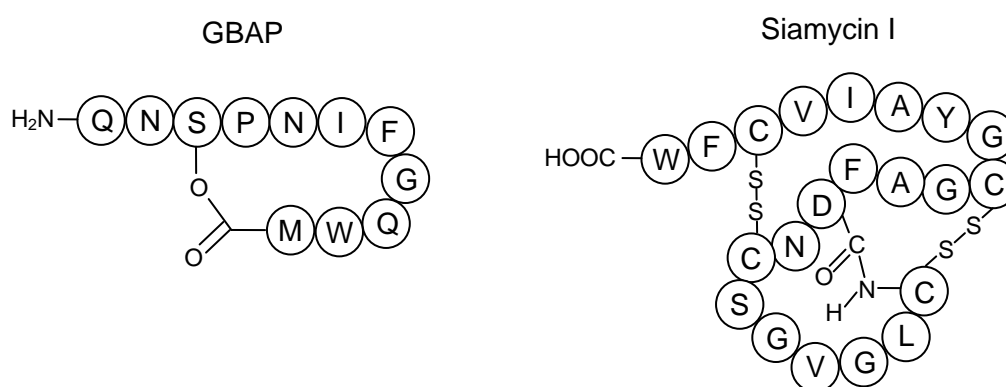


Figure 5.43 Structures of GBAP and siamycin I. GBAP is an eleven residue cyclic peptide containing a lactone structure (Nakayama *et al.*, 2001). Siamycin I is a twenty one residue peptide (Detlefsen *et al.*, 1995; Tsunakawa *et al.*, 1995).

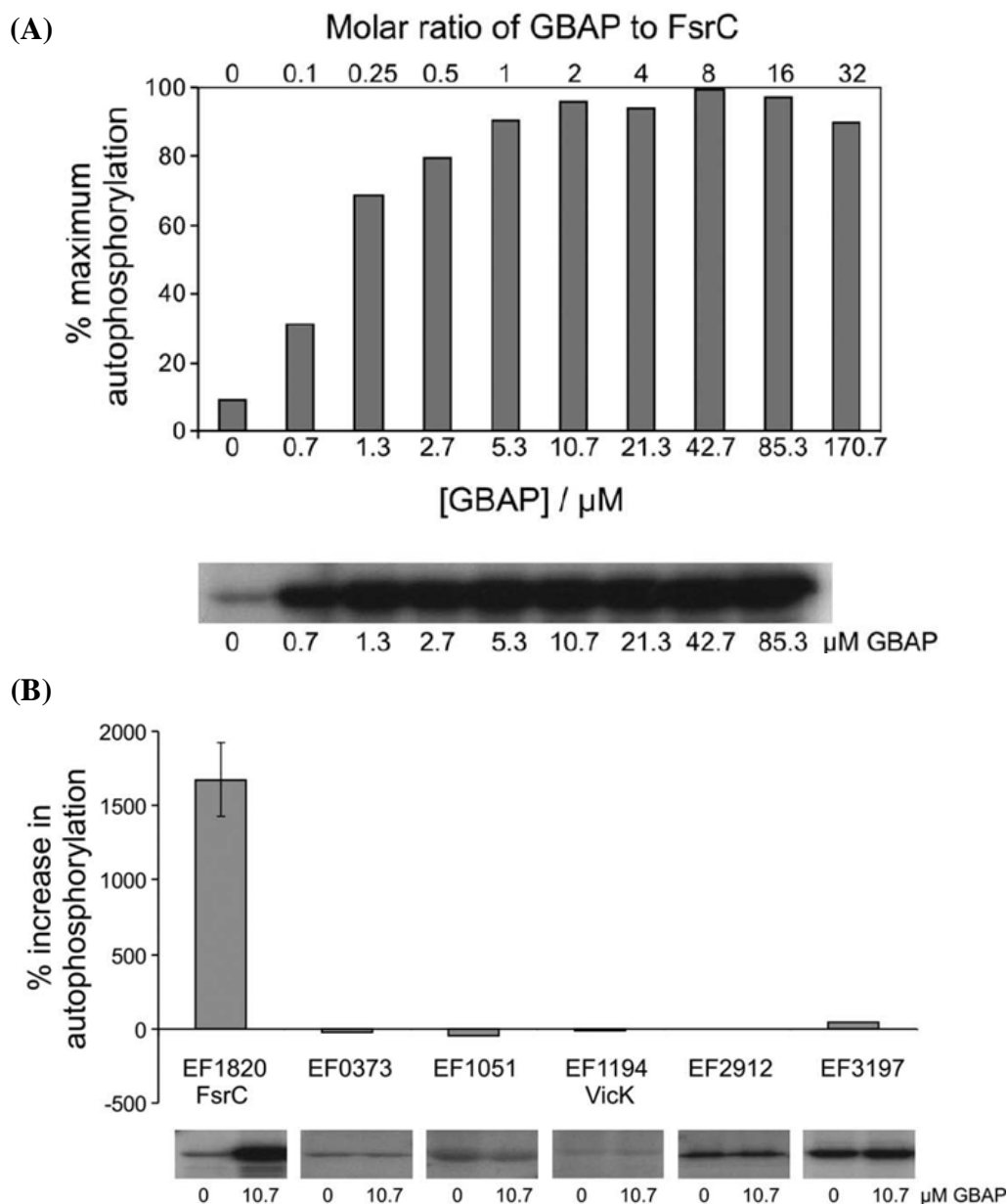


Figure 5.44 Effect of the GBAP peptide pheromone on autophosphorylation activity of FsrC and other purified histidine kinases of *E. faecalis*. Reactions (15 μl) containing 80 pmoles purified proteins were preincubated in reaction buffer for 20 minutes in the presence or absence of synthetic GBAP, followed by initiation of reactions (Section 2.5.5.3). After 60 minutes, 5 μl 4x SDS-PAGE loading buffer was added to terminate the reactions. Phosphorylated protein was determined by SDS-PAGE followed by autoradiography and phosphoimager analysis. All reactions, including controls, contained a final reaction concentration of 3.4% acetonitrile required for dissolving GBAP. (A) FsrC-P levels in response to increasing GBAP. Means of duplicate data are shown (except for 0.7, 21.3 and 170.7 μM GBAP datapoints, for which single measurements were made), and averaged data points varied by ± 0.07 -5.43%. (B) Comparison of responses to 10.7 μM GBAP by six purified histidine kinases. Single data measurements are shown for all proteins except FsrC (EF1820) which shows the mean and standard deviation derived from triplicate data. Autoradiographs of individual data are shown underneath the graph. These results have been published in Ma *et al.* (2008) and are solely my own work.

GBAP was previously chemically synthesised in an active form (Nakayama *et al.*, 2001b) and demonstrated to successfully induce transcription of *fsrB* and *fsrC*, as well as *gelE* and *sprE*, at 1 to 30 nM concentrations (Nakayama *et al.*, 2001; Nakayama *et al.*, 2007), though the exact mechanism is not yet known. To test whether the *in vitro* approach adopted throughout the study of histidine kinases is successful for identifying cognate signals amongst purified intact histidine kinases in autophosphorylation assays, and to determine whether synthetic GBAP induces activation through direct interactions with FsrC, and that FsrC is able to respond to its expected GBAP signal *in vitro*, purified FsrC was pre-incubated with a range of synthetic GBAP concentrations for twenty minutes, followed by autophosphorylation assays (Section 2.5.5.3). Mass spectrometry confirmed GBAP is in the active cyclic form, described in Section 5.4.8.

Upon addition of increasing concentrations of GBAP, an increased level of phosphorylated FsrC was observed (Figure 5.44). A maximum level of autophosphorylation was reached using ~10 μ M GBAP, that is a two-fold molar excess of GBAP over FsrC (Figure 5.44A). Further increase in GBAP concentrations up to 170.7 μ M elicited only small further changes in the levels of phosphorylated FsrC (Figure 5.44A). To test if the action of GBAP is specific for the FsrC histidine kinase, other purified *E. faecalis* histidine kinases with no known or predicted roles in quorum sensing were tested as described above using 10.7 μ M GBAP. Little change in the activities of the other histidine kinases of *E. faecalis* was observed in presence of GBAP (Figure 5.44B), confirming that GBAP specifically exerts a strong effect on FsrC.

5.4.6. Siamycin I inhibits GBAP-induced autophosphorylation of FsrC

Quorum sensing, which has been shown to be involved in virulence regulation of many Gram-positive bacteria, has been proposed as a new target for antimicrobial drug therapy (Smith & Iglewski, 2003; Otto, 2004; Raffa *et al.*, 2005; Martin *et al.*, 2008; Kociolek, 2009). Drugs that were effective in inhibiting quorum sensing were also shown to be effective for attenuating virulence without bactericidal or bacteriostatic activities, implying that the generation of drug resistance might be minimal. In a screen for inhibitors of *fsr* quorum sensing, Nakayama *et al.*, (2007) showed that siamycin I (Figure 5.43) inhibited both gelatinase and GBAP production at submicromolar concentrations without inhibiting *E. faecalis* cell growth, and that cell growth was inhibited above micromolar concentrations. At sublethal concentrations of siamycin I, both *fsrBDC* and *gelE-sprE* transcripts were suppressed, and an overdose of GBAP

added to the culture did not cause any effect, suggesting siamycin inhibited GBAP signalling via the FsrCA two-component system (Nakayama *et al.*, 2007).

To test if siamycin I exerts a direct effect on the FsrC protein, assays (15 μ l) employing purified FsrC (80 pmole) and GBAP (160 pmole) were preincubated for twenty minutes in the presence or absence of up to a ten-fold molar excess of siamycin I (0 – 100 μ M), and reactions were allowed to proceed for 30 minutes (Section 2.5.5.3). The results are shown in Figure 5.45 and reveal that whilst no inhibitory effect of siamycin I on GBAP-induced autophosphorylation at 0 – 12.5 μ M (0 – 1.3 fold excess) siamycin I was detectable, at higher concentrations of inhibitors (25 – 100 μ M), an inhibitory effect was observed (Figure 5.45). The proportion of phosphorylated FsrC was reduced by 29, 76 and 91% respectively, indicating that siamycin I does exert a direct inhibitory effect on FsrC autophosphorylation activity.

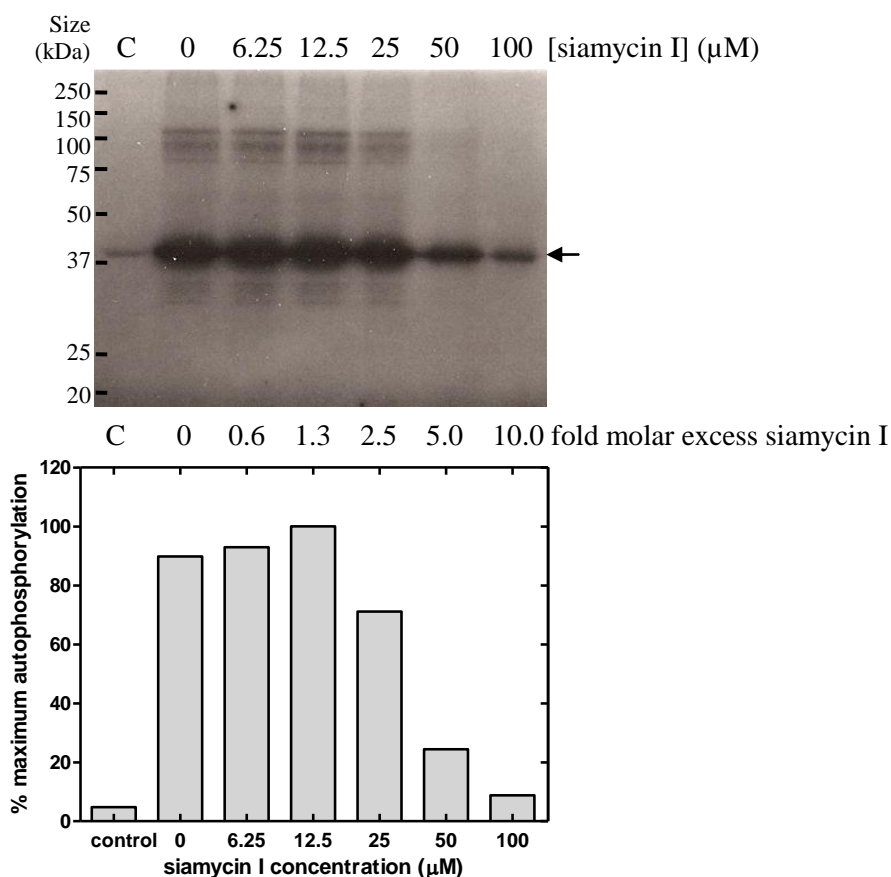


Figure 5.45 Inhibition of GBAP-induced autophosphorylation of FsrC by siamycin I. Purified FsrC (80 pmoles) and a 2-fold molar excess of GBAP (10.7 μ M) were preincubated for 20 minutes in a final reaction volume of 15 μ l in the presence or absence of siamycin I (0 – 100 μ M), followed by initiation of activity reactions (Section 2.5.5.3). Reactions proceeded for 30 minutes. Autoradiographic film (top) and graphical representations (bottom) of phosphorylated FsrC are shown. Control (C), which contains no GBAP and no siamycin I. Arrow indicates the position of FsrC.

5.4.7. Time course of FsrC autophosphorylation in the presence and absence of GBAP and/or siamycin I

Time course experiments were then performed to further characterise the effects of GBAP and siamycin I on autophosphorylation rates of FsrC (Figure 5.46). Assays were performed (Section 2.5.5.2) using purified FsrC in the presence and absence of GBAP (10.7 μM) and/or siamycin I (50 μM). An approximate seven-fold increase in phosphorylated FsrC was obtained after 60 minutes when GBAP was present in the assays compared with when GBAP was absent (Figure 5.46).

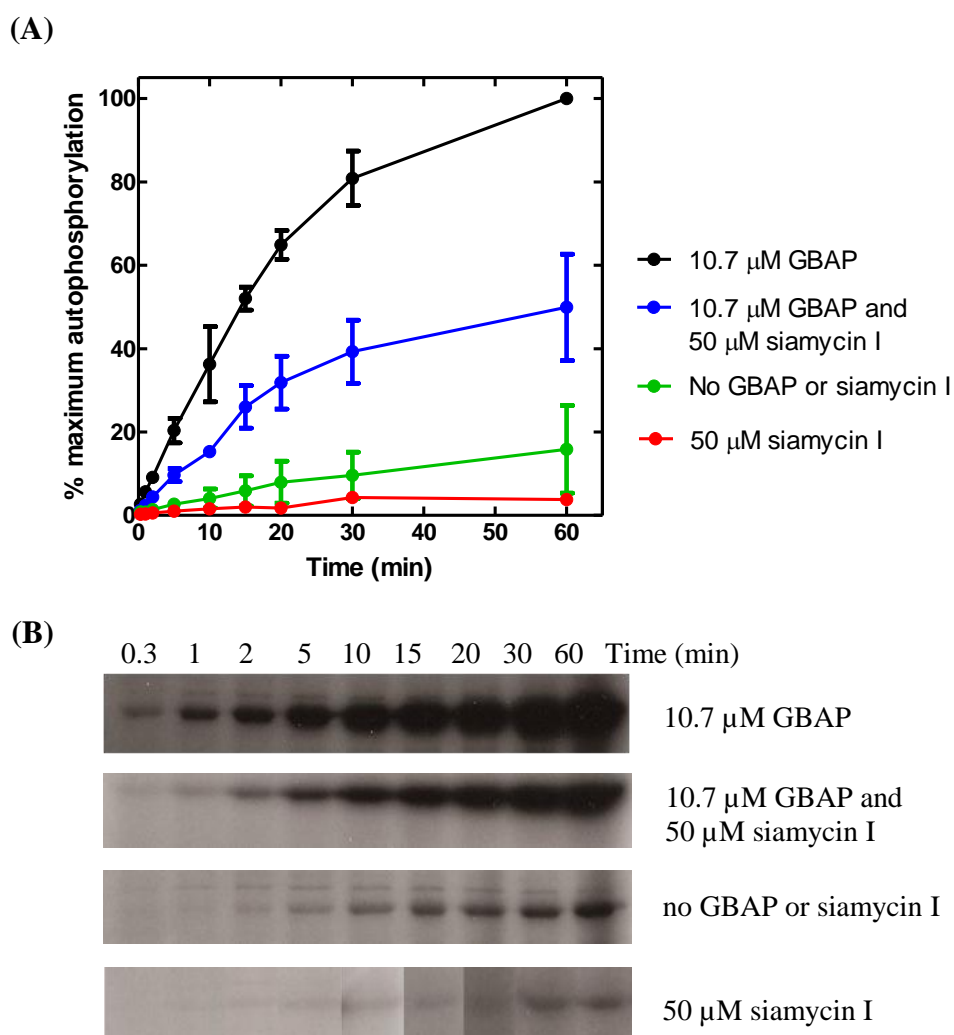


Figure 5.46 Time course of autophosphorylation of FsrC in the presence and absence of GBAP and/or siamycin I. Reactions (150 μl) containing 800 pmoles purified proteins was preincubated in reaction buffer for 20 minutes in the presence and absence of two-fold molar excess of GBAP (10.7 μM) and/or a 5-fold molar excess of siamycin I (50 μM), followed by initiation of reactions (Section 2.5.5.2). Samples (15 μl) were removed at time intervals between 0 and 60 min. All samples were analysed as described in Section 2.5.5.1. Data represent the mean of triplicate experiments. (A) Graphical representation of data. (B) Autoradiographic film.

The presence of 50 μM siamycin I in the presence of GBAP resulted in reduced FsrC-P levels, which were 50% of the controls at the 60 minute time point (Figure 5.46). With siamycin I alone, the level of FsrC activity also appeared to be lower than the control, suggesting siamycin I could perhaps inhibit constitutive activity of FsrC as well as GBAP-stimulated activity, but since the error bars on the control data overlap with the siamycin I only data, this data is not statistically significant. Further tests such as use of higher concentration of siamycin I ($\geq 100 \mu\text{M}$) could perhaps be useful to determine whether any statistical significance occurs in the absence of GBAP. Nevertheless, the assays here further confirm that siamycin I has a significant inhibitory effect on GBAP-induced activity of FsrC.

5.4.8. Mass spectrometry of synthetic GBAP

The synthetic GBAP used in this project was prepared by Kenzo Nishiguchi (Kyushu University, Japan). Active GBAP contains a cyclic lactone ring (Figure 5.43) and has a molecular weight of 1302.6 (Nakayma *et al.*, 2001). The GBAP used for the autophosphorylation assays in the present study were confirmed as cyclic and intact by ESI-mass spectrometry (Figure 5.47). Figure 5.47A shows GBAP in the 1+ and 2+ charge states. A peak which is present next to the 1303.58 peak was analysed and confirmed to be a sodium atom attached to GBAP (Figure 5.47B).

5.4.9. Circular dichroism spectroscopy of FsrC

The secondary structure integrity of purified FsrC was examined by circular dichroism (Section 2.6.1). The resultant spectrum is shown in Figure 5.48 and reveals the presence of alpha helices and therefore of intact secondary structure in the purified protein. Any observable changes in the secondary structure of FsrC upon ATP and GBAP binding was then investigated by further CD experiments. Firstly, a spectrum of FsrC alone was measured (dark blue line), followed by addition of 50 μM ATP (orange line) and 10.7 μM GBAP (grey line), accumulatively. Figure 5.49 shows that a change in the secondary structure of FsrC was detectable upon addition of ATP (Figure 5.49) and that this could be a slight relaxation of the protein. Addition of GBAP to the protein caused a negative shift in the curve (Figure 5.49). To determine if the shift is caused by changes in the structure of FsrC or by absorbance of GBAP, a spectrum of GBAP alone was made (light blue line) (Figure 5.50). GBAP absorbed at 190 – 240 nm,

shown by the negative indentations, which matches with a recent published CD spectrum of GBAP (Nishiguchi *et al.* 2009).

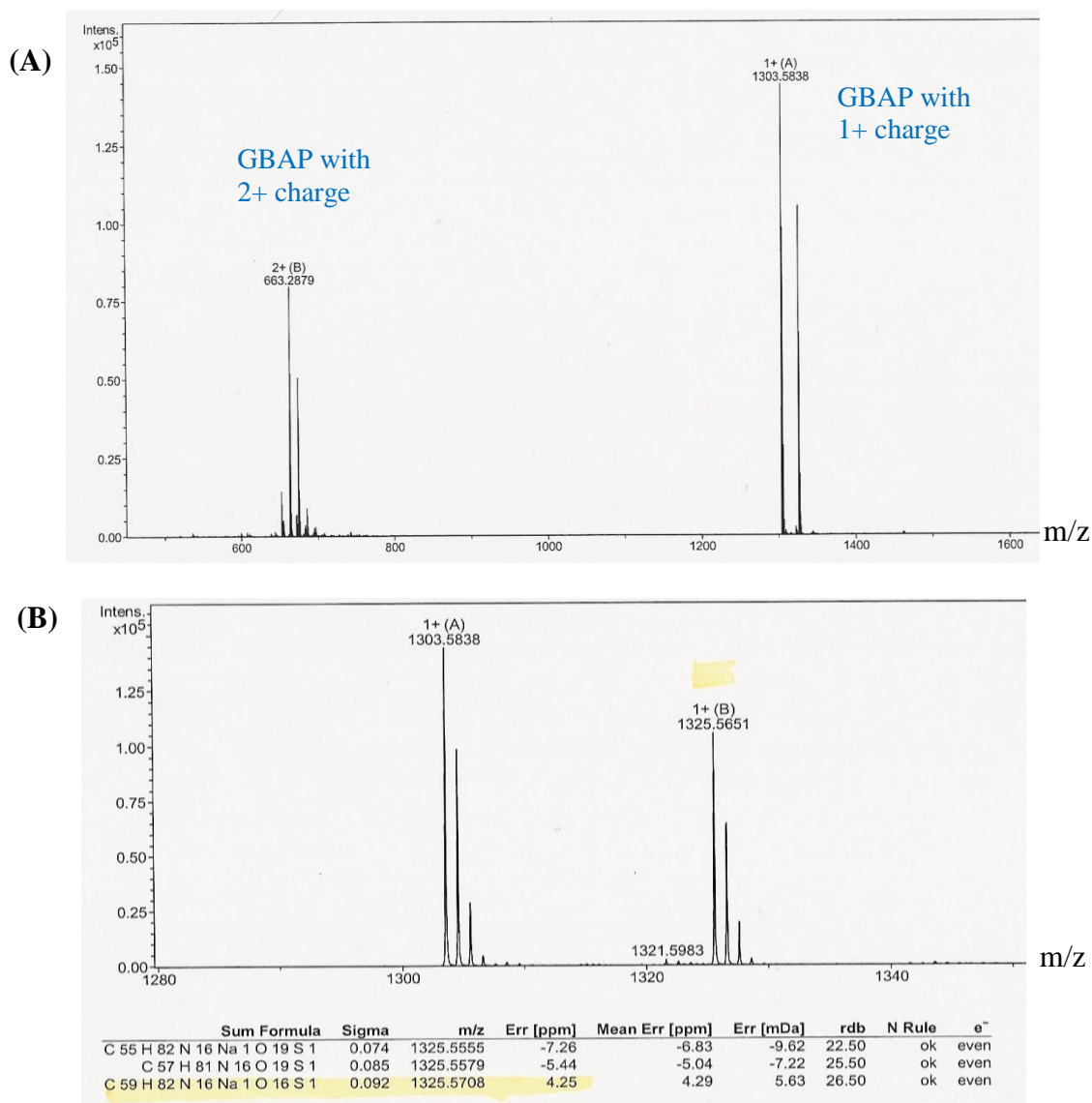


Figure 5.47 Electrospray ionisation mass spectrometry of GBAP. Samples were prepared in 3.4% acetonitrile and submitted for mass spectrometry analysis by School of Chemistry mass spectrometry service, University of Leeds. (A) positive ESI-mass spectrum; (B) close up view of (A).

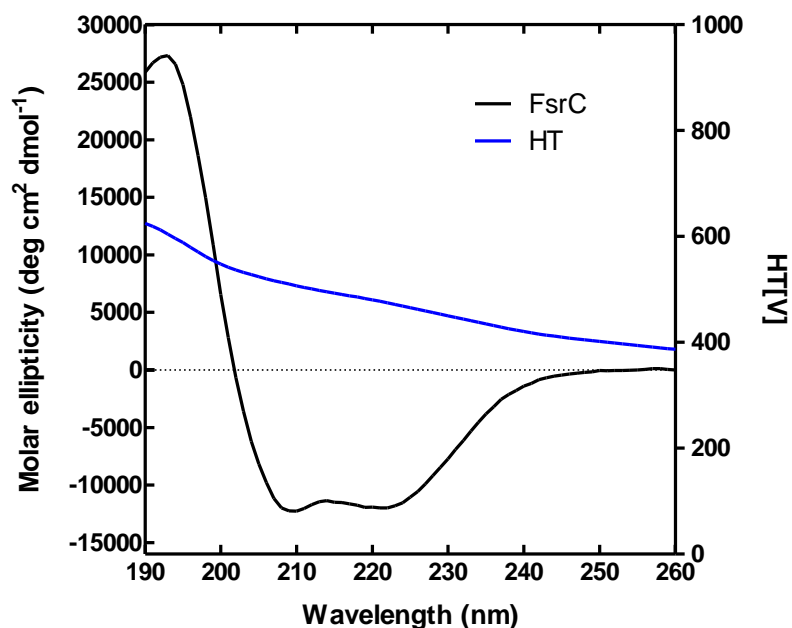


Figure 5.48 Circular dichroism spectrum of purified FsrC. Purified FsrC (0.05mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of FsrC was performed using a Jasco J-715 spectropolarimeter at 18°C with constant nitrogen flushing. The sample were analysed in Hellma quartz-glass cell of 1 mm path length. Spectrum was recorded with 1 nm step resolution at a scan rate of 10 nm/min. The response time was set at 1 second with a sensitivity of 20 mdeg and bandwidth of 1.0 nm. The spectrum represents an accumulation of ten scans, from which the buffer contribution was subtracted. The blue line represents the voltage applied to the photomultiplier.

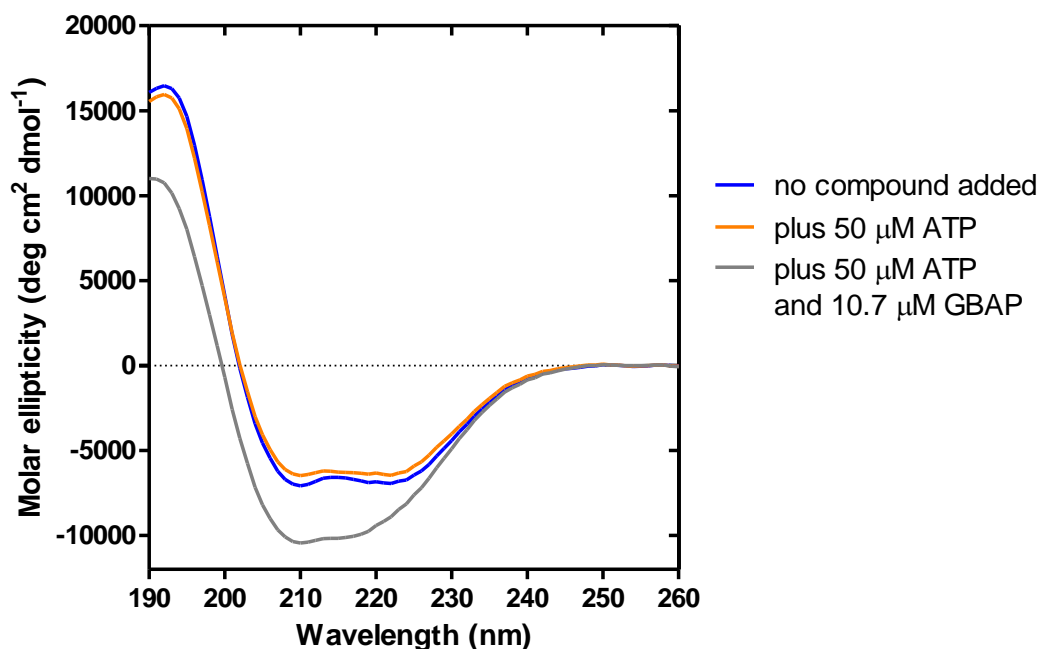


Figure 5.49 Circular dichroism spectrum of purified FsrC with ATP and GBAP. Purified FsrC (0.05mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of FsrC was performed as described in legend of Figure 5.48.

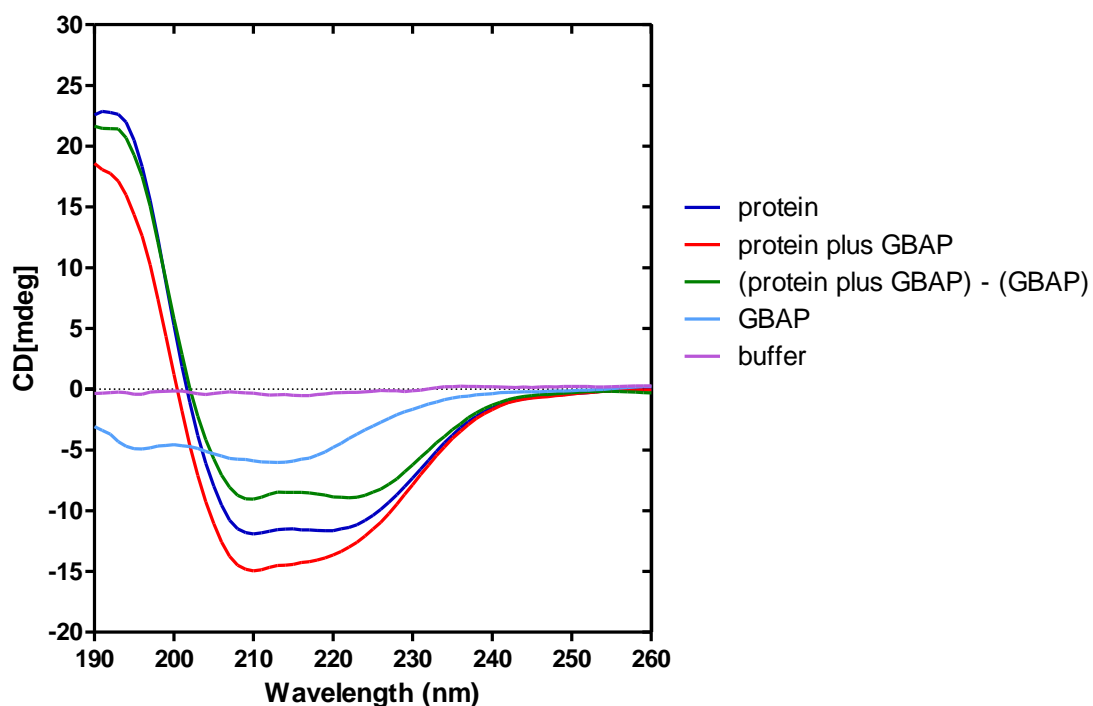


Figure 5.50 Circular dichroism spectrum of purified FsrC with GBAP. Purified FsrC (0.05mg/ml) was buffer-exchanged into 10 mM potassium phosphate buffer pH 7.6 and 0.05% DDM. CD spectral analysis of FsrC was performed as described in legend of Figure 5.48.

The effect of GBAP on the secondary structure of FsrC was further investigated. A spectrum of buffer alone was measured, and as expected it showed low level of absorbance (Figure 5.50). Next, FsrC was added, the spectrum measured (dark blue line) and this was followed by accumulative addition of 10.7 μ M GBAP and a measurement made (red line). The spectrum of ‘GBAP alone’ was subtracted from ‘FsrC plus GBAP’, giving rise to the change in secondary structure of FsrC upon GBAP binding (green line) (Figure 5.50). Comparing this to ‘FsrC alone’ (dark blue line), a reduction of the α -helical content of FsrC was observed in presence of GBAP (green line) (Figure 5.50).

5.4.10. Fluorimetry study of FsrC

There are four predicted tryptophan residues in the FsrC protein. Three are predicted to be located in the putative transmembrane segments and one in the cytoplasmic domain (Figure 5.37). Fluorimetry measurements of intrinsic tryptophan residues of FsrC were made upon titration by ATP and GBAP (Figure 5.51 and 5.52). Titration by ATP alone was made, and changes in fluorescence reaching a maximum intensity of \sim 3800 were observed (Figure 5.51). A second titration was made with the

FsrC protein, for which the ATP fluorescence is subtracted to give FsrC fluorescence alone (Figure 5.51). Firstly, concentrations of up to 5 μM ATP caused a rapid decline in FsrC fluorescence level, suggesting ATP binding could cause tighter packing of the FsrC structure leading to less tryptophan residues being exposed. Further addition of ATP then led to a slow increase; this could be caused by hydrolysis of ATP to ADP leading to a reverse effect.

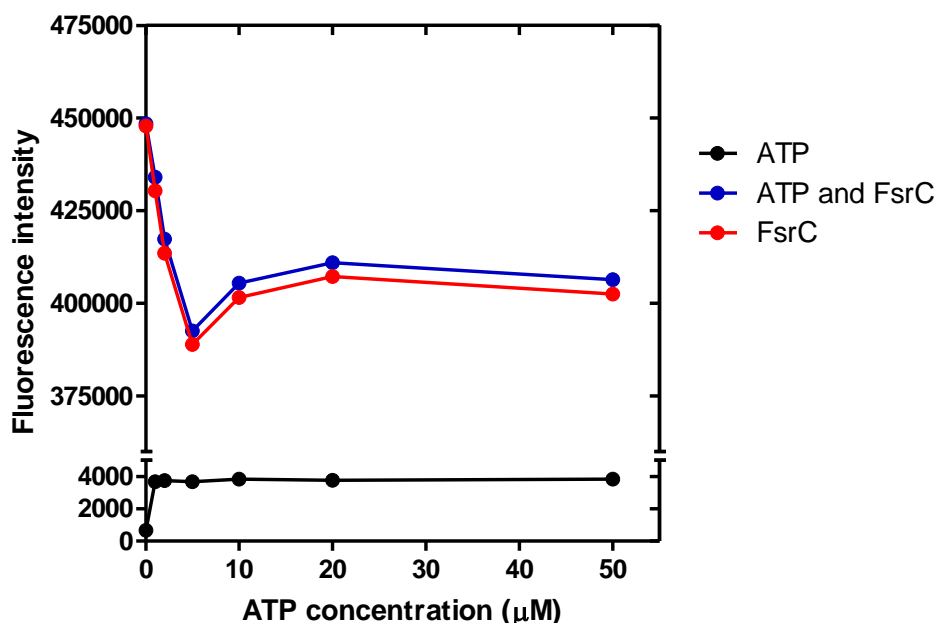


Figure 5.51 Tryptophan fluorescence measurements of FsrC titrated by ATP. FsrC was titrated with ATP and the changes in the tryptophan fluorescence at 348 nm monitored (Section 2.6.3.4). The measurements were performed with titration by ATP alone (●), followed by titration by ATP in the presence of FsrC (●). The fluorescence of ATP alone was then subtracted from this to produce fluorescence of FsrC alone (●).

In the titration experiments using GBAP alone, GBAP produced a rapid increase in fluorescence (Figure 5.52); this is caused by the presence of a tryptophan residue in the GBAP peptide (Figure 5.43). Another titration by GBAP in the presence of FsrC was made, and the GBAP fluorescence was subtracted from this to produce values for FsrC fluorescence alone (Figure 5.52). A gradual decrease in FsrC fluorescence was observed with increasing GBAP concentrations, suggesting that GBAP could be less exposed and more buried in the FsrC protein, thereby resulting in exposure of fewer tryptophan residues. A tryptophan-free FsrC, in which the activity is not affected, could be used to investigate further the binding site of GBAP, and whether it binds between the helical bundles and/or periplasmic loops.

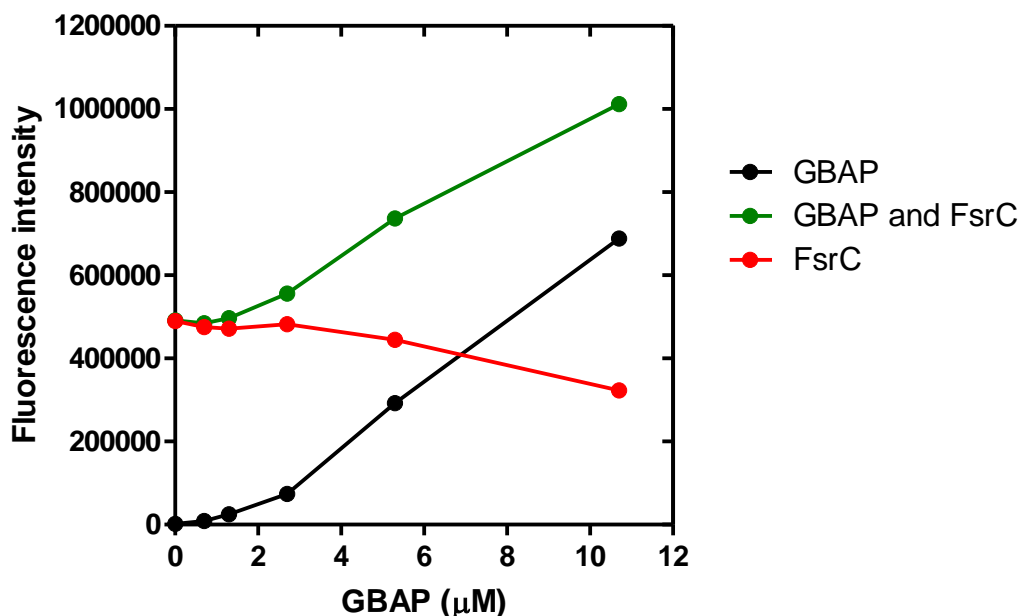


Figure 5.52 Tryptophan fluorescence measurements of FsrC titrated with GBAP. The FsrC solution was titrated with GBAP and the changes in the tryptophan fluorescence at 348 nm monitored (Section 2.6.3.5). The measurements were performed with titration by GBAP alone (●), followed by titration by GBAP in the presence of FsrC (●). The fluorescence of GBAP alone was then subtracted from this to produce FsrC fluorescence alone (●).

5.4.11. Cloning of the putative GBAP transporter: EF1821 (FsrB)

In a rabbit endophthalmitis model, it was shown that *fsrB* was required for function of the Fsr quorum sensing system of *E. faecalis* (Mylonakis *et al.*, 2002). A nonpolar deletion mutation in *fsrB* was shown to significantly reduce virulence compared to wild type, and complementation restored the virulence (Mylonakis *et al.*, 2002). It appears that FsrB is involved in the catalysing and release of GBAP (Nakayama *et al.*, 2006), and therefore this study included the cloning of this membrane protein for further characterisation.

The *fsrB* (EF1820) gene was therefore amplified by PCR from the chromosomal DNA of *E. faecalis* V583 and cloned into plasmid pTTQ18-His₆ to generate pTTQ18(*fsrB*)-His₆. DNA sequencing of the resultant clone confirmed that the gene had been successfully cloned (Appendix 12).

5.4.12. Discussion: study of the Fsr quorum sensing pathway of *E. faecalis*

Enterococci have become recognised as important global causes of nosocomial infections, being major contributors to hospital-acquired infections, associated with urinary tract infections, endocarditis, pelvic infections and surgical wound infections. Their ability to form biofilms on medical devices is an important aspect of pathogenesis in the hospital environment. Over 60% of microbial infections in the body are associated with biofilms, and since bacteria in biofilms are resistant to phagocytosis, this makes it difficult to eradicate them from the living host (Lewis, 2001). Clinical isolates growing as biofilms are more resistant than in their planktonic state to antibiotics such as ampicillin, linezolid and the last line glycopeptide vancomycin (Hancock & Perego, 2004a; Sandoe *et al.*, 2006). A mature biofilm is able to tolerate antibiotics at concentrations of 10 – 1000 times greater than those required to kill planktonic bacteria.

The importance of *fsr* of *E. faecalis* in virulence regulation and biofilm development have been documented in a few studies (Qin *et al.*, 2000 Nakayama *et al.*, 2001; Mylonakis *et al.*, 2002; Hancock & Perego, 2004; Nakayama *et al.*, 2006; Thomas *et al.*, 2008). The FsrC histidine kinase regulates production of extracellular proteases GelE and SprE, involved in regulating proteolysis and release of eDNA, a critical component for *E. faecalis* biofilm formation (Thomas *et al.*, 2008). GelE was shown to activate lysis of a subpopulation of bacteria and consequently catalyse the release of eDNA; and SprE opposed the phenotype of GelE showing higher rates of lysis (Thomas *et al.*, 2008). Three autolysins (AtlA, AtlB and AtlC) were identified to be secreted by *E. faecalis* and one of these, AtlA, is thought to be important for biofilms (Kristich *et al.*, 2008; Mesnage *et al.*, 2008). A recent study showed GelE together with SprE regulated enterococcal fratricide leading to eDNA release and biofilm development and AtlA was demonstrated to be the target of these proteases (Thomas *et al.*, 2009). A model was suggested by Thomas *et al.* (2009), in which GelE mediates fratricide in *E. faecalis* in response to GBAP; the release of GelE extracellularly binds to cells surface mediating fratricide, and SprE that is co-secreted was suggested to act as an immunity protein where binding to cell surface prevented autolysis and fratricide (Thomas *et al.*, 2009).

In this work it was demonstrated that purified full-length FsrC exhibited an approximately ten-fold increase in its autophosphorylation activity in response to a two-fold molar excess of synthetic GBAP. This provides the first evidence for direct and

specific interactions between GBAP and FsrC and in fact for any peptide pheromone with its quorum sensing kinase. It is therefore possible that future work will be expected to be successful for identifying signals for other intact histidine kinases expressed and purified in this study (Ma *et al.*, 2008), though there are some factors to be borne in mind. For example, as mentioned earlier (Section 5.3.4), some detergents were found to be inhibitory towards the activities of some kinases which was first identified by Dr Phillips-Jones for VicK and later shown to be the case also for EF1051 in the present study (Figure 5.31). Also high concentrations of some ligands are needed to test direct interactions with histidine kinases; for example, relatively high concentrations of reducing agents (20 – 100 mM DTT or dithionite) were required to elicit a visible effect on the autophosphorylation of histidine kinase EF3197 (Figure 5.17), and this was the same for redox sensors RegB (Potter *et al.*, 2006).

The study described will not only allow for identification of signals, but also for identification of inhibitors. This study highlights the potential of one candidate inhibitor, siamycin I which is also identified for its anti-HIV activities (Lin *et al.*, 1996). Concentrations of 25 – 100 μ M (2.5 – 10 fold excess) siamycin I inhibited GBAP-induced autophosphorylation of FsrC (Figure 5.45 and 5.46).

Future work on FsrC might include further tests, such as phosphotransfer in response to its cognate response regulator FsrA, and DNA-binding assay for phosphorylated FsrA could be done in the future work. Also, the threshold for GBAP-induced autophosphorylation of purified FsrC could be determined by testing lower concentrations of GBAP – following the experiment described in Figure 5.44A. The assay using purified kinase could be disadvantageous for screening expensive ligands, but nevertheless, it provides direct interaction of protein and its ligand and thus for signal/ligand screening.

The *frsB* (truncated EF1821) gene with no *frsD* portion was cloned in this study (Appendix 12). A previous study by Nakayama *et al.* (2009) showed that ambuic acid, a fungal metabolite, targeted FsrB and that it inhibited the quorum sensing-mediated gelatinase production of *E. faecalis* without influence on cell growth. It was also demonstrated using His₆-tagged FsrD (GBAP propeptide) that ambuic acid prevented the release of GBAP (Nakayama *et al.*, 2009). Further study of this clone, together with FsrC, will enable deeper understanding of the *frs* quorum sensing system of *E. faecalis*.

Bacteria are able to live in a planktonic form, but they are advantageous to live within a biofilm. Biofilms provide protection for the bacteria in a hostile environment

while maintaining nutrients for inhabitants, protecting them from immune response, as well as an inherent antibiotic resistance that would not be present in their planktonic form (Sandoe *et al.*, 2006). This resistance is a major problem in persistent and chronic bacterial infections, and underscores the urgent need for innovative therapeutics. To date, many medically relevant pathogens have been identified to have quorum sensing systems related to biofilm formation, and therefore inhibiting biofilm formation has become medically important.

Traditional treatment of infectious diseases often involved compounds that kill or inhibit bacterial growth, and the major concern with these approaches are that bacteria are becoming resistance to antimicrobial compounds. The discovery of quorum sensing in bacteria, which coordinate important chronological events during the infection process, have presented an important novel means to alleviate bacterial infections, and is now a target for novel therapeutics (Hentzer & Givskov, 2003; Raffa *et al.*, 2005; Khmel & Metlitskaya, 2006; Martin *et al.*, 2008; Njoroge & Sperandio, 2009; Raina *et al.*, 2009). Quorum sensing regulation is species-specific and therefore drugs that selectively inhibit the pathogen would not have any interference effect on other naturally habiting microflora. There are a few ways to inhibit quorum sensing: inhibition of autoinducer production, inhibition of autoinducer binding with the receptor protein, and acylated homoserine lactone degradation in Gram-negative bacteria (Khmel & Metlitskava, 2006). Continued investigation of the Fsr quorum sensing system of *E. faecalis* will increase our understanding of enterococcal pathogenesis, and will aid the development of novel drugs that may overcome the problem of antibiotic resistance and for the treatment of hospital-acquired infections. In particular, FsrC that senses, and FsrB that is required for the processing and transport of, the autoinducing peptide, may be potential therapeutic targets, and possible elucidation of their structure to promote drug development against enterococcal infections.

In the case of FsrC, not only is there a potential for designing inhibitors of FsrC activity to prevent expression of gelatinase (GelE) and serine protease (SprE) involved in adherence in hospital strains, but there is also a potential to design cyclic pheromones or other peptides that increase FsrC activity, rather than inhibit it, in probiotic strains. Enterococci (in which FsrC is found) form part of the lactic acid bacteria and are natural members of the microbiota of human and animal intestinal tracts, and are used in health supplements and human probiotics by the food industries (e.g. *E. faecalis* Symbioflor D). They are also used as starter cultures in the food industry, where they play a beneficial

role in the production of various fermented foods such as traditional cheeses and sausages, playing important roles in ripening and aroma development. Modulation (particularly increases) of adherence characteristics through modulation of FsrC activities will be of potential benefit for increasing the survival and longevity of probiotic strains in the intestine and of food strains during food processing. A knowledge of GBAP interactions with the molecular structure of FsrC will play an important role in the design of new peptides and/or engineered FsrC proteins that can be used by the food industry to activate or increase adherence characteristics.

Chapter 6

Conclusions and future perspectives

6.1. Summary of results – towards structural and functional determination of bacterial transport and sensory membrane proteins

In both prokaryotes and eukaryotes, approximately 25 – 35% of the open reading frames in sequenced genomes encode for membrane proteins, which perform many important biological functions, from transport of solutes to signal transduction. Many successes in functional determination of these proteins have been established. However, there is a lack of detailed understanding of their structure-activity relationships, which is of paramount importance to facilitate drug design and development.

In this thesis, the importance of the NCS1 family of transporters and the histidine kinases of the two-component system are described (Chapter 1). The ultimate aim of this research is to facilitate elucidation of their structure-activity relationships. The results presented in this research describe a genomic approach for the study of these membrane proteins, and the methods used to accomplish this research are described in Chapter 2.

In Chapter 3, experiments are described to investigate the NCS1 transporters. A genomic approach was utilised to produce full-length His₆-tagged protein for structural and functional study (Section 3.1). Thirteen of the selected targets were cloned, and another seven are undergoing the cloning stages. Twelve of the thirteen cloned proteins were successfully expressed in *E. coli* membranes, and substrates transported by five of them were confirmed or identified. Three of these transporters: CodB – a cytosine transporter (Section 3.2); PucI – an allantoin transporter (Section 3.3); and PA0443 – a uracil transporter (Section 3.4) were characterised, and their substrate specificity and kinetics were also determined. Bacteria are known to have many transport proteins with different kinetics, and transport of a particular substrate may occur through possession of multiple distinct transport systems. This presumably permits bacteria to switch to transport systems that work better at a particular substrate concentration. For example, if low concentrations of substrate prevail then bacteria are predicted to utilise transport system of lower K_m .

The molecular mechanisms for substrate transport by the NCS1 family of proteins are not completely understood, but three structures of Mhp1 (Weyand *et al.*, 2008; Shimamura *et al.*, 2010), each in a different state, provides fundamental information regarding their transport mechanism. The finding that transporters of the NSS, SSS, APC and BCCT families are structurally similar to those in the NCS1 family,

imply that these sodium-coupled symporters function by the same molecular mechanism.

The study described in Chapter 4 was aimed at determining the identity of the cation coupled to benzylhydantoin uptake by Mhp1. Radiolabelled transport assay and fluorimetry experiments were performed in this study to elucidate this identity, and these revealed that binding of benzylhydantoin to Mhp1 is sodium-dependent, with an increased affinity for benzylhydantoin when 15 mM NaCl was present.

Chapter 5 reports an investigation of the genome complement of predicted histidine kinases of *E. faecalis*. Using the same genomic approach for studying NCS1 transporters, each of the sixteen genes encoding the full-length membrane histidine kinase of *E. faecalis* were successfully cloned and subjected to expression trial (Section 5.1). Fifteen of the sixteen histidine kinases were expressed successfully in *E. coli*, fourteen were solubilised from the inner membrane fractions, and thirteen successfully purified (Section 5.1). Eleven of the fifteen membrane-bound proteins and twelve of the thirteen purified proteins were active in autophosphorylation assays, a striking success for production of active membrane histidine kinases, and in fact this is the first reported study of all intact membrane histidine kinases from a complete bacterial genome. Additional understanding for three histidine kinases EF3197, EF1051 and EF1820 (FsrC) was achieved through further studies, including the following progress and findings: (1) EF3197 responds to increasing concentrations of reducing agents, suggesting a possible role in redox-sensing (Section 5.2); (2) EF1051 was purified to 90 % purity and subjected to gel filtration, followed by promising crystallisation trials for future structural determination (Section 5.3); and (3) FsrC, the quorum sensing kinase of *E. faecalis* responded to increasing concentrations of its pheromone signal, GBAP, and GBAP-induced activity was inhibited by siamycin I (Section 5.4). The work with FsrC and its known GBAP signal established unequivocally that use of intact sensor kinases in *in vitro* activity assays was sufficiently sensitive and specific to identify cognate environmental signals, paving the way for screening of purified and membrane-bound proteins for systematic signal identification.

6.2. Impact of this research and potential applications

Drug discovery is an extensive and risky process with many pitfalls. Classic drug discovery has involved selection of appropriate compounds, followed by biological evaluation in lead discovery, which is a time-consuming process (Lombardino & Lowe,

2004). More recently, a number of high-throughput *in vitro* screening methods and the synthesis of combinatorial compound libraries have further advanced the field. Structure-based drug discovery is shown to improve and shorten the drug development process (Blundell, 1996; Nicola & Abagyan, 2009; Paul *et al.*, 2010), accelerating lead optimisation, and therefore creating a therapeutic demand for protein structures.

As up to 70% of currently successful drug targets are membrane proteins, many structural genomics networks (Lundstrom, 2006) were established to exploit membrane protein targets further, by improving the technologies for solving the bottlenecks in membrane protein structural determination. However, to date there are only 641 (245 unique) structures of membrane proteins in comparison to >60,000 structures of water-soluble proteins in the Protein Data Bank (www.pdb.org; Berman *et al.*, 2000). The majority of these high resolution structures are bacterial proteins, with very few structures of eukaryotic proteins. The number of membrane protein structures has been modest in comparison to soluble proteins, owing to the three major bottlenecks, which are expression, purification and structure determination (Figure 6.1). However, the results presented in this thesis demonstrated that many of these bottlenecks have been overcome successfully, at least for the proteins included in this study, resulting in the successful production of milligram quantities of many of these membrane proteins, which lead to characterisation and crystallisation studies. These methods could now be applied to further membrane proteins.

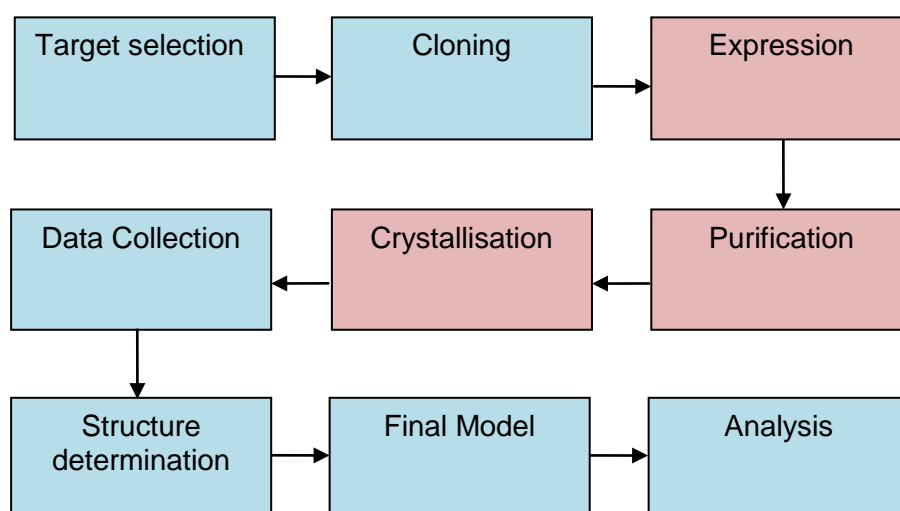


Figure 6.1 Structure determination pathway. Each box represents a stage in the structure determination pathway of membrane protein. The pink boxes are the rate-limiting steps.

This study has also increased our understanding of the behaviour and properties of the membrane proteins that were investigated, and these will be vital for future studies. Elucidation of their structures will contribute important information on their molecular mechanisms and thus will aid the development of inhibitors for diseases or bacterial infections related to these proteins. The NCS1 family of transporters forms part of the essential salvage pathway of microorganisms. They are also commercially important and structurally similar to bacterial homologues of neurotransmitter transporters. With the recent structures of three states of the transport mechanism of a NCS1 protein (Mhp1) being elucidated (Weyand *et al.*, 2008; Shimamura *et al.*, 2010), further understanding to complete the transport cycle would be of great importance for the design of antidepressants. In the case of histidine kinases of *E. faecalis*, understanding of their mechanisms are of medical importance because of their involvement in multiple aspects of bacterial regulation and their absence in animals make them ideal target for development of new classes of antimicrobial drug, and of course for treatment of enterococcal infections, the major cause of hospital acquired infection.

6.3. Limitations and future perspectives

Based on the work described in this thesis, a number of future studies can be proposed. These studies would endeavour to overcome some of the limitations described herein to improve the success rate of the genomic approach, as well as to improve our understanding of membrane proteins further.

6.3.1. Future studies related to expression and purification

Isolation of genes encoding the membrane proteins of interest was successful for the majority of the chosen targets. All histidine kinases genes were successfully isolated from *E. faecalis*, and most NCS1 targets were cloned. For NCS1 genes from high %GC organisms, such as *Streptomyces* and *Pseudomonas*, addition of DMSO and lowering of annealing temperature were required to improve the yields of PCR-amplified gene products. Amplification of *Streptomyces* genes, for example SCO0572 and SCO7500, produced smeared bands on the agarose gel (Figure 3.3), possibly resulting from low specific binding of the primers which is affected by the high %GC content of the genes.

High-level production of some of the His₆-tagged membrane proteins in *E. coli* was unsuccessful, including the NCS1 transporters PA0438 and PPA0619, and histidine kinases EF2219 and EF1335. Factors such as growth media, pH, temperature, induction time, expression host strains, induction systems (such as autoinduction; Grabski *et al.*, 2005), expression vector, could therefore be tested to attempt optimising protein expression levels. To assess the quality of expressed membrane proteins, the method described by Geertsma *et al.* (2008) could be applied; this study used a fused green fluorescent protein as indication of successful protein folding. Since GFP only fluoresces when it is correctly folded and active (in which case the fused recombinant protein is also assumed to be likewise correctly folded), this system could be used to distinguish well folded (and potentially active) proteins from aggregated proteins.

All but one (Pden0678) of the NCS1 targets of this study possessed C-termini that were predicted to be located in the cytoplasm. However, any protein (including Pden0678), possessing a His₆-tag attached to a C-terminus that is located to the periplasm of *E. coli* will be expected to fail expression trials (Rahman *et al.*, 2007). It is unlikely that the His₆-tag would be able to cross the membrane to take up its topological position and therefore presumably the protein would fail to integrate correctly into the membrane. Therefore, to study Pden0678, the gene encoding this transporter should be incorporated into a vector that permits production of a Strep-tag instead of His₆-tag.

6.3.2. Future studies related to determination of function

Further screening is required in order to identify substrates for some of the NCS1 proteins studied here. This will require a wider selection of substrates to be screened and completing the gaps in Table 3.4. For the histidine kinases used in this study, signal screening by *in vitro* assays has now been shown to be sufficiently sensitive and specific and therefore this approach can now be used with a wide range of candidate signals, including those signal(s) known to be perceived by homologues of these kinases in other bacterial species (Hancock & Perego, 2002). Also, a reduced concentration of candidate signal in the assays should be used in the initial screenings. Although use of a 2-fold molar excess of candidate signals were used in the more detailed experiments included in this study, the initial screenings using 10 mM signal in the presence of 60 pmoles of sensor protein in 10 ul assays represents a ratio of signal:protein of approx 1800:1. Screening for the signals perceived by the histidine kinases should therefore involve use of lower concentrations of candidate signals. In any

cases where a range of different concentrations is to be trialled in addition to a wide range of different compounds, then use of ^{33}P -ATP in these assays could be expensive. Methods for detection of phosphorylated proteins by other means could be used, which were developed and described by Kinoshita *et al.* (2006). The methods involve visualisation of phosphorylated protein using alkoxide-bridged dinuclear metal (Zn^{2+} or Mn^{2+}) complexes as novel phosphate-binding tag (Phos-tag) molecules (Kinoshita *et al.*, 2006; Kinoshita *et al.*, 2009).

Substrates for five of the NCS1 transporters were identified in this research (Table 3.5). Mutagenesis experiments could now be performed to identify residues important for their substrate or cation transport. This could also give us clues to which amino acid could lock the transporter in a particular form for crystallisation studies. For histidine kinases, the same could be performed.

6.3.3. Future studies related to structural determination

Mutagenesis studies may turn out to be important, as mentioned above, for stabilising some of the proteins sufficiently for successful future crystallisation, as it could overcome the dynamic problems which many transporters face for crystallisation. This is probably the same for the histidine kinases, as mutants may reduce or completely abolish the flexibility of the kinase for crystallisation.

A major factor in facilitating crystallisation is improvement of the thermal stability of the membrane protein. Warne *et al.* (2008) were able to construct mutants that gave rise to more stable β_1 -adrenergic G-protein coupled receptors suitable for crystallisation, and these mutant proteins were successfully used to obtain crystals and a crystal structure. Therefore, it is important to test the thermostability of the purified proteins prior to crystallisation. This could be done by recording the thermal melt of the protein by circular dichroism as well as the method described in Postis *et al.* (2008).

All proteins purified in this research could now proceed to size exclusion chromatography to obtain monodispersed proteins for crystallisation trials. The buffer components including pH, salts and detergent could be optimised to stabilise the protein. Screens such as MemGoldTM could also be tested as well as the ones mentioned in the research.

6.4 Concluding remarks

It costs pharmaceutical industries US \$ 1.3 – 1.8 billion for development of a new drug (Collier, 2009; Paul *et al.*, 2010), and methods are being developed to shorten and reduce the cost of the drug developmental process. Many high-throughput *in vitro* screening methods and structure-based drug design protocols have played important roles in optimising these processes. Given the high percentage of drug targets being membrane proteins, development of structural genomic approaches to tackle issues relating to structural determination of membrane proteins (expression, purification and crystallisation) are important. This study represents one of the successful ones. Successful approaches for the study of two different families of membrane proteins are shown, including descriptions for the production of milligram quantities (~10 mg batches) of some of these important membrane proteins for their future structure determinations.

References

- Aakra, A., Vebø, H., Snipen, L., Hirt, H., Aastveit, A., Kapur, V., Dunny, G., Murray, B. & Nes, I.F. (2005).** Transcriptional response of *Enterococcus faecalis* V583 to Erythromycin. *Antimicrobial Agents and Chemotherapy*, 2246-2259.
- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R. & Iwata, S. (2003).** Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*, 301(5633), 610-615.
- Abramson, J. & Wright, E.M. (2009).** Structure and function of Na⁺-symporters with inverted repeats. *Current Opinion in Structural Biology*, 19, 425-432.
- Ahmed, S. & Booth, I.R. (1983).** The use of valinomycin, nigericin and trichlorocarbanilide in control of the protonmotive force in *Escherichia coli* cells. *Journal of Biochemistry*, 212, 105-112.
- Ahrne, S., Nobaek, S., Jeppsson, B., Adlerberth, I., Wold, A.E. & Molin, G. (1998).** The normal *Lactobacillus* flora of healthy human rectal and oral mucosa. *Journal of Applied Microbiology*, 85, 88-94.
- Albanesi, D., Martin, M., Trajtenberg, F., Mansilla, M.C., Haouz, A., Alzari, P.M., de Mendoza, D. & Buschiazzo, A. (2009).** Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proceedings in National Academy of Science USA*, 106, 16185-16190.
- Albin, R., Weber, R. & Silverman, P.M. (1986).** The Cpx protein of *Escherichia coli* K12. Immunologic detection of the chromosomal *cpxA* gene product. *Journal of Biological Chemistry*, 261, 4698-4705.
- Allesen-Holm, M., Barken, K.B., Yang, L., Klausen, M., Webb, J.S., Kjelleberg, S., Molin, S., Givskov, M. & Tolker-Nielsen, T. (2006).** A characterisation of DNA release in *Pseudomonas aeruginosa* cultures and biofilms. *Molecular Microbiology*, 59, 1114-1128.
- Altenbuchner, J., Siemann-Herzberg, M. & Syldatk, C. (2001).** Hydantoinases and related enzymes as biocatalysts for the synthesis of unnatural chiral amino acids. *Current Opinions in Biotechnology*, 12, 559-563.
- Altendorf, K.H. & Staehelin, L.A. (1974).** Orientation of membrane vesicles from *Escherichia coli* as detected by freeze-cleave electron microscopy. *Journal of Bacteriology*, 117(2), 888-899.
- Andersen, P.S., Frees, D., Fast, R. & Mygind, B. (1995).** Uracil uptake in *Escherichia coli* K-12: Isolation of *uraA* mutants and cloning of the gene. *Journal of Bacteriology*, 2008-2013.
- Anderson, L., Kilstrup, M. & Neuhard, J. (1989).** Pyrimidine, purine and nitrogen control of cytosine deaminase synthesis in *Escherichia coli* K 12. Involvement of the

glnLG and *purR* genes in the regulation of *coda* expression. Archives in Microbiology, 152, 115-118.

Aravind, L. & Ponting, C.P. (1999). The cytoplasmic linker domain of receptor histidine kinase and methyl-accepting protein is common to many prokaryotic signalling proteins. FEMS Microbiology Letters, 176, 111-116.

Archibald, A.R., Hancock, I.C. & Harwood, C.R. (1993). Cell wall structure, synthesis, and turnover. In *Bacillus subtilis* and other gram-positive bacteria (Sonenshein, A.L., Hoch, J.A. & Losick, R. Ed.), pp 381-410, American society for Microbiology, Washington, DC.

Attwood, P.V., Piggott, M.J., Zu, X.L. (2007). Focus on phosphohistidine. Amino Acids, 32, 145-156.

Autret, N., Raynaud, C., Dubail, I., Berche, P. & Charbit, A. (2003). Identification of the *agr* locus of *Listeria monocytogenes*: role in bacterial virulence. Infection and Immunity, 71, 4463-4471.

Baker, D.A. & Kelly, J.M. (2004). Structure, function and evolution of microbial adenylyl and guanylyl cyclases. Molecular Microbiology, 52, 1229-1242.

Barakat, M., Ortet, P., Jourlin-Castelli, C., Ansaldi, M., Méjean, V. & Whitworth, D.E. (2009). P2CS: a two-component system resource for prokaryotic signal transduction research. BMC Genomics, 10: 315.

Baraquet, C., Theraulaz, L., Guiral, M., Lafitte, D., Mejean, V. & Jourlin-Castelli, C. (2006). TorT: a member of a new periplasmic binding protein family, triggers induction of the tor respiratory system upon trimethylamine N-oxide electron-acceptor binding in *Escherchia coli*. Journal of Biological Chemistry, 281, 38189-38199.

Barford, D. (2004). The role of cysteine residues as redox-sensitive regulatory switches. Current opinion in structural biology, 14, 679-686.

Bassler, B.L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. Current Opinions in Microbiology, 2, 582-587.

Bauer, C.E., Elsen, S. & Bird, T.H. (1999). Mechanism for redox control of gene expression. Annual Review in Microbiology, 53, 495-523.

Belenky, P.A., Moga, T.G. & Brenner, C. (2008). *Saccharomyces cerevisiae* YOR071C encodes the high affinity nicotinamide riboside transporter Nrt1. The Journal of Biological Chemistry, 283, 8075-8079.

Brandes, N., Schmitt, S. & Jakob, U. (2009). Thiol-based redox switches in eukaryotic proteins. Antioxid Redox singal, 11(5), 997-1014.

Bayles, K.W. (2007). The biological role of death and lysis in biofilm development. Nature Reviews in Microbiology, 5, 721-726.

- Beier, L., Nygaard, P., Jarmer, H. & Saxild, H.H. (2002).** Transcription analysis of the *Bacillus subtilis* PucR regulon and identification of a *cis*-acting sequence required for PucR-regulated expression of genes involved in purine catabolism. *Journal of Bacteriology*, 184(12), 3232-3241.
- Bekker, M., Alexeeva, S., Laan, W., Sawers, G., Teixeira de Mattos, J. & Hellingwerf, K. (2010).** The ArcBA two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool. *Journal of Bacteriology*, 192, 746-754.
- Bender, R.A. (1991).** The role of the NAC protein in the nitrogen regulation of *Klebsiella aerogenes*. *Molecular Microbiology*, 5, 2575-2580.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000).** The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235-242.
- Bettaney, K. (2008).** Characterisation of membrane transport proteins of bacteria – from gene to crystal. Ph.D. Thesis, University of Leeds, Leeds, UK.
- Blobel, G. (2000).** Protein targeting (Nobel lecture). *ChemBiochem*, 1(2), 86-102.
- Bloch, J.C., Sychrova, H., Souciet, J.L., Jund, R. & Chevallier, M.R. (1992).** Determination of a specific region of the purine-cytosine permease involved in the recognition of its substrates. *Molecular Microbiology*, 6(20), 2989-2997.
- Blundell, T.L. (1996).** Structure-based drug design. *Nature*, 384S, 23-26.
- Bogel, G., Schrempf, H. & Lucana, D.O. (2009).** The heme-binding protein HbpS regulates the activity of the *Streptomyces reticuli* iron-sensing histidine kinase SenS in a redox-dependent manner. *Amino acids*, 37, 681-691.
- Bolotin, A., Wincher, P., Mauger, S., Malarme, K., Weissenbach, J., Ehrlich, S.D. & Sorokin, A. (2001).** The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Research*, 11, 731-753.
- Bommarius, A.S., Schwarm, M. & Drauz, K. (1998).** Biocatalysis to amino acid-based chiral pharmaceuticals – examples and perspectives. *Journal of Molecular Catalysis B Enzymatic*, 5, 1-11.
- Bongaerts, G.P.A., Sin, I.L., Peters, A.L.J. & Vogels, G.D. (1977).** Purine degradation in *Pseudomonas aeruginosa* and *Pseudomonas testosteroni*. *Biochimica et Biophysica Acta*, 499, 111-118.
- Boşgelmez-Tinaz, G. (2003).** Quorum sensing in Gram-negative bacteria. *Turkish Journal of Biology*, 27, 85-93.
- Bourgogne, A., Hilsenbeck, S.G., Dunny, G.M & Murray, B.E. (2006).** Comparison of OG1RF and an isogenic *fsrB* detection mutant by transcriptional analysis: the Fsr system of *Enterococcus faecalis* is more than the activator of gelatinase and serine protease. *Journal of Bacteriology*, 188, 2875-2884.

- Boyer, P.D., DeLuca, M., Ebner, K.E., Hultquist, D.E. & Peter, J.B. (1962).** Identification of phosphohistidine in digest from a probable intermediate of oxidative phosphorylation. *Journal of Biological Chemistry*, 237, 2206.
- Boyle-Vavra, S., Yin, S. & Daum, R.S. (2006).** The *VraS/VraR* two-component regulatory system required for oxacillin resistance in community-acquired methicillin-resistant *Staphylococcus aureus*. *FEMS Microbiology Letters*, 262(2), 163-171.
- Brèthes, D., Chirio, M.C., Napias, C., Chevallier, M.R., Lavie, J.L. & Chevallier, J. (1992).** *In vivo* and *in vitro* studies of the purine-cytosine permease of *Saccharomyces cerevisiae*. Functional analysis of a mutant with an altered apparent transport constant of uptake. *European Journal of Biochemistry / FEBS*, 204(2), 699-704.
- Brook, B.E. & Buchanan, S.K. (2008).** Signalling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Biochimica Biophysica Acta*, 1778(9), 1930-1945.
- Brunskill, E.W. & Bayles, K.W. (1996a).** Identification of *LytSR*-regulated genes from *Staphylococcus aureus*. *Journal of Bacteriology*, 178, 5810-5812.
- Brunskill, E.W. & Bayles, K.W. (1996b).** Identification and molecular characterisation of a putative regulatory locus that affects autolysis in *Staphylococcus aureus*. *Journal of Bacteriology*, 178, 611-618.
- Bürkle, L., Cedzich, A., Dopke, C., Stransky, H., Okumoto, S., Gillissen, B., Kuhn, C. & Frommer, W.B. (2003).** Transport of cytokinins mediated by purine transporter of the PUP family expressed in phloem hydathodes and pollen of *Arabidopsis*. *Plant Journal*, 34, 13-26.
- Carpenter, E.P., Beis, K., Cameron, A. & Iwata, S. (2008).** Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, 18(5), 581-586.
- Castañó, I., Bastarrachea, F. & Covarrubias, A.A. (1988).** *gltBDF* operon of *Escherichia coli*. *Journal of Bacteriology*, 170(2), 821-827.
- Castañó, I., Flores, N., Valle, F., Covarrubias, A.A. & Bolivar, F. (1992).** *gltF*, a member of the *gltBDF* operon of *Escherichia coli*, is involved in nitrogen-regulated gene expression. *Molecular Microbiology*, 6(18), 2733-2741.
- Chandramohan, L., Ahn, J., Weaver, K.E. & Bayles, K.W. (2009).** An overlap between the control of programmed cell death in *Bacillus anthracis* and sporulation. *Journal of Bacteriology*, 191(13), 4103-4110.
- Cheung, H. & Freese, E. (1985).** Monovalent cations enable cell wall turnover of the turnover-deficient *lyt-15* mutant of *Bacillus subtilis*. *Journal of Bacteriology*, 161, 1222-1225.
- Cheung, J., Bingman, C.A., Reingold, M., Hendrickson, W.A. & Waldburger, C.D. (2008).** Crystal structure of a functional dimer of the PhoQ sensor domain. *Journal of Biological Chemistry*, 283(20), 13762-13770.

- Cheung, J. & Hendrickson, W.A. (2008).** Crystal structures of C4-dicarboxylate ligand complexes with sensor domains of histidine kinases of DcuS and DctB. *Journal of Biological Chemistry*, 283, 30256-30265.
- Cheung, J. & Hendrickson, W.A. (2009).** Structural analysis of ligand stimulation of the histidine kinase NarX. *Structure*, 17, 190-201.
- Cheung, J.K., Awad, M.M., McGowan, S. & Rood, J.L. (2009).** Functional analysis of the VirSR phosphorelay from *Clostridium perfringens*. *PLoS ONE*, 4(6): e5849. doi:10.1371/Journal.pone.0005849
- Cheung, J., Le-Khac, M. & Hendrickson, W.A. (2009).** Crystal structure of a histidine kinase sensor domain with similarity to periplasmic binding proteins. *Proteins*, 77, 235-241.
- Cheung, J. & Hendrickson, W.A. (2010).** Sensor domains of two-component regulatory systems. *Current Opinion in Microbiology*, 13, 116-123.
- Chevallier, M.R., Jund, R. & Lacroute, F. (1975).** Characterisation of cytosine permeation in *Saccharomyces cerevisiae*. *Journal of Bacteriology*, 122(2), 629-641.
- Cho, U.S., Bader, M.W., Amaya, M.F., Daley, M.E., Klevit, R.E., Miller, S.L. & Xu, W.Q. (2006).** Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signalling. *Journal of Molecular Biology*, 356, 1193-1206.
- Cho, H.Y., Cho, H.J., Kim, Y.M., Oh, J.I. & Kang, B.S. (2009).** Structural insights into the heme-based redox sensing by DosS from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 284, 13057-13067.
- Christiansen, L.C., Schou, S., Nygaard, P. & Saxild, H.H. (1997).** Xanthine metabolism in *Bacillus subtilis*: characterisation of the *xpt-pbuX* operon and evidence for purine- and nitrogen-controlled expression of genes involved in xanthine salvage and catabolism. *Journal of Bacteriology*, 179, 2540-2550.
- Chung, C. T., Niemela, S. L. & Miller, R. H. (1989).** One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *The Proceedings of the National Academy of Science USA*, 86, 2172-2175.
- Claros, M.G. & von Heijne, G. (1994).** TopPred II: an improved software for membrane protein structure predictions. *CABIOS* 10, 685-686.
- Clausen, V.A., Bae, W., Throup, J., Burnham, M.K.R., Rosenberg, M. & Wallis, N.G. (2003).** Biochemical characterisation of the first essential two-component signal transduction system from *Staphylococcus aureus* and *Streptococcus pneumoniae*. *Journal of Molecular Microbiology and Biotechnology*, 5, 252-260.
- Collier, R. (2009).** Drug development cost estimates hard to swallow. *Canadian Medical Association Journal*, 180(3), 279-280.
- Comella, N. & Grossman, A.D. (2005).** Conservation of genes and processes controlled by the quorum response in bacteria: characterisation of genes controlled by

the quorum-sensing transcription factor ComA in *Bacillus subtilis*. *Molecular Microbiology*, 57, 1159-1174.

Comenge, Y., Quintiliani, R., Li, L., Dubost, L., Brouard, J.P., Hugonnet, J.E. & Arthur, M. (2003). The CroRS two-component regulatory system is required for intrinsic beta-lactam resistance in *Enterococcus faecalis*. *Journal of Bacteriology*, 185, 7184-7192.

Cotter, P.D., Emerson, N., Gahan, C.G. & Hill, C. (1999). Identification and disruption of *lisRK*, a genetic locus encoding a two-component signal transduction system involved in stress tolerance and virulence in *Listeria monocytogenes*. *Journal of Bacteriology*, 181, 6840-6843.

Cotter, P.A. & Stibitz, S. (2007). C-di-CMP-mediated regulation of virulence and biofilm formation. *Current Opinion in Microbiology*, 10, 17-23.

Craig, J.E., Zhang, Y., Gallagher, M.P. (1994). Cloning of the *nupC* gene of *Escherichia coli* encoding a nucleoside transport system, and identification of an adjacent insertion element, IS 186. *Molecular Microbiology*, 11(6), 1159-1168.

Crane, R.K., Miller, D. & Bihler, I. (1960). The restrictions on possible mechanism of intestinal transport of sugars. In: *Membrane Transport and Mechanism. Proceedings of a Symposium held in Prague, August 22-27, 1960.* Edited by Kleinzeller, A. & Kotyk, A. Czech Academy of Sciences, Prague, 1961, pp. 439-449. Model of cotransport on page 448.

Cruz-Ramos, H., Glaser, P., Wray, Jr. L.V. & Fisher, S.H. (1997). The *Bacillus subtilis ureABC* operon. *Journal of Bacteriology*, 179(10), 3371-3373.

Danielsen, S., Kilstrup, M., Barilla, K., Jochimsen, B. & Neuhard, J. (1992). Characterization of the *Escherichia coli codBA* operon encoding cytosine permease and cytosine deaminase. *Molecular Microbiology*, 6(10), 1335-1344.

Danielsen, S., Boyd, D. & Neuhard, J. (1995). Membrane topology analysis of the *Escherichia coli* cytosine permease. *Microbiology*, 141, 2905-2913.

Daruwala, R., Song, J., Koh, W.S., Rumsey, S.C. & Levine, M. (1999). Cloning and functional characterisation of the human sodium-dependent vitamin C transporters hSVCT1 and hSVCT2. *FEBS Letters*, 460, 480-484.

de Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *The Proceedings of the National Academy of Science USA*, 80, 21-25.

de Koning, H. & Diallinas, G. (2000). Nucleobase transporters (Review). *Molecular Membrane Biology*, 75, 75-94.

Detlefsen, D.J., Hill, S.E., Volk, K.J., Klohr, S.E., Tsunakawa, M., Furumai, T., Lin, P.F., Nishio, M., Kawano, K., Oki, T. & Lee, M.S. (1995). Siamycins I and II, new anti-HIV-1 peptides. II. Sequence analysis and structure determination of siamycin I. *Journal of Antibiotics*, 48, 1515-1517.

- Diallinas, G., Gorfinkiel, L., Arst, H.N., Jr., Ceccheto, G. & Scazzocchio, C. (1995).** Genetic and molecular characterisation of a gene encoding a wide specificity purine permease of *Aspergillus nidulans* reveals a novel family of transporters conserved in prokaryotes and eukaryotes. *The Journal of Biological Chemistry*, 270, 8610-8622.
- Doan, T., Servant, P., Tojo, S., Yamaguchi, H., Lerondel, G., Yoshida, K., Fujita, Y. & Aymerich, S. (2003).** The *Bacillus subtilis* ywkA gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate. *Microbiology*, 149, 2331-2343.
- Dobrovetsky, E., Lu, M.L., Andorn-Broza, R., Khutoreskaya, G., Bray, J.E., Savchenko, A., Arrowsmith, C.H., Edwards, A.M. & Koth, C.M. (2005).** High-throughput production of prokaryotic membrane proteins. *Journal of Structural and Functional Genomics*, 6, 33-50.
- Dohan, O., De la Vieja, A., Paroder, V., Riedel, C., Artani, M., Reed, M., Ginter, C.S. & Carrasco, N. (2003).** The sodium/iodide symporter (NIS): characterisation, regulation, and medical significance. *Endocrinology Reviews*, 24, 48-77.
- Doroshchuk, N.A., Gelfand, M.S. & Rodionov, D.A. (2006).** Regulation of nitrogen metabolism in gram-positive bacteria. *Molecular Biology*, 40(5), 829-836.
- Dubrac, S. & Msadek, T. (2004).** Identification of genes controlled by the essential YycG/YycF two-component system of *Staphylococcus aureus*. *Journal of Bacteriology*, 186, 1175-1181.
- Dufour, P., Jarraud, S., Vandenesch, F., Greenland, T., Novick, R.P., Bes, M., Etienne, J. & Lina, G. (2002).** High genetic variability of the *agr* locus in *Staphylococcus* species. *Journal of Bacteriology*, 1180-1186.
- Dunham, C.M., Dioum, E.M., Tuckerman, J.R., Gonzalez, G., Scott, W.G. & Gilles-Gonzalez, M.A. (2003).** A distal arginine in oxygen-sensing Heme-PAS domain is essential to ligand binding, signal transduction, and structure. *Biochemistry*, 42, 7701-7708.
- Dunman, P.M., Murphy, E., Haney, S., Palacios, D., Tucker-Kellogg, G., Wu, S., Brown, E.L., Zagursky, R.J., Shlaes, D. & Projan, S.J. (1990).** Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *Journal of Bacteriology*, 7341-7353.
- Dunny, G.M. & Leonard, B.A. (1997).** Cell-cell communication in gram-positive bacteria. *Annual Review in Microbiology*, 51, 527-564.
- Dutta, R., Qin, L. & Inouye, M. (1999).** Histidine kinases: a diversity of domain organisation. *Molecular Microbiology*, 34, 633-640.
- Dutzler, R., Campbell, E.B., Cadene, M, Chait, B.T. & MacKinnon, R. (2002).** X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature*, 514(6869), 287-294.

- Eberl, L. (1999).** *N*-acyl homoserinelactone-mediated gene regulation in Gram-negative bacteria. *Systems in Applied Microbiology*, 22, 493-506.
- Eckhart, W., Hutchinson, M. & Hunter, T. (1979).** An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell*, 18, 925-933.
- Edman, P. (1950).** Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica*, 4, 283-293.
- Emami, K., Topakas, E., Nagy, T., Henshaw, J., Jackson, K.A., Nelson, K.E., Mongodin, E.F., Murray, J.W., Lewis, R.J. & Gilbert, H.J. (2009).** Regulation of the xylan-degrading apparatus of *Cellvibrio japonicas* by a novel two-component system. *Journal of Biological Chemistry*, 284, 1086-1096.
- Engbrecht, J. & Silverman, M. (1984).** Identification of genes and gene products necessary for bacterial bioluminescence. *Proceedings in National Academy of Sciences USA*, 81, 4154-4158.
- Engel, K., Zhou, M. & Wang, J. (2004).** Identification and characterisation of a novel monoamine transporter in the human brain. *The Journal of Biological Chemistry*, 279, 50042-50049.
- Enjo, F., Nosaka, K., Ogata, M., Iwashima, A. & Nishimura, H. (1997).** Isolation and characterisation of a thiamine transport gene, THI10, from *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, 272, 191165-19170.
- Evers, S. & Courvalin, P. (1996).** Regulation of VanB-type vancomycin resistance gene expression by the VanS(B)-VanR(B) two-component regulatory system in *Enterococcus faecalis* V583. *Journal of Bacteriology*, 178, 1302-1309.
- Faaland, C.A., Race, J.E., Ricken, G., Warner, F.J., Williams, W.J. & Holtzman, E.J. (1998).** Molecular characterisation of two novel transporters from human and mouse kidney and from LLC-PK₁ cells reveals a novel family that is homologous to bacterial and *Aspergillus* nucleobase transporters. *Biochimica et Biophysica Acta*, 1442, 353-360.
- Fabret, C., Feher, V.A. & Hoch, J.A. (1999).** Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *Journal of Bacteriology*, 181, 1975-1983.
- Facey, S.J. & Kuhn, A. (2010).** Biogenesis of bacterial inner-membrane proteins. *Cell and Molecular Life Sciences*, DOI 10.1007/s00018-010-0303-0 (published online 3rd March 2010).
- Faham, S., Watanabe, A., Besserer, G.M., Cascio, D., Specht, A., Hirayama, B.A., Wright, E.M. & Abramson, J. (2008).** The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science*, 321(5890), 810-814.

- Fang, Y., Jayaram, H., Shane, T., Kolmakova-Partensky, L., Wu, F., Williams, C., Xiong, Y. & Miller, C. (2009).** Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature*, 460, 1040-1043.
- Federle, M.J., McIver, K.S. & Scott, J.R. (1999).** A response regulator that represses transcription of several virulence operons in the group A *Streptococcus*. *Journal of Bacteriology*, 181, 3649-3657.
- Ferreira, T., Brèthes, D., Pinson, B., Napias, C. & Chevallier, J. (1997).** Functional analysis of mutated purine-cytosine permease from *Saccharomyces cerevisiae*. A possible role of the hydrophilic segment 371-377 in the active carrier conformation. *The Journal of Biological Chemistry*, 272(15), 9697-9702.
- Ferretti, J.J., McShan, W.M., Ajdic, D., Savic, D.J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, N., Kenton, S., Lai, H.S., Lin, S.P., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. & McLaughlin, R. (2001).** Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proceedings in National Academy of Science USA*, 98, 4658-4663.
- Ferson, A.E., Wray, Jr. L.V. & Fisher, S.H. (1996).** Expression of the *Bacillus subtilis* *gabP* gene is regulated independently in response to nitrogen and amino acid availability. *Molecular Microbiology*, 22(4), 693-701.
- Fisher, S.H., Rohrer, K. & Ferson, A.E. (1996).** Role of CodY in regulation of the *Bacillus subtilis* *hut* operon. *Journal of Bacteriology*, 178, 3779-3784.
- Fisher, S.H. (1999).** Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la difference! *Molecular Microbiology*, 32, 223-232.
- Fisher, S.H. & Wray, Jr. L.V. (2002).** *Bacillus subtilis* 168 contains two differentially regulated genes encoding L-asparaginase. *Journal of Bacteriology*, 184(8), 2148-2154.
- Fuhrmann, M., Hausherr, A., Ferbitz, L., Schodl, T., Heitzer, M. & Hegemann, P. (2004).** Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Molecular Biology*, 55(6), 869-881.
- Fujii, T., Ingham, C., Nakayma, J., Beerthuyzen, M., Kunuki, R., Molenaar, D., Sturme, M., Vaughan, E., Kleerebezem, M. & de Vos, W. (2008).** Two homologous Agr-like quorum-sensing systems in cooperatively control adherence, cell morphology, and cell viability properties in *Lactobacillus plantarum* WCFS1. *Journal of Bacteriology*, 180, 7655-7665.
- Fukuchi, K., Sasahara, K., Asai, K., Kobayashi, K., Moriya, S. & Ogasawara, N. (2000).** The essential two-component regulatory system encoded by *ycyF* and *ycyG* modulates expression of the *ftsAZ* operon in *Bacillus subtilis*. *Microbiology*, 146, 1573-1583.
- Fuqua, C. & Greenberg, E.P. (1998).** Self perception in bacteria: quorum sensing with acylated homoserine lactones. *Current Opinions in Microbiology*, 1, 183-189.

- Fuqua, C., Parsek, M.R., Greenberg, E.P. (2001).** Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. *Annual Review in Genetics*, 35, 439-468.
- Gale, E.F. & Taylor, E.S. (1947).** Action of tyrocidin and some detergent substances in releasing amino acids from the internal environment of *Streptococcus faecalis*. *Journal of General Microbiology*, 1, 77-84.
- Galperin, M.Y. (2004).** Bacterial signal transduction network in a genomic perspective. *Environmental Microbiology*, 6, 552-567.
- Gama-Castro, S., Jiménez-Jacinto, V., Peralta-Gil, M., Santos-Zavaleta, A., Pènalzoza-Spinola, M.I., Contreras-Moreira, B., Segura-Salazar, J., Muniz-Rascado, L., Martínez-Flores, I., Salgado, H., Bonavides-Martínez, C., Abreu-Goodger, C., Rodríguez-Penagos, C., Miranda-Ríos, J., Morett, E., Merino, E., Huerta, A.M., Trevino-Quintanilla, L. & Collado-Vides, J. (2008).** RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. *Nucleic Acids Research*, 36(Database issue), D120-D124.
- Ganesan, A.T. & Lederberg, J. (1965).** A cell-membrane bound fraction of bacterial DNA. *Biochemical and Biophysical Research Communications*, 18, 824.
- Gao, R. & Stock, A.M. (2009).** Biological insights from structures of two-component proteins. *Annual Reviews in Microbiology*, 63, 133-154.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D. & Bairoch, A. (2003).** ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31, 3784-3788.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. & Bairoch, A. (2005).** Protein identification and analysis tools on the ExPASy server. In: *The proteomics protocol handbook* (Walker, J.M., Ed) Humana Press.
- Geertsma, E.R., Groeneveld, M., Slotboom, D.-J. & Poolman, B. (2008).** Quality control of overexpressed membrane proteins. *Proceedings in National Academy of Sciences*, 105(15), 5722-5727.
- Geisinger, E., George, E.A., Muir, T.W. & Novick, R.P. (2008).** Identification of ligand specificity determinants in AgrC, the *Staphylococcus aureus* quorum-sensing receptor. *The Journal of Biological Chemistry*, 283(14), 8930-8938.
- Georgellis, D., Kwon, O. & Lin, E.C.C. (2001).** Quinones as the redox signal for the Arc two-component system of bacteria. *Science*, 292(5525), 2314-2316.
- Ghim, S.-Y. & Neuhard, J. (1994).** The pyrimidine biosynthesis operon of the thermophile *Bacillus caldolyticus* includes genes for uracil phosphoribosyltransferase and uracil permease. *Journal of Bacteriology*, 176, 3698-3707.

- Gilles-Gonzalez, M.A., Caceres, A.L., Sousa, E.H.S., Tomchick, D.R., Brautigam, C.A., Gonzalez, C. & Machius, M. (2006).** A proximal arginine R206 participates in switching of the *Bradyrhizobium japonicum* FixL oxygen sensor. *Journal of Molecular Biology*, 360, 80-89.
- Gillissen, B., Bürkle, L., André, B., Kühn, C., Rentsch, D., Brandl, B. & Frommer, W.B. (2000).** A new family of high-affinity transporters for adenine, cytosine, and purine derivatives in Arabidopsis. *Plant Cell*, 12, 291-300.
- Giridhara Upadhyaya, P.M., Ravikumar, K.L. & Umapathy, B.L. (2009).** Review of virulence factors of Enterococcus: An emerging nosocomial pathogen. *Indian Journal of Medical Microbiology*, 27(4), 301-305.
- Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fsihi, H., Portillo, F.G., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vincente, J.M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez Diaz, J.C., Purcell, R., Rimmel, B., Rose, M., Schleuter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehland, J. & Cossart, P. (2001).** Comparative genomics of *Listeria* species. *Science*, 294, 849-852.
- Gong, W., Hao, B., Mansy, S.S., Gonzalez, G., Gilles-Gonzalez, M.A. & Chan, M.K. (1998).** Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proceedings in National Academy of Science USA*, 95, 15177-15182.
- Gong, W.M., Hao, B. & Chan, M.K. (2000).** New mechanistic insights from structural studies of the oxygen-sensing domain of *Bradyrhizobium japonicum* FixL. *Biochemistry*, 39, 3955-3962.
- Gordeliy, V.I., Labahn, J., Moukhametzianov, R., Efremov, R., Granzin, J., Schlesinger, R., Buldt, G., Savopol, T., Scheidig, A.J., Klare, J.P. & Engelhard, M. (2002).** *Nature*, 419, 484-487.
- Gorfinkiel, L., Diallinas, G. & Scazzocchio, C. (1993).** Sequence and regulation of the *uapA* gene encoding a uric acid-xanthine permease in the fungus *Aspergillus nidulans*. *The Journal of Biological Chemistry*, 268, 23376-23381.
- Goudela, S., Tsilivi, H. & Diallinas, G. (2006).** Comparative kinetic analysis of AzgA and Fcy21p, prototypes of the two major fungal hypoxanthine-adenine-guanine transporter families. *Molecular Membrane Biology*, 23, 291-303.
- Grabski, A., Mehler, M. & Drott, D. (2005).** The overnight express autoinduction system: high-density cell growth and protein expression while you sleep. *Nature methods*, 2(3), 223-235.
- Gram, H.C. (1884).** Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte der Medizin*, 2, 185-189.

- Gray, K.M. (1997).** Intercellular communication and group behavior in bacteria. *Trends in Microbiology*, 5, 184-188.
- Grebe, T.W. & Stock, J.B. (1999).** The histidine kinase superfamily. *Advances in Microbial Physiology*, 41, 139-227.
- Green, J. & Paget, M.S. (2004).** Bacterial redox sensors. *Nature Reviews in Microbiology*, 2, 954-967.
- Griffiths, M., Yao, S.Y.M., Abidi, F., Phillips, S.E.V., Cass, C.E., Young, J.D. & Baldwin, S.A. (1997).** Molecular cloning and characterisation of a nitrobenzylthioinosine-insensitive (*ei*) equilibrative nucleoside transporter from human placenta. *Journal of Biochemistry*, 328, 739-743.
- Grisshammer, R. & Tate, C.G. (1995).** Overexpression of integral membrane proteins for structural studies. *Quarterly Reviews of Biophysics*, 28(3), 315-422.
- Groicher, K.H., Firek, B.A., Fujimoto, D.F. & Bayles, K.W. (2000).** The staphylococcus aureus *lrgAB* operon modulates murein hydrolase activity and penicillin tolerance. *Journal of Bacteriology*, 182, 1794-1801.
- Grosjean, H. & Fier, W. (1982).** Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene*, 18(3), 199-209.
- Gründling, A., Bläsi, U. & Young, R.Y. (2000).** Genetic and biochemical analysis of dimer and oligomer interaction of the λ S holin. *Journal of Bacteriology*, 182, 6082-6090.
- Guan, L., Mirza, O., Verner, G., Iwata, S. & Kaback, H.R. (2007).** Structural determination of wild-type lactose permease. *Proceedings of the National Academy of Sciences of the USA*, 104(39), 15294-15298.
- Gutmann, D.A.P., Mizohata, E., Newstead, S., Ferrandon, S., Henderson, P.J.F., Van Veen, H.W. & Byrne, B. (2007).** A high-throughput method for membrane protein solubility screening: The ultracentrifugation dispersity sedimentation assay. *Protein Science*, 16, 1422-1428.
- Hancock, L.E. & Perego, M. (2002).** Two-component signal transduction in *Enterococcus faecalis*. *Journal of Bacteriology*, 5819-5825.
- Hancock, L.E. & Perego, M. (2004a).** The *Enterococcus faecalis* *fsr* two-component system controls biofilm development through production of gelatinase. *Journal of Bacteriology*, 186, 5629-5639.
- Hancock, L.E. & Perego, M. (2004b).** Systemic inactivation and phenotypic characterisation of two-component signal transduction systems of *Enterococcus faecalis* V583. *Journal of Bacteriology*, 186, 7951-7958.

- Hansen, M.E., Wangari, R., Hansen, E.B., Mijakovic, I. & Jensen, P.R. (2009).** Engineering of *Bacillus subtilis* 168 for increased nisin resistance. *Applied and environmental microbiology*, 75(21), 6688-6695.
- Hao, B., Isaza, C., Arndt, J., Soltis, M. & Chan, M.L. (2002).** Structure-based mechanism of O₂-sensing and ligand discrimination by the FixL heme domain of *Bradyrhizobium japonicum*. *Biochemistry*, 41, 12952-12958.
- Hardman, A.M., Steward, G.S. & Williams, P. (1998).** Quorum sensing and the cell-cell communication dependent regulation of gene expression in pathogenic and non-pathogenic bacteria. *Antonie Van Leeuwenhoek*, 74, 199-210.
- Hayaish, O. & Kornberg, A. (1952).** Metabolism of cytosine, thymine, uracil, and barbituric acid by bacterial enzymes. *The Journal of Biological Chemistry*, 197(2), 717-732.
- Hebert, S.C., Mount, D.B. & Gamba, G. (2004).** Molecular physiology of cation-coupled Cl⁻ cotransport: the SLC12 family. *Pflügers Archiv: European Journal of Physiology*, 447, 580-593.
- Heiman, M. (2000).** Finding restriction sites (Webcutter). *Genomic Biology* 1, reports2048.
- Hellingwerf, K.J. & Konings, W.N. (1985).** The energy flow in bacteria: the main free energy intermediates and their regulatory role. *Advances in Microbial Physiology*, 26, 125-154.
- Hellingwerf, K.J., Crielaard, W.C., Joost Teixeira de Mattos, M., Hoff, W.D., Kort, R., Verhamme, D.T. & Avignone-Rossa, C. (1998).** Current topics in signal transduction in bacteria. *Antonie Van Leeuwenhoek*, 74, 211-227.
- Henderson, P.J.F. & Macpherson, A.J.S. (1986).** Assay, genetics, proteins and reconstitution of proton-linked galactose, arabinose and xylose transport systems of *Escherichia coli*. *Methods in Enzymology*, 125, 387-429.
- Hentzer, M. & Givskov, M. (2003).** Pharmacological inhibition of quorum sensing for the treatment of chronic bacterial infections. *Journal of Clinical Investigation*, 112(9), 1300-1307.
- Hess, J.F., Bourret, R.B. & Simon, M.I. (1988).** Histidine phosphorylation and phosphoryl group transfer in bacterial chemotaxis. *Nature*, 336, 139-143.
- Hess, J.F., Oosawa, K., Kaplan, N. & Simon, M.I. (1988).** Phosphorylation of three proteins in the signalling pathway of bacterial chemotaxis. *Cell*, 53, 79-87.
- Hirayama, B.A., Loo, D.D.F., Diez-Sampedro, A., Leung, D.W., Meinild, A.K., Lai-Bing, M., Turk, E. & Wright, E.M. (2007).** Sodium-dependent reorganisation of the sugar-binding site of SGLT1. *Biochemistry*, 46(13), 13391-13406.

- Ho, Y.S., Burden, L.M. & Hurley, J.H. (2000).** Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *Journal of EMBO*, 19, 5288-5299.
- Hoch, J.A. & Silhavy, T.J. (1995).** Two-component signal transduction. Washing, DC: ASM Press.
- Hockney, R.C. (1994).** Recent developments in heterologous protein production in *Escherichia coli*. *Trends in Biotechnology*, 12, 456-463.
- Hoitje, J.V. & Turmanen, E.I. (1991).** The murein hydrolases of *Escherichia coli*: properties, functions and impact on the course of infections in vivo. *Journal of General Microbiology*, 137, 441-454.
- Howell, A., Dubrac, S., Andersen, K.K., Noone, D., Fert, J., Msadek, T. & Devine, K. (2003).** Genes controlled by the essential YycG/YycF two-component system *Bacillus subtilis* revealed through a novel hybrid regulator approach. *Molecular Microbiology*, 49(6), 1639-1655.
- Hoyle, C. J. (2000).** Recombinant expression, purification and characterisation of bacterial multidrug efflux proteins. Ph.D. Thesis. University of Leeds.
- Huang, Y., Lemieux, M.J., Song, J., Auer, M. & Wang, D.-N. (2003).** Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science*, 301(5633), 616-620.
- Hulett, F.M. (1996).** The signal-transducing network for Pho regulation in *Bacillus subtilis*. *Molecular Microbiology*, 19, 933-939.
- Hulko, M., Berndt, F., Gruber, M., Linder, J.U., Truffault, V., Schultz, A., Martin, J., Schultz, J.E., Lupas, A.N. & Coles, M. (2006).** The HAMP domain structure implies helix rotation in transmembrane signalling. *Cell*, 126, 929-940.
- Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E. & Michel, H. (2005).** Structure of a Na⁺/H⁺ antiporter and insights into mechanism of action and regulation by pH. *Nature*, 435(7046), 1197-1202.
- Inoue, H., Nojima, H. & Okayama, H. (1990).** High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, 96, 23-28.
- Ioanoviciu, A., Yukl, E.T., Moenne-Loccoz, P., Ortiz de Montellano, P.R. (2007).** DevS, a heme-containing two-component oxygen sensor of *Mycobacterium tuberculosis*. *Biochemistry*, 46, 4250-4260.
- Ishii, T., Yoshida, K., Terai, G., Fujita, Y. & Nakai, K. (2001).** DBTBS: a database of *Bacillus subtilis* promoters and transcription factors. *Nucleic Acids Research*, 29, 278-280.
- Jansen, A., Tuerck, M., Szekat, C., nagel, M., Clever, I. & Bierbaum, G. (2007).** Role of insertion elements and yycFG in the development of decreased susceptibility to

- vancomycin in *Staphylococcus aureus*. International Journal of Medical Microbiology, 297, 205-215.
- Jardetzky, O. (1966).** Simple allosteric model for membrane pumps. Nature, 221, 969-970.
- Jarraud, S., Lyon, G.J., Figueiredo, A.M., Lina, G., Vandenesch, F. & Etienne, J. (2000).** Exfoliatin-producing strains define a fourth *agr* specificity group in *Staphylococcus aureus*. Journal of Bacteriology, 182, 6517-6522.
- Jensen, R.D., Winzer, K., Clarke, S.R., Chan, W.C. & Williams, P. (2008).** Differential recognition of *Staphylococcus aureus* quorum-sensing signals depends on both extracellular loops 1 and 2 of the transmembrane sensor AgrC. Journal of Molecular Biology, 381, 300-309.
- Jett, B.D., Huycke, M.M. & Gilmore, M.S. (1994).** Virulence of enterococci. Clinical Microbiology Review, 7, 462-478.
- Ji, G., Beavis, R.C. & Novick, R.P. (1995).** Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. Proceedings in National Academy of Science USA, 92, 12055-12059.
- Ji, G., Beavis, R. & Novick, R.P. (1997).** Bacterial interference caused by autoinducing peptide variants. Science, 276(5321), 2027-2030.
- Johansen, L.E., Nygaard, P., Lassen, C., Agero, Y. & Sazild, H.H. (2003).** Definition of a second *Bacillus subtilis pur* regulon comprising the *pur* and *xpt-pbuX* operons plus *pbuG*, *nupG* (*yxjA*), and *pbuE* (*ydhL*). Journal of Bacteriology, 185, 5200-5209.
- Jones, M.E., Draghi, D.C., Thornsberry, C., Karlowsky, J.A., Sahm, D.F. & Wenzel, R.P. (2004).** Emerging resistance among bacterial pathogens in the intensive care unit – a European and North American Surveillance study (2000-2002). Annual Clinical Microbiology and Antimicrobial, 3, 14.
- Judge, R.A., Swift, K. & González, C. (2005).** An ultraviolet fluorescence-based method for identifying and distinguishing protein crystals. Acta Crystallography Biological Crystallography, D61, 60-61.
- Kallipolitis, B.H. & Ingmer, H. (2001).** *Listeria monocytogenes* response regulators important for stress tolerance and pathogenesis. FEMS Microbiology Lett. 204, 111-115.
- Kanehisa, M & Goto, S. (2000).** KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 28(1), 27-30.
- Kanehisa, M., Susumu, G., Furumichi, M., Tanebe, M. & Hirakawa, M. (2010).** KEGG for presentation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Research, 38, D355-D360.

- Kantardjieff, K.A. & Rupp, B. (2004).** Protein isoelectric point as a predictor for increased crystalliation screening efficiency. *Bioinformatics*, 20(14), 2162-2168.
- Katzen, F. & Beckwith, J. (2003).** Role and location of the unusual redox-active cysteines in the hydrophobic domain of the transmembrane electron transporter DsbD. *PNAS*, 100(18), 10471-10476.
- Key, J. & Moffat, K. (2005).** Crystal structures of deoxy and CO-bound bjFixLH reveal details of ligand recognition and signalling. *Biochemistry*, 44, 4627-4635.
- Key, J., Srajer, V., Pahl, R. & Moffat, K. (2007).** Time-revealed crystallographic studies of the heme domain of the oxygen sensor FixL: structural dynamics of ligand rebinding and their relation to signal transduction. *Biochemistry*, 46, 4706-4715.
- Keyes, M.H., Gray, D.N., Kreh, K.E. & Sanders, C.R. (2003).** Solubilising detergents for membrane proteins. In *Methods and results in crystallisation of membrane proteins – Part II Principles and techniques in membrane protein crystallisation* (Iwata, S. Ed.), pp 15-33. International University Line, USA.
- Khmel, I.A. & Metlitskaya, A.Z. (2006).** Quorum sensing regulation of gene expression: a promising target for drugs against bacterial pathogenicity. *Molecular Biology*, 40(2), 169-182.
- Kilstrup, M., Meng, L.M., Neuhard, J. & Nygaard, P. (1989).** Genetic evidence for a repressor of synthesis of cytosine deaminase and purine biosynthesis enzymes in *Escherichia coli*. *Journal of Bacteriology*, 171, 2124-2127.
- Kim, S. & West, T.P. (1991).** Pyrimidine catabolism in *Pseudomonas aeruginosa*. *FEMS Microbiology Letters*, 61(2-3), 175-179.
- Kim, D. & Forst, S. (2001).** Genomic analysis of the histidine kinase family in bacteria and archaea. *Microbiology*, 147, 1197-1212.
- Kim, B.H. & Gadd, G.M. (2008).** Composition and structure of prokaryotic cells. In *Bacterial physiology and metabolism*, pp 7-34, Cambridge University Press, U.K.
- Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K. & Koike, T. (2006).** Phosphate-binding tag, a new tool to visualise phosphorylated proteins. *Molecular & Cellular Proteomics*, 5, 749-757.
- Kinoshita, E., Kinoshita-Kikuta, E. & Koike, T. (2009).** Separation and detection of large phosphoproteins using Phos-tag SDS-PAGE. *Nature Protocols*, 4(10), 1514-1521.
- Kleerebezem, M., Quadri, L.E., Kuipers, O.P. & de Vos, W.M. (1997).** Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Molecular Microbiology*, 24, 895-904.
- Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O.P., Leer, R., Tarchini, R., Peters, S.A., Sandbrink, H.M., Fiers, M.W.E.J., Stiekema, W., Lankhorst, R.M.K., de Vries, M., Ursing, B., de Vos, W.M. & Siezen, R.J.**

- (2003). Complete genome sequence of *Lactobacillus plantarum* WCFS1. Proceedings in National Academy of Science USA, 100-1990-1995.
- Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K-I., Ogasawara, N., Tanaka, T. & Fujita, Y. (2001).** Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. *Journal of Bacteriology*, 183(24), 7365-7370.
- Kociolek, M.G. (2009).** Quorum-sensing inhibitors and biofilms. *Anti-Infective Agents in Medicinal Chemistry*, 8(4), 315-326.
- Konings, W.N., Poolman, B. & van Veen, H.W. (1994).** Solute transport and energy transduction in bacteria. *Antonie Van Leeuwenhoek*, 65(4), 369-380.
- Kornblum, J., Kreishwirth, B., Projan, S.J., Ross, H. & Novick, R.P. (1990).** Agr: A polycistronic locus regulating exoprotein synthesis in *Staphylococcus aureus*. In *Molecular biology of the staphylococci* (Novick, R.P. Ed.), pp 373-402, VCH Publishers, New York.
- Krebs, E.G. & Fischer, E.H. (1956).** The phosphorylase *b* to *a* converting of rabbit skeletal muscle. *Biochimica et Biophysica Acta*, 20, 150-157.
- Kristich, C.J., Nguyen, V.T., Barnes, Le, T., Barnes, A.M., Grindle, S. & Dunny, G.M. (2008).** Development and use of an efficient system for random mariner transposon mutagenesis to identify novel genetic determinants of biofilm formation in the core *Enterococcus faecalis* genome. *Applied Environmental Microbiology*, 74, 3377-3386.
- Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L. (2001).** Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, 305(3), 567-580.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F. (2001).** The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390, 249-256.
- Kuo, J.F. & Greengard, P. (1969).** An adenosine 3', 5'-monophosphate-dependent protein kinase from *Escherichia coli*. *Journal of Biological Chemistry*, 244, 3417-3419.
- Kuroda, M., Kuwahara-Arai, K. & Hiramatsu, K. (2000).** Identification of the up- and down-regulated genes in vancomycin-resistant *Staphylococcus aureus* strains Mu3 and Mu50 by cDNA differential hybridisation method. *Biochemical and Biophysical Research Communications*, 269, 485-490.
- Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T.,**

- Hattori, M., Ogasawara, N., Hayashi, H. & Hiramatsu, K. (2001).** Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet*, 357, 1225-1240.
- Kuroda, M., Kuroda, H., Oshima, T., Takeuchi, F., Mori, H. & Hiramatsu, K. (2003).** Two-component system VraSR positively modulates the regulation of cell-wall biosynthesis pathway in *Staphylococcus aureus*. *Molecular Microbiology*, 49(3), 807-821.
- Lamy, M.-C., Zouine, M., Fert, J., Vergassola, M., Courve, E., Pellegrini, E., Glaser, P., Kunst, F., Msadek, T., Trieu-Cout, P. & Poyart, C. (2004).** CovS/CovR of group B streptococcus: a two-component global regulatory system involved in virulence. *Molecular Microbiology*, 54(5), 1250-1268.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2007).** ClustalW and ClustalX version 2. *Bioinformatics*, 23(21), 2947-2948.
- Laskowska, E., Bohdanowicz, J., Kuczyńska-Wiśnik, D., Matuszewska, E., Kedzierska, S. & Taylor, A. (2004).** Aggregation of heat-sock-denatured endogenous proteins and distribution of the IbpA/B and Fda marker-proteins in *Escherichia coli* WT and *grpE280* cells. *Microbiology*, 150, 247-259.
- Lazazzera, B.A. & Grossman, A.D. (1998).** The ins and outs of peptide signaling. *Trends in Microbiology*, 6, 288-294.
- Le Brenton, Y., Boël, G., Benachour, A., Prévost, H., Auffray, Y. & Rincé, A. (2003).** Molecular characterisation of *Enterococcus faecalis* two-component signal transduction pathways related to environmental stresses. *Environmental Microbiology*, 5, 329-337.
- Le Maire, M., Champeil, P. & Møller, J.V. (2000).** Interaction of membrane proteins and lipids with solubilising detergents. *Biochimica et Biophysica Acta*, 1508, 86-111.
- Lee, J.M., Cho, H.Y., Cho, H.J., Do, I.J., Park, S.W., Balk, H.S., Oh, J.H., Eom, C.Y., Kim, Y.M. & Kang, B.S. (2008).** O₂- and NO-sensing mechanism through the DevSR two-component system in *Mycobacterium smegmatis*. *Journal of Bacteriology*, 190, 6795-6804.
- Lees, J.G. & Wallace, B.A. (2002).** Synchrotron radiation circular dichroism and conventional circular dichroism spectroscopy: A comparison. *Spectroscopy*, 16, 121-125.
- Leroux, B., Yanofsky, M.F., Winans, S.C., Ward, J.E., Ziegler, S.F. & Nester, E.W. (1987).** Characterisation of the *virA* locus of *Agrobacterium tumefaciens*: a transcriptional regulator and host range determinant. *Embo Journal*, 6, 849-856.
- Letunic, I., Doerks, T. & Bork, P. (2009).** SMART6: recent updates and new developments. *Nucleic Acids Research*, 37(Database issue):D229-32.

- Levin, J.C. & Wessels, M.R. (1998).** Identification of *csrR/csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A *Streptococcus*. *Molecular Microbiology*, 30, 209-219.
- Lewis, K. (2001).** Riddle of biofilm formation resistance. *Antimicrobial Agents and Chemotherapy*, 45, 999-1007.
- Liang, W.J., Wilson, K.J., Xie, H., Knol, J., Suzuki, S., Rutherford, N., Henderson, P.J.F., Jefferson, R.A. (2005).** The *gusBC* Genes of *Escherichia coli* Encode a Glucuronide Transport System. *Journal of Bacteriology*, 187(7), 2377-2385.
- Lin, P.-F., Samanta, H., Bechtold, C.M., Deminie, C.A., Patick, A.K., Alam, M., Riccardi, K., Rose, R.E., White, R.J. & Colonna, R.J. (1996).** Characterisation of siamycin I, a human immunodeficiency virus fusion inhibitor. *Antimicrobial agents and chemotherapy*, 40(1), 133-138.
- Lipmann, F.A. & Levene, P.A. (1932).** Serinephosphoric acid obtained on hydrolysis of vitellinic acid. *Journal of Biological Chemistry*, 98, 109-114.
- Loewen, S.K., Ng, A.M., Mohabir, N.N., Baldwin, S.A., Cass, C.E. & Young, J.D. (2003).** Functional characterisation of a H⁺/nucleoside co-transporter (CaCNT) from *Candida albicans*, a fungal member of the concentrative nucleoside transporter (CNT) family of membrane proteins. *Yeast*, 20, 661-675.
- Loh, K.D., Gyaneshwar, P., Papadimitriou, E.M., Fong, R., Kim, K-W., Parales, R., Zhou, Z., Inwood, W. & Kustu, S. (2006).** A previously undescribed pathway for pyrimidine catabolism. *PNAS*, 103(13), 5114-5119.
- Lombardino, J.G. & Lowe, J.A. III (2004).** The role of the medicinal chemistry in drug discovery – then and now. *Nature Reviews Drug Discovery*, 3, 853-862.
- Lundstrom, K. (2006).** Structural genomics for membrane proteins. *Cellular and Molecular Life Sciences*, 63, 2597-2607.
- Lyon, G.J. & Novick, R.P. (2004).** Peptide signalling in *Staphylococcus aureus* and other Gram-positive bacteria. *Peptides*, 25, 1389-1403.
- Ma, P., Yuille, H.M., Blessie, V., Göhring, N., Iglói, Z., Nishiguchi, K., Nakayama, J., Henderson, P.J.F. & Phillips-Jones, M.K. (2008).** Expression, purification and activities of the entire family of intact membrane sensor kinases from *Enterococcus faecalis*. *Molecular Membrane Biology*, 25(6-7), 449-473.
- Madiraju, M.V., Brunner, D.P. & Wilkinson, B.J. (1987).** Effects of temperature, NaCl, and methicillin on penicillin-binding proteins, growth, peptidoglycan synthesis, and autolysis in methicillin-resistance *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 31, 1727-1733.
- Magnuson, R., Solomon, J. & Grossman, A.D. (1994).** Biochemical and genetic characterisation of a competence pheromone from *B. subtilis*. *Cell*, 77, 207-216.

- Malpica, R., Franco, B., Rodriguez, C., Kwon, O. & Georgellis, D. (2004).** Identification of a quinone-sensitive redox switch in the ArcB sensor kinase.
- Manefield, M. & Turner, S.L. (2002).** Quorum sensing in context: out of molecular biology and into microbial ecology. *Microbiology*, 148, 3762-3764.
- Mann, E.E., Rice, K.C., Boles, B.R., Endres, J.L. & Ranjit, D. (2009).** Modulation of eDNA release and degradation affects *Staphylococcus aureus* biofilm maturation. *PLoS ONE* 4(6): e5822. doi:10.1371/journal.pone.0005822.
- Marina, A., Waldburger, C.D. & Hendrickson, W.A. (2005).** Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *The EMBO Journal*, 24, 4247-4259.
- Martin, C.A., Hoven, A.D. & Cook, A.M. (2008).** Therapeutic frontiers: preventing and treating infectious diseases by inhibiting bacterial quorum sensing. *European Journal of Clinical Microbiology and Infectious Diseases*, 27, 635-642.
- Martin, M., Albanesi, D., Alzari, P.M. & de Mendoza, D. (2009).** Functional *in vitro* assembly of the integral membrane bacterial thermosensor DesK. *Protein Experiment and Purification*, 66, 39-45.
- Martinez, S.E., Beavo, J.A. & Hol, W.G. (2002).** GAF domains: two-billion-years-old molecular switches that bind cyclic nucleotides. *Molecular Interventions*, 2, 317-323.
- Mascher, T., Margulis, N., Wang, T., Ye, R.W. & Helmann, J.D. (2003).** Cell wall stress response in *Bacillus subtilis*: the regulatory network of the bacitracin stimulon. *Molecular Microbiology*, 50(5), 1591-1604.
- Mäser, P., Sütterlin, C. Kralli, A. & Kaminsky, R. (1999).** A nucleoside transporter from *Tyranosoma brucei* involved in drug resistance. *Science*, 285, 242-244.
- Mates, S.M., Eisenberg, E.S., Mandel, L.J., Patel, L., Kaback, H.R. & Miller, M.H. (1982).** Membrane potential and gentamicin uptakes in *Staphylococcus aureus*. *Proceedings in National Academy of Science USA*, 79, 6693-6697.
- Mathiopoulos, C., Mueller, J.P., Slack, F.J., Murphy, C.G., Patankar, S., Bukusoglu, G. & Sonenshein, A.L. (1991).** A *Bacillus subtilis* dipeptide transport system expressed early during sporulation.
- Matsushita, M. & Janda, K.D. (2002).** Histidine kinases as targets for new antimicrobial agents. *Bioorganic and Medicinal Chemistry*, 10, 855-867.
- McPherson, M.J. & Wootton, J.C. (1983).** Complete nucleotide sequence of the *Escherichia coli* *gdhA* gene. *Nucleic Acids Research*, 11(15), 5257-5266.
- Meroueh, S.O., Bencze, K.Z., Heseck, D., Lee, M., Fisher, J.F., Stemmler, T.L. & Mobashery, S. (2006).** Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proceedings in National Academy of Science USA*, 103(12), 4404-4409.

- Mesnage, S., Chau, F., Dubost, L. & Arthur, M. (2008).** Role of N-acetylglucosaminidase and N-acetylmuramidase activities in *Enterococcus faecalis* peptidoglycan metabolism. *The Journal of Biological Chemistry*, 283, 19845-19853.
- Miller, M.V. & Bassler, B.L. (2001).** Quorum sensing in bacteria. *Annual Review in Microbiology*, 55, 165-199.
- Mirza, O., Guan, L., Verner, G., Iwata, S. & Kaback, H.R. (2006).** Structural evidence for induced fit and a mechanism for sugar/H⁺ symport in LacY. *The EMBO Journal*, 25, 1177-1183.
- Mitchell, P. (1957).** A general theory of membrane transport from studies of bacteria. *Nature*, 180, 134-136.
- Mitchell, P. (1961).** Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature*, 191, 144-148.
- Mitchell, P. (1966).** Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biology Reviews in Cambridge Philosophy Society*, 41, 445-502.
- Miyamoto, C.M., Boylan, M., Graham, A.F. & Meighen, E.A. (1988).** Organisation of the *lux* structural genes of *Vibrio harveyi*. Expression under the T7 bacteriophage promoter. *Molecular Microbiology*, 36, 594-607.
- Mizuno, T., Wurtzel, E.T. & Inouye, M. (1982).** Osmoregulation of gene expression. II. DNA sequence of the *envZ* gene of the *ompB* operon of *Escherichia coli* and characterisation of its gene product. *Journal of Biological Chemistry*, 257, 13692-13698.
- Moglich, A., Ayers, R.A. & Moffat, K. (2009).** Structure and signalling mechanism of Per-Arnt-Sim domains. *Structure*, 17, 1282-1294.
- Mohamed, J.A., Huang, D.B. (2007).** Biofilm formation by enterococci. *Journal of Medical Microbiology*, 56, 1581-1588.
- Mohedano, M.L., Overweg, K., de la Fuente, A., Reuter, M., Aktabem S., Mulholland, F., de Mendoza, D., Lopez, P. & Wells, J.M. (2005).** Evidence that the essential response regulator YycF in *Streptococcus pneumoniae* modulates expression of fatty acid biosynthesis genes and alters membrane composition. *Journal of Bacteriology*, 187, 2357-2367.
- Moore, J.O. & Hendrickson, W.A. (2009).** Structural analysis of sensor domains from the TMAO-responsive histidine kinase receptor TorS. *Structure*, 17, 1195-1204.
- Muller, C., Sanguinetti, M., Riboulet, E., Hébert, L., Posteraro, B., Fadda, G., Auffray, Y. & Rincé, A. (2008).** Characterization of two signal transduction systems involved in intracellular macrophage survival and environmental stress response in *Enterococcus faecalis*. *Journal of Molecular Microbiology and Biotechnology*, 14, 59-66.

- Mullner, M., Hammel, O., Mienert, B., Schlag, S., Bill, E. & Unden, G. (2008).** A PAS domain with an oxygen labile [4Fe-4S(2+)] cluster in the oxygen sensor kinase NreB of *Staphylococcus carnosus*. *Biochemistry*, 47, 13921-13932.
- Murakami, S., Nakashima, R., Yamashita, E. & Yamaguchi, A. (2002).** Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*, 419, 587-593.
- Murray, B.E. (1990).** The life and times of the Enterococcus. *Clinical Microbiology Review*, 3, 46-65.
- Muse, W.B. & Bender, R.A. (1998).** The *nac* (nitrogen assimilation control) gene from *Escherichia coli*. *Journal of Bacteriology*, 180(5), 1166-1173.
- Muse, W.B., Rosario, C.J. & Bender, R.A. (2003).** Nitrogen regulation of the *codBA* (cytosine deaminase) operon from *Escherichia coli* by the nitrogen assimilation control protein, NAC. *Journal of Bacteriology*, 185, 2920-2926.
- Mylonakis, E., Engelbert, M., Qin, X., Sifri, C.D., Murray, B.E. & Ausubel, F.M. (2002).** The *Enterococcus faecalis* *fsrB* gene, a key component of the *fsr* quorum-sensing system, is associated with virulence in the rabbit endophthalmitis model. *Infection and Immunity*, 70, 4678-4681.
- Nakano, M.M., Yang, F., Hardin, P. & Zuber, P. (1995).** Nitrogen regulation of *nasA* and the *nasB* operon, which encodes genes required for nitrate assimilation in *Bacillus subtilis*. *Journal of Bacteriology*, 177, 573-579.
- Nakano, M.M., Hoffman, T., Zhu, Y. & Jahn, D. (1998).** Nitrogen and oxygen regulation of *Bacillus subtilis* *nasDEF* encoding NADH-dependent nitrite reductase by TnrA and ResDE. *Journal of Bacteriology*, 180, 5344-5350.
- Nakayama, J., Cao, Y., Horii, T., Sakuda, S., Akkermans, A.D.L. & De Vos, W.M. (2001a).** Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates a quorum sensing in *Enterococcus faecalis*. *Molecular Microbiology*, 41, 145-154.
- Nakayama, J., Cao, Y., Horii, T., Sakuda, S. & Nagasawa, H. (2001b).** Chemical synthesis and biological activity of the gelatinase biosynthesis-activating pheromone of *Enterococcus faecalis* and its analogs. *Bioscience Biotechnology and Biochemistry*, 65, 2322-2325.
- Nakayama, J., Chem, S.M., Oyama, N., Nishiguchi, K., Azab, E.A., Tanaka, E., Kariyama, R. & Sonomoto, K. (2006).** Revised model for *Enterococcus faecalis* *fsr* quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal AgrD. *Journal of Bacteriology*, 188, 8321-8326.
- Nakayama, J., Uemura, Y., Nishiguchi, K., Yoshimura, N., Igarashi, Y. & Sonomoto, K. (2009).** Ambuic acid inhibits the biosynthesis of cyclic peptide quorumones in Gram-positive bacteria. *Antimicrobial agents and chemotherapy*, 53(2), 580-586.

- National Nosocomial Infections Surveillance (2004).** System report, data summary from January 1992 through June 2004, issued October 2004. A report from the NNIS system. *American Journal of Infection Control*, 32(8), 470-485.
- Nealson, K.H. & Hastings, J.W. (1979).** Bacterial bioluminescence: its control and ecological significance. *Microbiology Review*, 43, 496-518.
- Neiditch, M.B., Ferderle, M.J., Miller, S.T, Bassler, B.L. & Hughson, F.M. (2005).** Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. *Molecular Cell*, 18, 507-518.
- Neiditch, M.B., Federle, M.J., Pompeani, A.J., Kelly, R.C., Swem, D.L., Jeffrey, P.D., Bassler, B.L., Hughson, F.M. (2006).** Ligand-induced asymmetry in histidine sensor kinase complex regulates quorum sensing. *Cell*, 126, 1095-1108.
- Nes, I.F., Diep, D.B., Havarstein, L.S., Brurberg, M.B., Eijsink, V. & Holo, H. (1996).** Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek*, 70, 113-128.
- Newby, Z. E.R., O'Connell, J.D., Gruswitz, F., Hays, F.A., Harries, W.E., Harwood, I.M., Ho, J.D., Lee, J.K., Savage, D.F., Miercke, L.J.W. & Stroud, R.M. (2009).** A general protocol for the crystallisation of membrane proteins for X-ray structural investigation. *Nature Protocols*, 4, 619-637.
- Ng, E. (2001).** Rare codon caltor. <http://www.doe-mbi.ucla.edu/~sumchan/caltor.html>. Accessed on 20/08/2009
- Ng, W.-L., Robertson, G.T., Kazmierczak, K.M., Zhao, J., Gilmour, R. & Winkler, M.E. (2003).** Constitutive expression of PcsB suppresses the requirement for the essential VicR (YycF) response regulator in *Streptococcus pneumoniae* R6. *Molecular Microbiology*, 50, 1647-1663.
- Ng, W.-L., Kazmierczak, K.M. & Winkler, M.E. (2004).** Defective cell wall synthesis in *Streptococcus pneumoniae* R6 depleted for the essential PcsB murein hydrolase or the VicR (YycF) response regulator. *Molecular Microbiology*, 53, 1161-1175.
- Ng, W.-L., Tsui, H.-C.T. & Winkler, M.E. (2005).** Regulation of the *pspA* virulence factor and essential *pcsB* murein biosynthetic genes by the phosphorylated VicR (YycF) response regulator in *Streptococcus pneumoniae*. *Journal of Bacteriology*, 187, 7444-7459.
- Ng, W.-L. & Bassler, B.L. (2009).** Bacterial quorum-sensing network architectures. *Annual Review in Genetics*, 43, 197-222.
- Nicola, G. & Abagyan, R. (2009).** Structure-based approaches to antibiotic drug discovery. *Current Protocols in Microbiology*, Chapter 17: Unit 17.2.
- Nikaido, H. & Vaara, M. (1985).** Molecular basis of bacterial outer membrane permeability. *Microbiology Reviews*, 49(1), 1-32.

- Nikaido, H. & Saier Jr, M.H. (1992).** Transport proteins in bacteria: common themes in their design. *Science*, 258(5084), 936-942.
- Nikaido, H. (2003).** Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews*, 67(4), 593-656.
- Ninfa, A.J., Ninfa, E.G., Lupas, A.N., Stock, A., Magasanik, B. & Stock, J. (1988).** Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. *Proceedings in National Academy of Science USA*, 85, 5492-5496.
- Nishiguchi, K., Nagata, K., Tanokura, M., Sonomoto, K. & Nakayama, J. (2009).** Structure-activity relationship of gelatinase biosynthesis-activating pheromone of *Enterococcus faecalis*. *Journal of Bacteriology*, 191(2), 641-650.
- Nixon, B.T., Ronson, C.W. & Ausubel, F.M. (1986).** Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. *Proceedings in National Academy of Science USA*, 83, 7850-7854.
- Njoroge, J. & Sperandio, V. (2009).** Jamming bacterial communication: New approaches for the treatment of infectious diseases. *EMBO Molecular Medicine*, 1(4), 201-210.
- Nolling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J, Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, E.V. & Smith, D.R. (2001).** Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *Journal of Bacteriology*, 183, 4823-4838.
- Nomura, M & Hosoda, J. (1956).** Nature of the primary action of the autolysin of *Bacillus subtilis*. *Journal of Bacteriology*, 72, 573-581.
- Novak, R., Henriques, B., Charpentier, E., Normark, S. & Tuomanen, E. (1999).** Emergence of vancomycin tolerance in *Streptococcus pneumoniae*. *Nature*, 399, 590-593.
- Novick, R.P. (2003).** Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Molecular Microbiology*, 48, 1429-1449.
- Nygaard, P., Duckert, P. & Saxild, H.H. (1996).** Role of adenine deaminase in purine salvage and nitrogen metabolism and characterisation of the *ade* gene in *Bacillus subtilis*. *Journal of Bacteriology*, 178, 846-853.
- Nygaard, P., Bsted, S.M., Anderson, K.A.K. & Saxild, H.H. (2000).** *Bacillus subtilis* guanine deaminase is encoded by the *yknA* gene and is induced during growth with purines as the nitrogen source. *Microbiology*, 146, 306103069.

- Ogawa, J. & Shimizu, S. (1997).** Diversity and versatility of microbial hydantoin-transforming enzymes. *Journal of Molecular Catalysis B: Enzymatic*, 2, 163-176.
- Oh, J. -I. & Kaplan, S. (2000).** Redox signaling: globalization of gene expression. *EMBO*, 19, 4237-4247.
- Oncu, S., Punar, M. & Eraksoy, H. (2004).** Susceptibility patterns of enterococci causing infections. *Tohoku Journal of Experimental Medicine*, 202, 23-29.
- Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. (2006)** How many drug targets are there? *Nature Review Drug Discovery*, 5, 993-996.
- Otto, M. (2004).** Quorum-sensing control in staphylococci – a target for anti-microbial drug therapy? *FEMS Microbiology Letters*, 241, 135-141.
- Paget, M.S.B. & Buttner, M.J. (2003).** Thiol-based regulatory switches. *Annual Review in Genetics*, 37, 91-121.
- Pantazopoulou, A. & Diallinas, G. (2007).** Fungal nucleobase transporters. *FEMS Microbiology Reviews*, 31(6), 657-675.
- Pappalardo, L., Janausch, I.G., Vijayan, V., Zientz, E., Junker, J., Peti, W., Zweckstetter, M., Unden, G. & Griesinger, C. (2003).** The NMR structure of the sensory domain of the membranous two-component fumarate sensor (histidine protein kinase) DcuS of *Escherichia coli*. *Journal of Biological Chemistry*, 277, 39185-39188.
- Parkinson, J.S. & Kofoid, E.C. (1992).** Communication modules in bacterial signalling proteins. *Annual Reviews in Genetic*, 26, 71-112.
- Parsek, M.R., Val, D.L., Hanzelka, B.L., Cronan, J.E. Jr & Greenberg, E.P. (1999).** Acyl homoserine-lactone quorum-sensing signal generation. *Proceedings in National Academy of Science USA*, 96, 4360-4365.
- Patton, T. G., Rice, K.C., Foster, M.K. & Bayles, K.W. (2005).** The *Staphylococcus aureus cidC* gene encodes a pyruvate oxidase that affects acetate metabolism and cell death in stationary phase. *Molecular Microbiology*, 56, 1664-1674.
- Patton, T.G., Yang, S.J. & Bayles, K.W. (2006).** The role of protein motive force in expression of the *Staphylococcus aureus cid* and *lrg* operons. *Molecular Microbiology*, 59, 1395-1404.
- Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., Persinger, C.C., Munos, B.H., Lindborg, S.R. & Schacht, A.L. (2010).** How to improve R& D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, 9(3), 203-214.
- Paulsen, I. T., Sliwinski, M. K. & Saier, M. H., Jr. (1998).** Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *Journal of Molecular Biology*, 277, 573-592.

Paulsen, I.T., Nguyen, L., Sliwinski, M.K., Rabus, R. & Saier, M.H., Jr (2000). Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *Journal of Molecular Biology*, 301, 75-100.

Paulson, I.T., Banerjee, L., Myers, G.S., Nelson, K.E., Seshadri, R., Read, T.D., Fouts, D.E., Eisen, J.A., Gill, S.R., Heidelberg, J.F., Tettelin, H., Dodson, R.J., Umayam, L., Brinkac, L., Beanan, M., Daugherty, S., DeBoy, R.T., Durkin, S., Kolonay, J., Madupu, R., Nelson, W., Vamathevan, J., Tran, B., Upton, J., Hansen, T., Shetty, J., Khouri, H., Utterback, T., Radune, D., Ketchum, K.A., Dougherty, B.A. & Fraser, C.M. (2003). Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science*, 299(5615), 2071-2074.

Pearson, W.R. (2000). Flexible sequence similarity searching with the FASTA3 program package. *Methods in Molecular Biology*, 132, 185-219.

Perkins, H.R. (1990). The bacterial autolysins. In *Microbial cell walls* (Rogers, H.J., Perkins, H.R. & Ward, J.B. Ed.). pp 437-456. Chapman and Hall, London, UK.

Pestova, E.V., Havarstein, L.S. & Morrison, D.A. (1996). Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. *Molecular Microbiology*, 21, 853-862.

Pillai, S.K., Sakoulas, G., Eliopoulos, G.M., Moellering, R.C., Murray, B.E., Jr. & Inouye, R.T. (2004). Effects of glucose on *fsr*-mediated biofilm formation in *Enterococcus faecalis*. *Journal of Infections and Diseases*, 190, 967-970.

Pinson, B., Napias, C., Chevallier, J. Van den Broek, P.J.A. & Brèthes, D. (1997). Characterization of the *Saccharomyces cerevisiae* cytosine transporter using energizable plasma membrane vesicles. *The Journal of Biological Chemistry*, 272(46), 28918-28924.

Podbielski, A. & Kreikemeyer, B. (2004). Cell density – dependent regulation: basic principles and effects on the virulence of Gram-positive cocci. *International Journal of Infectious Diseases*, 8, 81-95.

Podust, L.M., Ioanoviciu, A. & Ortiz de Montellano, P.R. (2008). 2.3 Å X-ray structure of the heme-bound GAF domain of sensory histidine kinase DosT of *Mycobacterium tuberculosis*. *Biochemistry*, 47(47), 12523-12531.

Poncet, S., Soret, M., Mervelet, P., Deutscher, J. & Noirot, P. (2009). Transcriptional activator YesS is stimulated by histidine-phosphorylated HPr of the *Bacillus subtilis* phosphotransfer system. *The Journal of Biological Chemistry*, 284(41), 28188-28197.

Poolman, B. & Konings, W.N. (1993). Secondary solute transport in bacteria. *Biochimica et Biophysica Acta*, 1183, 5-39.

Porter, J.R. (1976). Antony van Leeuwenhoek: tercentenary of his discovery of bacteria. *Bacteriology Reviews*, 40(2), 260-269.

- Postis, V., Deacon, S., Roach, P., Wright, G., Xia, X., Ingram, J., Hadden, J., Henderson, P., Phillips, S., McPherson, M. & Baldwin, S. (2008).** A high-throughput assay of membrane protein stability. *Molecular Membrane Biology*, 25(8), 617-624.
- Potter, C.A., Ward, A., Laguri, C., Williamson, M.P., Henderson, P.J.F. & Phillips-Jones, M.K. (2002).** Expression, purification and characterization of full-length histidine protein kinase RegB from *Rhodobacter sphaeroides*. *Journal of Molecular Biology*, 320(2), 201-213.
- Potter, C.A., Jeong, E-J., Williamson, M.P., Henderson, P.J.F. & Phillips-Jones, M.K. (2006).** Redox-responsive in vitro modulation of the signaling state of the isolated PrrB sensor kinase of *Rhodobacter sphaeroides* NCIB 8253. *FEBS Letters*, 580(13), 3206-3210.
- Purcell, E.B., Siegal-Gaskins, D., Rawling, D.C., Fiebig, A. & Crosson, S. (2007).** A photosensory two-component system regulates bacterial cell attachment. *Proceedings in National Academy of Science USA*, 104, 18241-18246.
- Qi, F. & Turnbough, C.L. Jr. (1995).** Regulation of codBA operon expression in *Escherichia coli* by UTP-dependent reiterative transcription and UTP-sensitive transcriptional start site switching. *Journal of Molecular Biology*, 254(4), 552-565.
- Qin, X., Singh, K.V., Weinstock, G.M. & Murray, B.E. (2000).** Effects of *Enterococcus faecalis* *fsr* genes on production of gelatinase and a serine protease and virulence. *Infection and Immunity*, 68, 2579-2586.
- Qin, X., Singh, K.V., Weinstock, G.M. & Murray, B.E. (2001).** Characterisation of *fsr*, a regulator controlling expression of gelatinase and serine protease in *Enterococcus faecalis* OG1RF. *Journal of Bacteriology*, 183, 3372-3382.
- Raffa, R.B., Iannuzzo, R., Levine, D.R., Saeid, K.K., Schwartz, R.C., Sucic, N.T., Terleckyj, O.D. & Young, J.M. (2005).** Bacterial communication ("Quorum Sensing") via ligands and receptors: A novel Pharmacological target for the design of antibiotic drugs. *Journal of Pharmacological Experiment and Therapy*, 312, 417-423.
- Rahman, M., Ismat, F., McPherson, M.J. & Baldwin, S.A. (2007).** Topology-informed strategies for the overexpression and purification of membrane proteins. *Molecular Membrane Biology*, 24(5-6), 407-418.
- Raina, S., de Vizio, D., Odell, M., Clements, M., Vanhulle, S. & Keshavarz, T. (2009).** Microbial quorum sensing: a tool or a target for antimicrobial therapy? *Biotechnology in Applied Biochemistry*, 54, 65-84.
- Rajagopal, S. & Moffat, K. (2003).** Crystal structure of a photoactive yellow protein from a sensor histidine kinase: conformational variability and signal transduction. *Proceedings in National Academy of science USA*, 100, 1649-1654.
- Rajan, D.P., Huang, W., Dutta, B., Devoe, L.D., Leibach, F.H., Ganapathy, V. & Prasad, P.D. (1999).** Human placental sodium-dependent vitamin C transporter (SVTC2): Molecular cloning and transport function. *Biochemical Biophysical Research Communications*, 262, 762-768.

- Reinelt, S., Hofmann, E., Gerharz, T., Bott, M. & Madden, D.R. (2003).** The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. *Journal of Biological Chemistry*, 278, 39189-39196.
- Ren, Q., Kang, K.H. & Paulsen, I.T. (2004).** TransportDB: a relational database of cellular membrane transport systems. *Nucleic Acids Research*, 32, D284-D288.
- Ren, Q. & Paulsen, I.T. (2005).** Comparative analyses of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. *PLoS Computational Biology*, 1, 190-201.
- Ren, Q., Chen, K. & Paulsen, I.T. (2007).** TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Research*, 35, Database issue doi:10.1093/nar/gk1925.
- Ressl, S., Terwisscha van Scheltinga, A.C., Vonrhein, C., Ott, V. & Ziegler, C. (2009).** Molecular basis of transport and regulation in the Na⁺/betaine symporter BetP. *Nature*, 458(7234), 47-52.
- Rice, K.C., Firek, B.A., Nelson, J.B., Yang, S.J., Patton, T.G. & Bayles, K.W. (2003).** The *Staphylococcus aureus* *cidAB* operon: evaluation of its role in regulation of murein hydrolase activity and penicillin tolerance. *Journal of Bacteriology*, 185, 2635-2643.
- Rice, K.C., Nelson, J.B., Patton, T.G., Yang, S.J. & Bayles, K.W. (2005).** Acetic acid induces expression of *Staphylococcus aureus* *cidABC* and *lrgAB* murein hydrolase regulator operons. *Journal of Bacteriology*, 187, 813-821.
- Rice, K.C., Mann, E. E., Endres, J.L., Weiss, E.C., Cassat, J.E., Smeltzer, M.S. & Bayles, K.W. (2007).** The *cidA* murein hydrolase regulator contributes to DNA release and biofilm development in *Staphylococcus aureus*. *Proceedings in National Academy of Science USA*, 104, 8113-8118.
- Rice, K.C. & Bayles, K.W. (2003).** Death's toolbox: examining the molecular components of bacterial programmed cell death. *Molecular Microbiology*, 50(3), 729-738.
- Rice, K.C. & Bayles, K.W. (2008).** Molecular control of bacterial death and lysis. *Microbiology and molecular biology reviews*, 85-109.
- Ritzel, M.W.L., Yao, S.Y.M., Huang, M.-Y., Elliot, J.F., Cass, C.E. & Young, J.D. (1997).** Molecular cloning and functional expression of cDNAs encoding a human Na⁺-nucleoside cotransporter (hCNT1). *The American Journal of Physiology*, 272, C707-C714.
- Rodionov, D.A., Vitreschak, A.G., Mironov, A.A. & Gelfand, M.S. (2002).** Comparative genomics of thiamine biosynthesis in prokaryotes. *The Journal of Biological Chemistry*, 277(50), 48949-48959.

Rodriguez, C., Bloch, J.C. & Chevallier, M.R. (1995). The immunodetected yeast purine-cytosine permease is not N-linked glycosylated, nor are glycosylation sequences required to have a functional permease. *Yeast*, 11, 15-23.

Rolfes, R.J. & Zalkin, H. (1988). *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence, and interaction with the *purF* operator. *The Journal of Biological Chemistry*, 263(36), 19653-19661.

Ronson, C.W., Astwood, P.M., Nixon, B.T. & Ausubel, F.M. (1987). Deduced products of C4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene products. *Nucleic Acids Research*, 15, 7921-7934.

Rose, L. & Jenkins, A.T. (2007). The effect of the ionophores valinomycin on biomimetic solid supported lipid DTTTE/EPC membranes. *Bioelectrochemistry*, 70(2), 387-393.

Saidijam, M. (2004). Exploitation of molecular methods for elucidating the mechanisms and structure of pathogenic prokaryotic membrane transport protein. Ph.D. Thesis, University of Leeds, Leeds, UK.

Saidijam, M., Benedetti, G., Ren, Q.H., Xu, Z.Q., Hoyle, C.J., Palmer, S.L., Ward, A., Bettaney, K.E., Szakonyi, G., Mueller, J., Morrison, S., Pos, M.K., Butaye, P., Walravens, K., Langton, K., Herbert, R.B., Skurray, R.A., Paulsen, I.T., O'Reilly, J., Rutherford, N.G., Brown, M.H., Bill, R.M. & Henderson, P.J.F. (2006). Microbial drug efflux proteins of the major facilitator superfamily. *Current Drug Targets*, 7(7), 793-811.

Saidijam, M., Bettaney, K. E., Szakonyi, G., Psakis, G., Shibayama, K., Suzuki, S., Clough, J. L., Blessie, V., Abu-Bakr, A., Baumberg, S., Mueller, J., Hoyle, C. K., Palmer, S. L., Butaye, P., Walravens, K., Patching, S. G., O'Reilly, J., Rutherford, N. G., Bill, R. M., Roper, D. I., Phillips-Jones, M. K. & Henderson, P. J. (2005). Active membrane transport and receptor proteins from bacteria. *Biochemical Society Transactions*, 33, 867-87

Saidijam, M., Psakis, G., Clough, J. L., Mueller, J., Suzuki, S., Hoyle, C. J., Palmer, S. L., Morrison, S. M., Pos, M. K., Essenberg, R. C., Maiden, M. C., Abu-bakr, A., Baumberg, S. G., Neyfakh, A. A., Griffith, J. K., Stark, M. J., Ward, A., O'Reilly, J., Rutherford, N. G., Phillips-Jones, M. K. & Henderson, P. J. (2003). Collection and characterisation of bacterial membrane proteins. *FEBS Letters*, 555, 170-175.

Saier, M.H., Jr. (2000). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiology and Molecular Biology Reviews*, 64, 354-411.

Saier, M.H., Jr., Tran, C.V. & Barabote, R.D. (2006). TCDB: the transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Research*, 34, D181-186.

Sambrook, J., Fritsch, E. R. & Maniatis, T. (1989). *Molecular cloning - A laboratory manual*, 2nd Edition. Cold Spring Harbor Laboratory Press, New York.

- Sandoe, J.A., Wytome, J., West, A.P., Heritage, J. & Wilcox, M.H. (2006).** Measurement of ampicillin, vancomycin, linezolid, and gentamicin activity against enterococcal biofilms. *Journal of Antimicrobial Chemotherapy*, 57, 767-770.
- Sardiwal, S., Kendall, S.L., movahedzadeh, F., Rison, S.C., Stoker, N.G. & Djordjevic, S. (2005).** A GAF domain in the hypoxia/NO-inducible *Mycobacterium tuberculosis* DosS protein binds haem. *Journal of Molecular Biology*, 353, 929-936.
- Schaberg, D.R., Culver, D.H. & Gaynes, R.P. (1991).** Major trends in the microbial etiology of nosocomial infections. *American Journal of Medicine*, 91(Suppl. 3B): 72S-75S.
- Schaffner, W. & Weissman, C. (1973).** A rapid, sensitive and specific method for the determination of protein in dilute solution. *Analytical Biochemistry*, 56, 502-514.
- Scheffers, D.-J. & Pinho, M.G. (2005).** Bacterial cell wall synthesis: new insights from localisation studies. *Microbiology and Molecular Biology Reviews*, 69(4), 585-607.
- Scheu, P., Sdorra, S., Liao, Y.-F., Wegner, M., Basché, T., Uden, G. & Erker, W. (2008).** Polar accumulation of the metabolic sensory histidine kinases DcuS and CitA in *Escherichia coli*. *Microbiology*, 154, 2463-2472.
- Schreier, H.J., Rostkowski, C.A. & Kellner, E.M. (1993).** Altered regulation of the *glnRA* operon in a *Bacillus subtilis* mutant that produces methionine sulfoximine-tolerant glutamine synthetase. *Journal of Bacteriology*, 175, 892-897.
- Schultz, A., Nygaard, P. & Saxild, H.C. (2001).** Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. *Journal of Bacteriology*, 183(11), 3293-3302.
- Schultz, J., Mileptz, F., Bork, P. & Ponting, C. (1998).** SMART, a simple modular architecture research tool: Identification of signaling domains. *PNAS*, 95(11), 5857-5864.
- Screpanti, E. & Hunte, C. (2007).** Discontinuous membrane helices in transport proteins and their correlation with function. *Journal of Structural Biology*, 159(2), 261-267.
- Sears, C.L. (2005).** A dynamic partnership: celebrating our gut flora. *Anaerobe*, 11(5), 247-251.
- Senadheera, M.D., Guggenheim, B. Spatafora, G.A., Huang, Y.-C.C., Choi, J., Hung, D.C.I., Treglown, J.S., Goodman, S.D., Ellen, R.P. & Cvitkovitch, D.G. (2005).** A VicRK signal transduction system in *Streptococcus mutans* affects *gtfBCD*, *gfpB* and *ftf* expression, biofilm formation and genetic competence development. *Journal of Bacteriology*, 187, 4064-4076.

- Sevvana, M., Vijayan, V., Zweckstetter, M., Reinelt, S., Madden, D.R., Herbst-Irmer, R., Sheldrick, G.M., Bott, M., Griesinger, C. & Becker, S. (2008).** A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA. *Journal of Molecular Biology*, 377, 512-523.
- Shaffer, P., Goehring, A., Shankaranarayanan, A. & Gouaux, E. (2009).** Structure and mechanism of a Na⁺-independent amino acid transporter. *Science* 325, 1010-1014.
- Shankar, N., Baghdayan, A.S. & Gilmore, M.S. (2002).** Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature*, 417, 746-750.
- Sharma-Kuinkel, B.K., Mann, E.E., Ahn, J.S., Kuechenmeister, L.J., Dunman, P.M. & Bayles, K.W. (2009).** The *Staphylococcus aureus* LytSR two-component regulatory system affects biofilm formation. *Journal of Bacteriology*, 191(15), 4767-4775.
- Shimamura, T., Weyand, S., Beckstein, O., Rutherford, N.G., Hadden, J.M., Sharples, D., Sansom, M.S.P., Iwata, S., Henderson, P.J.F. & Cameron, A.D. (2010).** Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science*, 328, 470. DOI:10.1126/science.1186303.
- Shockman, G.D. & Barrett, J.F. (1983).** Structure, function, and assembly of cell walls of gram-positive bacteria. *Annual Reviews in Microbiology*, 37, 501-527.
- Shockman, G.D. & Holtje, J.V. (1994).** Microbial peptidoglycan (murein) hydrolases. In *Bacterial cell wall* vol. 27 (Ghuysen, J.-M. & Hakenbeck, R. Ed.). pp 131-166. Elsevier Science B.V., Amsterdam, The Netherlands.
- Shompole, S., Henon, K.T., Liou, L.E., Dziwanowska, K., Bohach, G.A. & Bayles, K.W. (2003).** Biphasic intracellular expression of *Staphylococcus aureus* virulence factors and evidence for Agr-mediated diffusion sensing. *Molecular Microbiology*, 49, 919-927.
- Shpakov, A.O. (2009).** Peptide autoinducers in bacteria. *Microbiology*, 78(3), 255-266.
- Sifri, C.D., Mylonakis, E., Singh, K.V., Qin, X., Garsin, D.A. & Murray, B.E. (2002).** Virulence effect of *Enterococcus faecalis* protease genes and the quorum-sensing locus *fsr* in *Caenorhabditis elegans* and mice. *Infection and Immunity*, 70, 5647-5650.
- Singer, S.J. & Nicolson, G.L. (1972).** The fluid mosaic model of the structure of cell membranes. *Science*, 175(4023), 720-731.
- Singh, S.K., Yamashita, A. & Gouaux, E. (2007).** Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature*, 448, 952-956.
- Slack, F.J., Serror, P., Joyce, E. & Sonenshein, A.L. (1995).** A gene required for nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *Molecular Microbiology*, 15, 689-702.

- Sleator, R.D. & Hill, C. (2005).** A novel tole for the LisRK two-component regulatory system in listeria osmotolerance. *Clinical Microbiology and Infections*, 11, 599-601.
- Smirnova, I.N., Kasho, V. & Kaback, R. (2008).** Protonation and sugar binding to LacY. *Proceedings in National Academy of Sciences of the USA*, 105(26), 8896-8901.
- Smith, D. & Hanawait, P.C. (1967).** Properties of the growing point region in the bacterial chromosome. *Biochimica et Biophysica Acta*, 148, 519.
- Smith, R.S. & Iglewski, B.H. (2003).** *Pseudomonas aeruginosa* quorum sensing as a potential antimicrobial target. *Journal of Clinical Investigation*, 112, 1460-1465.
- Smith, K.M., Slugoski, M.D., Loewen, S.K., Ng, A.M., Yao, S.Y., Chen, X.Z., Karpinski, E., Cass, C.E., Baldwin, S.A. & Young, J.D. (2005).** The broadly selective human Na⁺/nucleoside cotransporter (hCNT3) exhibits novel cation-coupled nucleoside transport characteristics. *The Journal of Biological Chemistry*, 280, 25436-25449.
- Smits, P.H.M., De Haan, M., Maat, C. & Grivell, L.A. (1994).** The complete sequence of a 33 kb fragment on the right arm of chromosome II from *Saccharomyces cerevisiae* reveals 16 open reading frame, including ten new open reading frames, five previously identified genes and a homologue of the SCO1 gene. *Yeast*, 10(Supplement A): S75-S80.
- Sobczak, I. & Lolkema, J.S. (2005).** Structural and mechanistic diversity of secondary transporters. *Current Opinion in Microbiology*, 8(2), 161-167.
- Sonders, M.S., Quick, M. & Javitch, J.A. (2005).** How did the neurotransmitter cross the bilayer? A closer view. *Current Opinions in Neurobiology*, 15, 296-304.
- Sousa, E.H., Tuckerman, J.R., Gonzalez, G. & Gilles-Gonzalez, M.A. (2007).** DosT and DevS are oxygen-switched kinases in *Mycobacterium tuberculosis*. *Protein Science*, 16(8), 1708-1719.
- Spoering, A.L. & Gilmore, M.S. (2006).** Quorum sensing and DNA release in bacterial biofilms. *Current Opinions in Microbiology*, 9, 133-137.
- Stark, M. J. (1987).** Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. *Gene* 51, 255-267.
- Staroń, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H. & Mascher, T. (2009).** The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Molecular Microbiology*, 74(3), 557-581.
- Stock, A., Chen, T., Welsh, D. & Stock, J. (1988).** CheA protein, a central regulator of bacterial chemotaxis, belongs to a family of proteins that control gene expression in response to changing environmental conditions. *Proceedings in National Academy of Science USA*, 85, 1403-1407.

- Stock, J.B., Ninfa, A.J. & Stock, A.M. (1989).** Protein phosphorylation and regulation of adaptive responses in bacteria. *Microbiology Review*, 53, 450-490.
- Stock, J.B., Surette, M.G., Levit, M. & Park, P. (1995).** Two-component signal transduction systems: structure function relationship and mechanisms of catalysis. In *Two-component signal transduction*, pp 25-51. Edited by Hoch, J.A. & Silhavy, T.J. Washington, D.C.: ASM.
- Stock, A.M., Robinson, V.I. & Goudreau, P.N. (2000).** Two-component signal transduction. *Annual Reviews in Biochemistry*, 69, 183-215.
- Stolz, J. & Vielreicher, M. (2003).** Tpn1p, the plasma membrane vitamin B₆ transporter of *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, 278, 18990-18996.
- Stover, C.K., Pham, X. Q., Erwin, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Largrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E.W., Lory, S. & Olson, M.V. (2000).** Complete genome sequence on *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, 406(6799), 959-964.
- Strauch, M.A., Aronson, A.I., Brown, S.W., Schreier, H.J., Sonenhein, A.L. (1988).** Sequence of the *Bacillus subtilis* glutamine synthase gene region. *Gene*, 71(2), 257-265.
- Sturme, M.H., Nakayama, J., Molenaar, D., Murakami, Y., Kunugi, R., Fujii, T., Vaughan, E.E., Kleerebezem, M. & de Vos, W.M. (2005).** An agr-like two-component regulatory system in *Lactobacillus plantarum* is involved in production of a novel cyclic peptide and regulation of adherence. *Journal of Bacteriology*, 187(15), 5224-5235.
- Su, Y.A., Sulavik, M.C., He, P., Makinen, K.K., Makinen, P.L., Fiedler, s., Wirth, R. & Clewell, D.B. (1991).** Nucleotide sequences of the gelatinase gene (*gelE*) from *Enterococcus faecalis* subsp. *liquefaciens*. *Infection and Immunity*, 59, 415-420.
- Sundaram, M., Yao, S.Y., Ng, A.M., Griffiths, M., Cass, C.E., Baldwin, S.A. & Young, J.D. (1998).** Chimeric constructs between human and rat equilibrative nucleoside transporters (hENT1 and rENT1) reveal hENT1 structural domains interacting with coronary vasoactive drugs. *The Journal of Biological Chemistry*, 273, 21519-21525.
- Surade, S., Klein, M., Stolt-Bergner, P.C., Muenke, C., Roy, A. & Michelet, H. (2006).** Comparative analysis and “expression space” coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sciences*, 15, 2178-2189.
- Suzuki, S., Onisha, N. & Yokozeki, K. (2005a).** Purification and characterisation of hydantoin racemase from *Microbacterium liquefaciens* AJ 3912. *Bioscience and Biotechnological Biochemistry*, 69(3), 530-536.

- Suzuki, S., Takenaka, Y., Onishi, N. & Yokozeki, K. (2005b).** Molecular cloning and expression of the *hyu* genes from *Microbacterium liquefaciens* AJ 3912, responsible for the conversion of 5-substituted hydantoins to α -amino acids, in *Escherichia coli*. *Bioscience and Biotechnological Biochemistry*, 69(8), 1473-1482.
- Suzuki, S. & Henderson, P.J.F. (2006).** The hydantoin transport protein from *Microbacterium liquefaciens*. *Journal of Bacteriology*, 118(9), 3328-3336.
- Swem, L.R., Kraft, B.J., Swem, D.L., Setterdahl, A.T., Masuda, S., Knaff, D.B., Zaleski, J. & Bauer, C.E. (2003).** Signal transduction by the global regulator RegB is mediated by a redox-active cysteine. *The EMBO journal*, 22(18), 4699-4708.
- Swem, L.R., Gong, X., Yu, C.-A. & Bauer, C.E. (2006).** Identification of a ubiquinone-binding site that affects autophosphorylation of the sensor kinase RegB. *The Journal of Biological Chemistry*, 281(10), 6768-6775.
- Swift, S., Throup, J.P., Williams, P., Salmond, G.P. & Stewart, G.S. (1996).** Quorum sensing: a population-density component in the determination of bacterial phenotypes. *Trends in Biochemical Science*, 21, 214-219.
- Syldatk, C., Laufer, A., Muller, R. & Hoke, H. (1990).** Production of optically pure D- and L- α -amino acids by bioconversion of D,L-5-monosubstituted hydantoin derivatives. *Advances in Biochemical Engineering and Biotechnology*, 41, 29-75.
- Syldatk, C., Muller, R., Siemann, M., Krohn, K. & Wagner, F. (1992).** Microbial and enzymatic production of D-amino acids from DL-5-monosubstituted hydantoins. *Biocatalytics production of amino acids and derivatives*. Hanser Publisher, Munich, 77-206.
- Syldatk, C., May, O., Altenbuchner, J., Mattes, R. & Siemann, M. (1999).** Microbial hydantoinases – industrial enzymes from the origin of life? *Applied Microbiology and Biotechnology*, 51, 293-309.
- Szurmant, H., White, R.A. & Hoch, J.A. (2007).** Sensor complexes regulating two-component signal transduction. *Current Opinion in Structural Biology*, 17, 706-715.
- Tanaka, K., Kobayashi, K. & Ogasawara, N. (2003).** The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YfIS, and is essential for utilisation of malate in minimal medium. *Microbiology*, 149, 2317-2329.
- Tanford, C. (1983).** Translocation pathway in the catalysis of active transport. *Proceedings in National Academy of Science USA*, 80, 3701-3705.
- Tang, L., Bai, L., Wang, W.-H. & Jiang, T. (2010).** Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nature Structural & Molecular Biology*, 17(4), 492-497.
- Tanghe, A., Van Dijck, P. & Thevelein, J.M. (2006).** Why do microorganisms have aquaporins? *Trends in Microbiology*, 14(2), 78-85.

Teng, F., Wang, L. Singh, K.V., Murray, B.E. & Weinstock, G.M. (2002). Involvement of PhoP-PhoS homologs in *Enterococcus faecalis* virulence. *Infection and Immunity*, 70, 1991-1996.

Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.M., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickinson, T., Hickey, E.K., Hollingshead, I. & Fraser, C.M. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*, 293, 498-506.

The UniProt Consortium. (2008). The Universal Protein Resources (UniProt). *Nucleic Acids Research*, 36, D190-D195.

Thomas, V.C., Thurlow, L.R., Boyle, D. & Hancock, L.E. (2008). Regulation of autolysis-dependent extracellular DNA release by *Enterococcus faecalis* extracellular proteases influences biofilm development. *Journal of Bacteriology*, 5690-5698.

Thomas, V.C., Hiromasa, Y., Harms, N., Thurlow, L., Tomich, J. & Hancock, L.E. (2009). A fratricidal mechanism is responsible for eDNA release and contribution to biofilm development of *Enterococcus faecalis*. *Molecular Microbiology*, 72(4), 1022-1036.

Thompson, J.D., Gibson, T.J. & Higgins, D.G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, Chapter 2, Unit 2.3.

Tobin, P.J., Mani, N. & Jayaswal, R.K. (1994). Effect of physiological conditions on the autolysis of *Staphylococcus aureus* strains. *Antonie Leeuwenhoek*, 65, 71-78.

Tommassen, J., de Geus, P., Lugtenberg, B., Hackett, J. & Reeves, P. (1982). Regulation of the *pho* regulon of *Escherichia coli* K-12. Cloning of the regulatory genes *phoB* and *phoR* and identification of their gene products. *Journal of Molecular Biology*, 157, 265-274.

Towbin, H., Staehelin, J. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets. Procedure and some applications. *The Proceedings of the National Academy of Sciences USA*, 76, 4350-4354.

Treuner-Lange, A., Kuhn, A. & Dürre, P. (1997). The *kdp* system of *Clostridium acetobutylicum*: cloning, sequencing, and transcriptional regulation in response to potassium concentration. *Journal of Bacteriology*, 179, 4501-4512.

Tsukaguchi, H., Tokui, T., Mackenzie, B., Berger, U.V., Chen, X.Z., Wang, Y., Brubaker, R.F. & Hediger, M.A. (1999). A family of mammalian Na⁺-dependent L-ascorbic acid transporter. *Nature*, 399, 70-75.

Tsunakawa, M., Hu, S.L., Hoshino, Y., Detlefson, D.J., Hill, S.E., Furumai, T., Whilte, R.J., Nishio, M., Kawano, K., Yamamoto, S., Fukagawa, Y. & Oki, T. (1995). Siamycins I and II, new anti-HIV peptides. I. Fermentation, isolation, biological activity and initial characterisation. *Journal of Antibiotics*, 48, 433-434.

- Turner, R.J., Lu, Y. Switzer, R. (1994).** Regulation of the *Bacillus subtilis* pyrimidine biosynthesis (pyr) gene cluster by an autonomous transcriptional attenuation mechanism. *Journal of Bacteriology*, 176, 3707-3722.
- Ueda, A., Attila, C., Whlteley & Wood, T. (2009).** Uracil influences quorum sensing and biofilm formation in *Pseudomonas aeruginosa* and fluorouracil is an antagonist. *Microbial Biotechnology*, 2(1), 62-74.
- Ukaegbu, U.E. & Rosenzweig, A.C. (2009).** Structure of the redox sensor domain of *Methylococcus capsulatus* (Bath) MmoS. *Biochemistry*, 48, 2207-2215.
- Ulrich, L.E., Koonin, E.V. & Zhulin, I.B. (2005).** One-component systems dominates signal transduction in prokaryotes. *Trends in Microbiology*, 13(2), 52-56.
- Vanderheiden, G.J., Fairchild, A.C. & Jago, G.R. (1970).** Construction of a laboratory press for use with the French Pressure cell. *Application in Microbiology*, 19(5), 875-877.
- Vasudevan, G., Carter, N.S., Drew, M.E., Beverley, S.M., Sanchez, M.A., Seyfang, A., Ullman, B. & Landfear, S.M. (1998).** Cloning of *Leishmania* nucleoside transporter genes by rescue of a transport-deficient mutant. *Proceedings in National Academy of Science USA*, 95, 9873-9878.
- Venter, H. (2001).** Applications of biophysical techniques to the analysis of membrane transport proteins. Ph.D. Thesis, University of Leeds (School of Chemistry), Leeds, UK.
- Vlanti, A. & Dhallinas, G. (2008).** The *Aspergillus nidulans* FcyB cytosine-purine scavenger is highly expressed during germination and in reproductive compartments and is down regulated by endocytosis. *Molecular Microbiology*, 68(4), 959-977.
- Voet-van-Vormizeele, J. & Groth, G. (2008).** Ethylene controls autophorylation of the histidine kinase domain in ethylene receptor ETR1. *Molecular plant*, 1, 380-387.
- von Heijne, G. (1986).** The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *The EMBO Journal*, 5(11), 3021-3027.
- Vuong, C., Saenz, H.L., Gotz, F. & Otto, M. (2000).** Impact of the agr quorum-sensing system on adherence to polystyrene in *Staphylococcus aureus*. *Journal of Infection and Disease*, 182, 1688-1693.
- Vuong, C., Gerke, C., Somerville, G.A., Fischer, E.R. & Otto, M. (2003).** Quorum-sensing control of biofilm factors in *Staphylococcus epidermidis*. *Journal of Infection and disease*, 188, 706-718.
- Wagner, R., de Montigny, J., de Wergifosse, P., Souciet, J.L. & Potier, S. (1998).** The ORF *YBL042* of *Saccharomyces cerevisiae* encodes a uridine permease. *FEMS Microbiology Letters*, 159, 69-75.

- Wagner, C., de Saizieu, A., Schonfeld, H.-J., Kamber, M., Lange, R., Thompson, C.J. & Page, M.G. (2002).** Genetic analysis and functional characterisation of the *Streptococcus pneumoniae* vic operon. *Infections and Immunity*, 70, 6121-6128.
- Wagner, J.R., Brunzelle, J.S., Forest, K.T. & Vierstra, R.D. (2005).** A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*, 438, 325-331.
- Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.-A., van Wijk, K.J. & de Gier, J.-W. (2007).** Consequences of membrane protein overexpression in *E. Coli*. *Molecular and Cellular Proteomics*, 6, 1527-1550.
- Walsh, G. (2001).** Protein sources. In *Proteins: Biochemistry and Biotechnology*, pp 51-85, John Wiley & Sons Ltd, West Sussex.
- Wang, J.Y. & Koshland, D.E., Jr. (1978).** Evidence for protein kinase activities in the prokaryote *Salmonella typhimurium*. *Journal of Biological Chemistry*, 253, 7605-7608.
- Wang, I.-N., Smith, D.L. & Young, R. (2000).** Holins: the protein clocks of bacteriophage infections. *Annual Reviews in Microbiology*, 54, 799-825.
- Wang, I.N., Deaton, J. & Young, R. (2003).** Sizing the holin lesion with an endolysin- β -galactosidase fusion. *Journal of Bacteriology*, 185, 779-787.
- Wang, D.-N., Safferling, M., Lemieux, M.J., Griffith, H., Chen, Y. & Li, S.-D. (2003).** Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochimica et Biophysica Acta*, 1610, 23-36.
- Ward, J.B. & Williamson, R. (1985).** Bacterial autolysins: specificity and function. In *Microbial cell wall synthesis and autolysis* (Nombela, C Ed.). pp 159-166. Elsevier Science Publishers, Amsterdam, The Netherlands.
- Ward, A., Sanderson, N.M., O'Reilly, J., Rutherford, N.G., Poolman, B. & Henderson, P. J. F. (2000).** The amplified expression, identification, purification, assay and properties of hexahistidine-tagged bacterial membrane transport proteins. In *Membrane Transport - A Practical Approach* (Baldwins, S.A., Ed.), pp 141-166, Blackwell's Press, Oxford.
- Ward, A., Hoyle, C.J., Palmer, S.A., O'Reilly, J., Griffiths, J.K., Pos, K.M., Morrison, S., Poolman, B., Gwynne, M. & Henderson, P.J.F. (2001).** Prokaryote multidrug efflux proteins of the major facilitator superfamily: amplified expression, purification and characterization. *Journal of Molecular Microbiology and Biotechnology*, 3, 193-200.
- Warne, T., Serrano-Vega, J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G.W., Tate, C. & Schertler, G.F.X. (2008).** Structure of a β_1 -adrenergic G-protein-coupled receptor. *Nature*, 454486-492.
- Waters, C.M. & Bassler, B.L. (2005).** Quorum sensing: cell-to-cell communication in bacteria. *Annual Review in Cell Development and Biology*, 21, 319-346.

- Westh Hansen, S.E., Jensen, N. & Munch-Petersen, A. (1987).** Studies on the sequence and structure of the *Escherichia coli* K-12 nupG gene, encoding a nucleoside-transport system. *European Journal of Biochemistry*, 168(2), 385-391.
- Weyand, S., Shimamura, T., Yajima, S., Suzuki, S., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J., Ma, P., Saidijam, M., Patching, S.G., Hope, R.J., Norbertczak, H.T., Roach, P.C., Iwata, S., Henderson, P.J. & Cameron, A.D. (2008).** Structure and molecular mechanism of a nucleobase-cation-symporter-1 family transporter. *Science*, 322, 709-713.
- Whitchurch, C.B., Tolker-Nielsen, T., Ragas, P.C. & Mattick, J.S. (2002).** Extracellular DNA required for bacterial biofilm formation. *Science*, 295, 1487.
- White, S.H. (2009)** http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html
- Whitehead, N.A., Barnard, A.M.L., Slater, H., Simpson, N.J.L., Salmond, G.P.C. (2001).** Quorum-sensing in Gram-negative bacteria. *FEMS Microbiology Reviews*, 25, 365-404.
- Whitman, W.B., Coleman, D.C. & Wiebe, W.J. (1998).** Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences USA*, 95(12), 6578-6583.
- Wilson, T.H. & Ding, P.Z. (2001).** Sodium-substrate cotransport in bacteria. *Biochimica et Biophysica Acta*, 1505, 121-130.
- Wiese, A., Pietzsch, M., Syldatk, C., Mattes, R. & Altenbuchner, J. (2000).** Hydantoin racemase from *Arthrobacter aurescens* DSM3747: heterologous expression, purification and characterisation. *Journal of Biotechnology*, 80, 217-230.
- Wray, Jr., L.V., Atkinson, M.R. & Fisher, S.H. (1994).** The nitrogen-regulated *Bacillus subtilis* nrgAB operon encodes a membrane protein and a protein highly similar to the *Escherichia coli* glnB-encoded PII protein. *Journal of Bacteriology*, 176(1), 108-114.
- Wray, Jr., L.V., Ferson, A.M., Rohrer, K. & Fisher, S.H. (1996).** TnrA, a transcription factor required for global nitrogen regulation in *Bacillus subtilis*. *Proceedings in National Academy of Science USA*, 93, 8841-8845.
- Wray, Jr., L.V., Ferson, A.E. & Fisher, S.H. (1997).** Expression of the *Bacillus subtilis* ureABC operon is controlled by multiple regulatory factors including CodY, GlnR, TnrA and Spo0H. *Journal of Bacteriology*, 179, 5494-5501.
- Wray, Jr., L.V., Zalieckas, J.M. & Fisher, S.H. (2000).** Purification and *in vitro* activities of the *Bacillus subtilis* TnrA transcription factor. *Journal of Molecular Biology*, 300, 29-40.
- Wray, Jr., L.V., Zalieckas, J.M. & Fisher, S.H. (2001).** *Bacillus subtilis* glutamine synthetase controls gene expression through a protein-protein interaction with transcription factor TnrA. *Cell*, 107, 427-435.

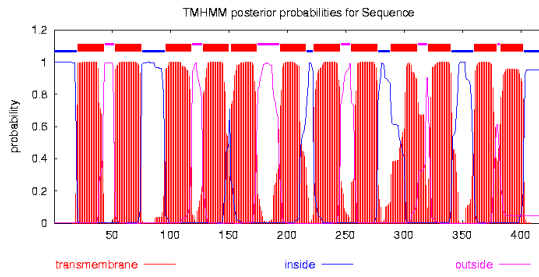
- Wright, E.M. & Turk, E. (2004).** The sodium/glucose cotransport family SLC5. *Pflügers Archiv: European Journal of Physiology*, 447, 510-518.
- Wright, E.M., Hirayama, B.A. & Loo, D.D.F. (2007).** Active sugar transport in health and disease. *Journal of Internal Medicine*, 261, 32-43.
- Wuster, A. & Babu, M.M. (2008).** Conservation and evolutionary dynamics of the *agr* cell-to-cell communication system across firmicutes. *Journal of Bacteriology*, 743-746.
- Xu, M., Stuck, D.K., Deaton, J., Wang, I.N. & Young, R. (2004).** A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proceedings in National Academy of Science USA*, 101, 6415-6420.
- Xu, M., Struck, Arulandu, A., Struck, D.K. Swanson, S., Sacchettini, J.C. & Young, R. (2005).** Disulfide isomerisation after membrane release of its SAR domain activates P1 lysozyme. *Science*, 307, 113-117.
- Yamamoto, K., Kitayama, T., Minagawa, S., Watanabe, T., Sawada, S., Okamoto, T. & Utsumi, R. (2001).** Antibacterial agents that inhibit histidine protein kinase YycG of *Bacillus subtilis*. *Bioscience and Biotechnological Biochemistry*, 65(10), 2306-2310.
- Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R. & Ishihama, A. (2005).** Functional characterisation in vitro of all two-component signal transduction systems from *Escherichia coli*. *The Journal of Biological Chemistry*, 280, 1448-1456.
- Yamamoto, K., Matsumoto, F., Minagawa, S., Oshima, T., Fujita, N., Ogasawara, N. & Ishihama, A. (2009).** Characterisation of CitA-CitB signal transduction activating genes involved in anaerobic citrate catabolism in *Escherichia coli*. *Bioscience and Biotechnological Biochemistry*, 73(2), 346-350.
- Yamashita, A., Singh, S.K., Kawate, T., Jin, Y. & Gouaux, E. (2005).** Crystal structure of a bacterial homologue of Na⁺/Cl⁻ dependent neurotransmitter transporters. *Nature*, 437, 215-223.
- Yang, S.J., Rice, K.C., Brown, R.J., Patton, T.G. & Liou, L.E. (2005).** A LysR-type regulator, CidR, is required for induction of the *Staphylococcus aureus* cidABC operon. *Journal of Bacteriology*, 187, 5893-5900.
- Yang, X., Kuk, J. & Moffat, K. (2008).** Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: photoconversion and signal transduction. *Proceedings in National Academy of Science USA*, 105, 14715-14720.
- Yang, X.J., Kuk, J. & Moffat, K. (2009).** Conformational differences between the Pfr and Pr states in *Pseudomonas aeruginosa* bacteriophytochrome. *Proceedings in National Academy of Science USA*, 106, 15639-15644.
- Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. (2004).** Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature*, 431, 811-818.
- Yin, S., Daum, R.S. & Boyle-Vavra, S. (2006).** VraSR two-component regulatory system and its role in induction of *pbp2* and *vraSR* expression by cell wall

- antimicrobials in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*, 50(1), 336-343.
- Yin, Y., He, X., Szewczyk, P., Nguyen, T. & Chang, G. (2006).** Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science*, 312(5774), 741-744.
- Yocum, R.R., Waxman, D.J., Rasmussen, J.R. & Strominger, J.L. (1979).** Mechanism of penicillin action: penicillin and substrate bind covalently to the same active site serine in two bacterial D-alanine carboxypeptidases. *Proceedings in National Academy of Science USA*, 76(6), 2730-2734.
- Yoo, H.S., Cunningham, T.S. & Cooper, T.G. (1992).** The allantoin and uracil permease gene sequences of *Saccharomyces cerevisiae* are nearly identical. *Yeast* 8, 997-1006.
- Young, R.Y. (1992).** Bacteriophage lysis: mechanism and regulation. *Microbiology review*, 56, 430-481.
- Young, I., Wang, I. & Roof, W.D. (2000).** Phages will out: strategies of host cell lysis. *Trends in Microbiology*, 8, 120-128.
- Yoshida, K., Sano, H., Seki, S., Oda, M., Fujimura, M. & Fujita, Y. (1995).** Cloning and sequencing of a 29 kb region of *Bacillus subtilis* genome containing the *hut* and *wapA* loci. *Microbiology*, 141(Pt2), 337-343.
- Yoshida, K., Yamaguchi, H., Kinehara, M., Ohki, Y.H., Nakaura, Y. & Fujita, Y. (2003).** Identification of additional TnrA-regulated genes of *Bacillus subtilis* associated with TnrA box. *Molecular Microbiology*, 49(1), 157-165.
- Zhang, J., Smith, K.M., Tackaberry, T., Sun, X., Carpenter, P., Slugoski, M.D., Robins, M.J., Nielsen, L.P., Nowak, I., Baldwin, S.A., Young, J.D. & Cass, C.E. (2006).** Characterisation of the transport mechanism and permeant binding profile of the uridine permease Fui1p of *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*, 281, 28210-28221.
- Zhou, Z., Zhen, J., Karpowich, N.K., Goetz, R.M., Law, C.J., Reith, M.E.A. & Wang, D.-N. (2007).** LeuT-Desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science*, 317(5843), 1390-1393.
- Zhou, Y., Guan, L., Freitas, A. & Kaback, H.R. (2008).** Opening and closing of the periplasmic gate in lactose permease. *Proceedings in National Academy of Science USA*, 105(10), 3774-3778.
- Zhou, Y.F., Nan, B.Y., Nan, J., Ma, Q.J., Panjekar, S., Liang, Y.H., Wang, Y.P. & Su, X.D. (2008).** C₄-dicarboxylates sensing mechanism revealed by the crystal structures of DctB sensor domain. *Journal of Molecular Biology*, 383, 49-61.
- Zhulin, I.B., Taylor, B.L. & Dixon, R. (1997).** PAS domain S-boxes in archaea, bacteria and sensors for oxygen and redox. *Trends in Biochemical Sciences*, 22(9), 331-333.

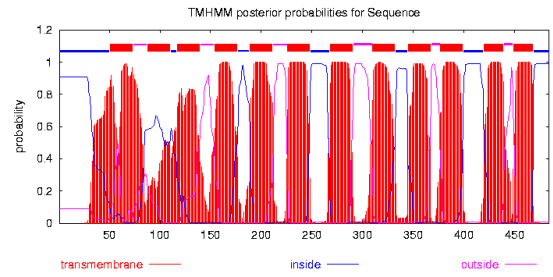
Appendices

Appendix 1 Topology prediction by TMHMM of the selected NCS1 proteins.

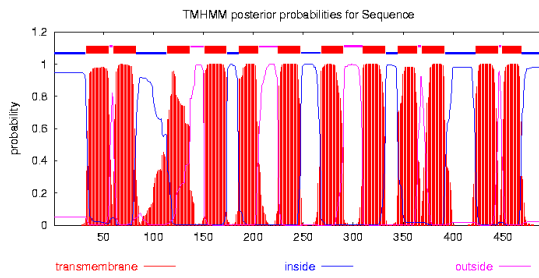
b0336



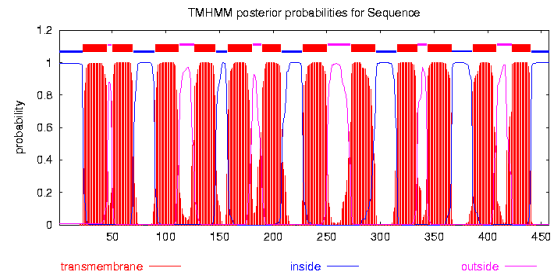
b0511



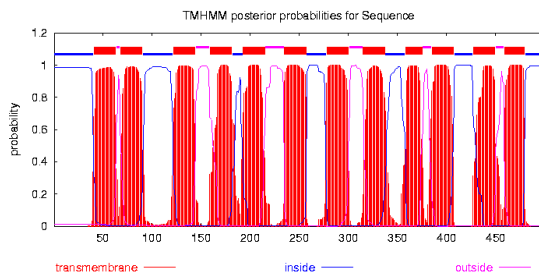
Bsu36470



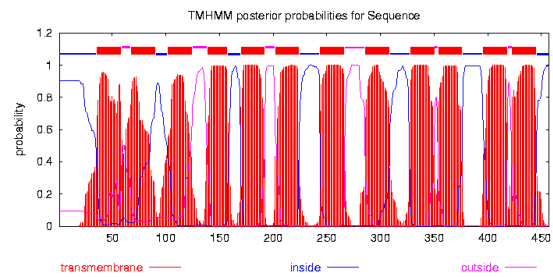
Bsu3867



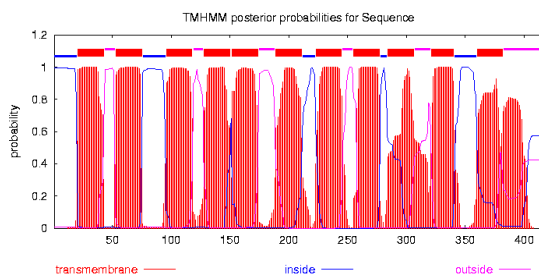
EF3000



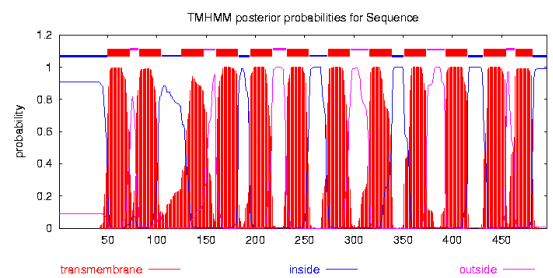
Ip3374



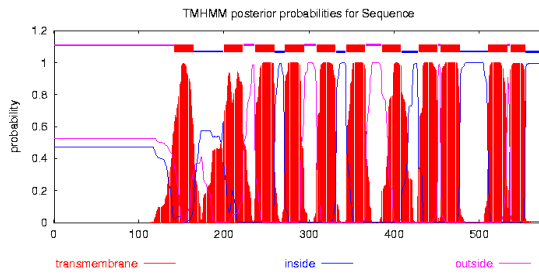
PA0438



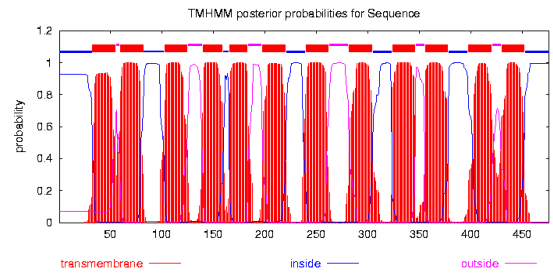
PA0443



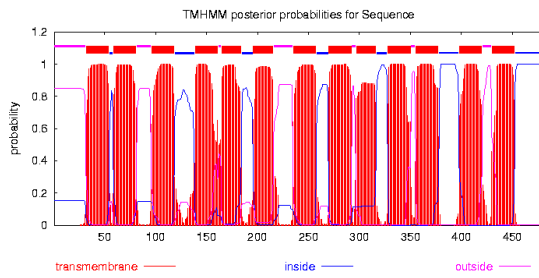
PA0476



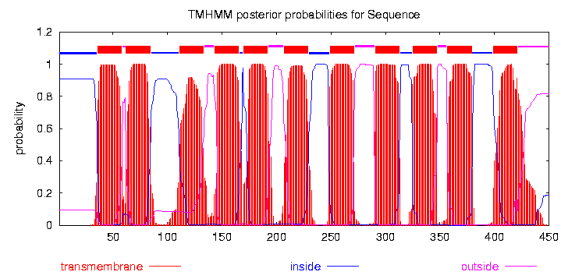
PA2073



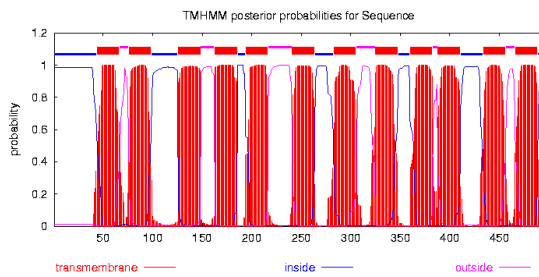
PA5099



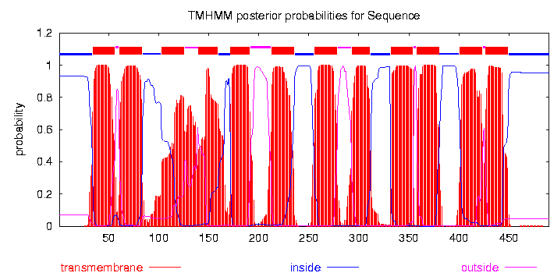
Pden0678



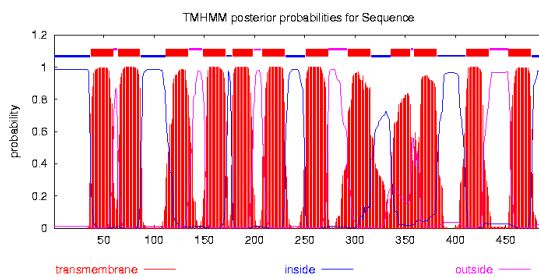
Pden1111



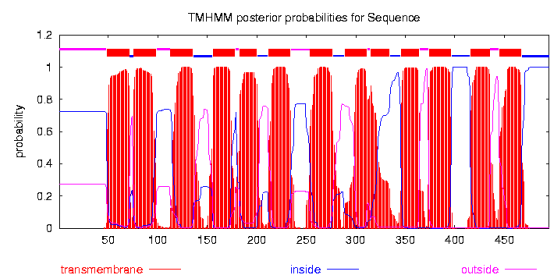
Pden4351



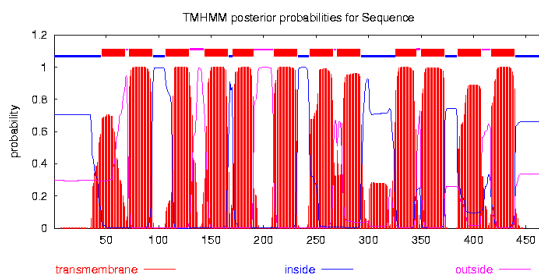
PPA0619



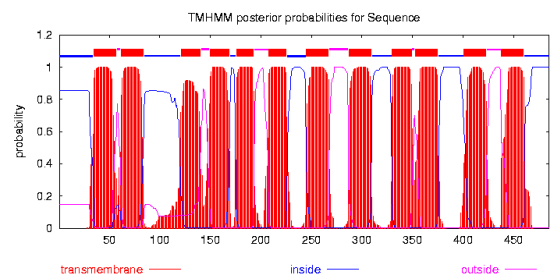
PPA2211



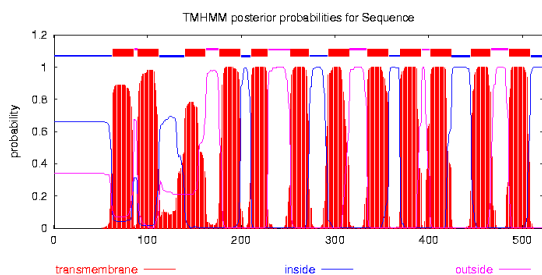
SCO0572



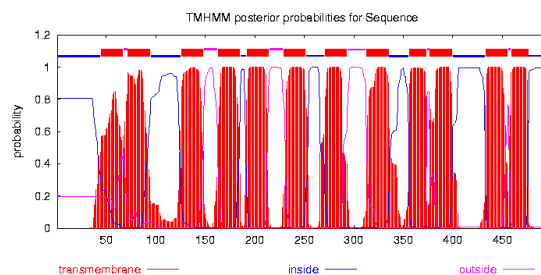
SCO5579



SCO6417



SCO7500



Appendix 2 Sequence alignment of Mhp1 and its homologues. Amino acid sequence of Mhp1 and its homologues were aligned using ClustalW (Thompson *et al.*, 2002; Larkin *et al.*, 2007) and similarity between the sequences were analysed using Geneious Pro 4.6.5 software. ■ 100% similar; ■ 80-100% similar; □ 60-80% similar; □ <60% similar. The residues that are involved in substrate recognition and cation coupling in Mhp1 are boxed in red. Note: Bsu3645 = Bsu36470 (PucI).

	1	10	20	30	40	50	60
Consensus				MAXESRXXLE			
Identity	■	■	■	■	■	■	■
PA0438				MSADSN			
b0336				MSQDNN			
SCO6417				MIGLPMTDTAPTAIPPSAQVTLDPGRVELAPGS			
Pden1111				MTQVEATAYEMLSRSGLPAD--DLDP			
PA0443				MQQSRSEVTERDGLVELSAGSDVLDS			
b0511				MEHQKLFQQRG			
EF3000				MEKNVSQVTAQEETALKARG			
Bsu3645				MKLKESQQQSNRL			
PA0476				MSSTPKPPASPREFSPSLLDLGLTLESLLDSLPRAGQRSAEFSASLLDLGTPASLDCLPRAG			
SCO7500				MTALEDRRSSDRTG			
PA2073				MSNNNEASRPKIE			
Bsu3867				MSNNNEASRPKIE			
PPA2211				MVSAAPNSAQDEVHRGNTAPTASGAGRGVE			
PA5099				MTQGS DSKPLIE			
Pden4351				MSMTERL DARAERS			
Mhp1				MSTTPIE EARSLLN			
Pden0678				MTTHEMAPDSAVIEHK			
SCO0572				MAMHKNPAGGAAAA EAATGAVG			
SCO5579				MSKTA ENEGALET			
PPA0619				MTQTATKTNATTHKG			
lp_3374				MQQESK			
Consensus						XNEDI LPVPAE ERTWGAWSL	
Identity						■	■
PA0438						YSQQPVPLAARKS	---ALAL
b0336						FSQGPVPQSARKG	---VLAL
SCO6417						PPPSGPHYANEDLLPVPVEGR	TWTTYNF
Pden1111						DLYNEDQLPPTVAERT	TWTWVSI
PA0443						PRYNHDIAPTRVEQR	TWNRWHI
b0511						YSEDLLPKQTQSQR	TWKTFNY
EF3000						YNEDLLPSSPKQR	TMGARNF
Bsu3645						SNEDLVPLGQEK	RTWKAMNF
PA0476						LRAAEFSAARLDLGLTPASLDCLPRASHHAGPELPLSPRLYSADL	APTKEGRRWGRYSL
SCO7500						DTAT	SSGLYTYDLAPTKEGRRWGAYNV
PA2073						VRSDYVPRH	ERHGVVHQ
Bsu3867						RRTEIYIPNE	ERHGKAKDL
PPA2211						KRTIDMVPDGE	ERHGTTPKSQ
PA5099						RRSIDYIPANE	ERHGRLYSQ
Pden4351						ALIEESILPIRAAQ	RPIGALGY
Mhp1						PSNAPTRYA	ERSVGPFFSL
Pden0678						AVGEESLAPQEA	RIMDPWSY
SCO0572						SDDYALSRVPR	RLRGFWTM
SCO5579						RGTIEQVPDH	ERTARTREL
PPA0619						IETTGIEIVKES	QRTAQPQDL
lp_3374						YQIEVVAPKH	RHMNSNWDN

Consensus Identity
) 130 140 150 160 170 180
 FALWVXGXNHNHATYXTGGGLIA-XGLSVWQALLAALLLGNLIXGXFMALNGXACPKYQJDP

PA0438 SVVMTGLLTFEFSASMWTTGGTGLT--GLSYDDEFLLAVLLGNLMLGLYTAALGFLICARTEEST
 b0336 TFVMLGLLTFEFSASMWTTGGTGLT--GLSYHDFFLAVLLIGNLMLGLYTAALGFLICARTEEST
 SCO6417 SALWVGMMAHNTASWTLASGLIA-VGMDWKQAVFTHALANLHVLAPMLLTGHACPKYQJDP
 Pden1111 SALWVGMVVCPTYLLASYLEIG-AGMDWKQAVITLLLANAHLVLIPLVLLTGHACPKYQJDP
 PA0443 TALWVGMVVCPTYLLGGLVLTAYFELSVGEALLAALLLANLHVLIPHTLNAFAGTRVYQJDP
 b0511 FTLLWVGMVHNPNYVAVGGFFI-LGLSPLQVMLAVVLSFFVATFVNLNGVAGSKYQJDP
 EF3000 ASLWVGMCIHNPTATVGGGLIA-LGLSPWQVLAHITASLHLLFGAMALNGHACPKYQJDP
 Bsu3645 FALWVNDVHNTANYSFAIGLFA-LGMSGAQILAAALGALHVVYGLMNLISGYMGLRTEQDP
 PA0476 FALWVNDVHNTANYSFAIGLFA-LGMSGAQILAAALGALHVVYGLMNLISGYMGLRTEQDP
 SCO7500 FTLLWVNDVHNTANYSFAIGLFA-LGMLNVWVILAAAFALASVHLLFLLMNLISGYMGLRTEQDP
 PA2073 APFVFTGNEVETTMVGTGTGPA-LGLGALYSILAAVVLGVCFGTFFMAAFHANQCPRLMTEQDP
 Bsu3867 FPLWVFGANHHTTLVGTGTPVA-MGLNLFWVMAAHLICGTLHGAIFMAHSAQCPRLMTEQDP
 PPA2211 FTLLWVGVNMQTAVVNGALAIV-FGADVVAALICHTFVIGQVHGGAVMAHSAQCPALMTEQDP
 PA5099 FTLLWVLANLQHTAIVTALAVV-LGSDVFWSLVGLLGLLGGAVMAHSAQCPALMTEQDP
 Pden4351 AWLWVGIAVIHTATVSLGATGIQ-GFVGLGTVALTLLANLHAGFMLLTADIGTEHGSIF
 Mhp1 AAIIWVFAAIQVAIEIAAGOMTS--SFQVQVIVAAHAGCTHAVILHGFSAATRWGAINF
 Pden0678 LFAWVLLGGCVSGHTETVGSGLV--TLLNLQTFVAHTIGCTVIGVAVMMNGAQCGRMTEQDP
 SCO0572 LLQWLLAQSGSLHQTTLGATIGV--GTMTEGDAFLAFTLGAHLLVAVVFAIGLACMRTEQDP
 SCO5579 FPTWVGANISVLLLTMGASLVVAYRIVFVSWQALVAVAGAAPVVSYGLVGLIAGCRGTEQDP
 PPA0619 FLPWVFNANVSLFGMSYGAFLMLG-FEVSFWQATAATLVGVIVSFGFCGIIAAGCRGTEQDP
 Ip_3374 FATWVIGANANGTWVYVGGVLA--CGLLTALKVIVASSTISYLCISLVGFMVYRTEQAT

Consensus Identity
) 190 200 210 220 230 240
 MVL SRAS - FGLRCAKLPALR - GMVAIGWVFGIQTXLGALAMNXLLGX - X - - - - -

PA0438 HLLAHYS-FGSKCSWTPESAIL-GGTQVGVWEGVGVAMFAIPMAKATG--
 b0336 HLLAHYS-FGSKCSWTPESAIL-GGTQVGVWEGVGVAMFAIPMAKATG--
 SCO6417 PVFARAS-FGLRCAKLPALR-ALVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Pden1111 PVLRLSS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PA0443 PVLRLSS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 b0511 AMILRAS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 EF3000 AMHLQST-YGSLCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Bsu3645 PVIIRAS-YGSLCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PA0476 PVMCRIA-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 SCO7500 PVMCRIA-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PA2073 PVMCRIA-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Bsu3867 MIIQSRAS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PPA2211 MIIQSRAS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PA5099 MIIQSRAS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Pden4351 AVYLRAP-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Mhp1 TVAARMP-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 Pden0678 MVQARSS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 SCO0572 PHLTRWAG-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 SCO5579 MALSRAP-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT
 PPA0619 MVL SRAS - FGLRCAKLPALR - GMVAIGWVFGIQTXLGALAMNXLLGX - X - - - - -
 Ip_3374 MSLIRGS-FGLRCAKLPALR-GIVACGWFVGIQTLWIGGEAHLFAGKLGIDGWT--NAGT

Consensus Identity
) 250 260 270 280 290 300
 - - - - - L P G W L X F L X F W A M Q V L I V F F C I E S I R R L E A I A A P L I I A V X X Y L X V W A M X X A G G S G

PA0438 - - - - - LDVNLILLIVSGAMTTLTVFFCIAGTLLSATAIVPAIVVLSVSVWLAVRDAGGLA
 b0336 - - - - - LDINLLIAVSGLEMTVTVVFFCISATVLSVIAVPAIACGLGGYSVWLAVNGMGGLD
 SCO6417 FGGHAWTMWLSFALEWATQMAIYIKRMTERRFENWAAPFVLVGAFLVWMLWADKAGGFG
 Pden1111 ILGINLTFELCELAFWGHTHLYFIKHTESTERWLETWSAPFLIAGLALLAWAVKAGGFG
 PA0443 ALG-GTGEVIGFELFWGHTHLYFIKHTESTERWLETWSAPFLIAGLALLAWAVKAGGFG
 b0511 LLGLSLPGLITFLIFLWLVNMGVFGGKVNKFTAILNPCIYVFGGMATIAWISLVG-IG
 EF3000 FFGLRRLPELMAETLEWLVNVAIGFGGSKINRFTAILSLIYVVIIGLTIWATRAGGGLT
 Bsu3645 ILGIHLSGLLSEFVFAWHLHLLHGMESIKRFEVWAGBLVYLVFVGGMVAWMDIAGGLG
 PA0476 ILGLSLLGWATFVAIWCVQLLIFAYEMEMVRRYVESFAGVILLTVAALAAWVYLRAGASI
 SCO7500 LLGQSTLGLWITFLVWAVQMLIVSYGMQMIRRYMAFAAPTLLTMCALAVWMEVRAAGSV
 PA2073 - - - - - HRAPVYV-LIGVAVLLAIVGMDLHTVQRWLTIVYVMIAGVGLVTVSAIRS-LEAD
 Bsu3867 - - - - - IPGSWSIIGLSAVCFLLTIFCHDLHKKMKILSWTSFAVFFAATLIFQLPIPAG
 PPA2211 - - - - - OSRWIGILVFAATTAIIAMFESYDLHKKLVKVSSTIGLLGLLYAVRFLQVSSFGP
 PA5099 - - - - - SD-SAGIFPAAFIIMLTVLGYRVIIHIGRIASVLGIIAFAYFESRLMLLSDIDG
 Pden4351 - - - - - FSNWFLWYAAFAAVQVANTMLGIRSVHRLAALAAPAIATIASVWVYFTELEGIAETKG
 Mhp1 - - - - - FTNLPLWVIFGAIQVMTTFYGITERRWMNVFASVLLAMGVYVYVYMLLDGADVSL
 Pden0678 - - - - - YSNWFAFFLGFQDQIGLSVLEFHGKIKWLENIGAVFLIGSLIYMFESVNRVYGAET
 SCO0572 - - - - - PSWLCCTAAGVAITVLLVIFSKYMAVFAKIVTLLFFMAVWSVTDALSDHGSFSE
 SCO5579 IRANSVLDWMTLAEFVATFAISGLGINAQKCNKYAAYLFGAFSLVLLGLYLVDTDWSA
 PPA0619 - - - - - SGNVVKICATLIVAFPIVGGAVAYHIIIMKLOAVLTVITGILTVIYLVMAVSHIDWHA
 Ip_3374 KPGGAKGLILGIEVMSVHLLSVSVQRSSVQMIERLGIILVILFVLWETVVVVFQTVSVHD

Consensus Identity
) 310 320 330 340 350 360
 XLLSXAPSEFXWG - - - - - ELVALALVVGFWXTLALNIXDETRVX - - - - - RSXRXQV

PA0438 ALQQVTPSAPLS - - - - - LSTAILALVVGSEFVSAGTLTADDEVREGE - - - - - RSARTAV
 b0336 ALKAVVPAQPLD - - - - - ENVALALVVGSEFVSAGTLTADDEVREGE - - - - - RNAKLAV
 SCO6417 PLFDQPSRLGWG - - - - - GDFWKLFWPSSLMGMIGFWSLTLNIEDEVREGE - - - - - KSQRAQT
 Pden1111 PMLSAPSQFAPGQPKGQFWSVWSPSLTGMIGYWATLALNIEDEVREGE - - - - - RSQKDFP
 PA0443 ELLAQPPKRPBG - - - - - ASVTGYEMAGLITAMVGFWATLSLNIIEDEVREGE - - - - - RSQKDOI
 b0511 PIFDYIPSGIQR - - - - - AENGGFLFELVINAVVAVWAPAVSASDEVREGE - - - - - HSREQA
 EF3000 PLYS-QVSGAIR - - - - - SVNPLVAMLIIFNSVAVWSAPGASVDEVREGE - - - - - RSTRAQV
 Bsu3645 PHYS-OPGKPH - - - - - FSETFWPFAAGMTGIIGIWTATLILNIEDEVREGE - - - - - ETQKEQI
 PA0476 AWSVSGPEMSSGSQ - - - - - MWLKI FGGAAFVWTTIYGTMLILNFCDEVREGE - - - - - PDRRTIS
 SCO7500 SMSVDAPLTTGGA - - - - - MWLQVLQAAALWVVIYGTFLVNFDEVREGE - - - - - RSRGSIV
 PA2073 AALGGTPAFSWS - - - - - AELVQLSAAAGYQISYAVYVSDYSRYLPHE - - - - - TPTRKVI
 Bsu3867 SWIPG--AIDL - - - - - IELVAMSAMATWQLAYAPYVADYSRYLPVK - - - - - TPASKTF
 PPA2211 QLTDV--HFSV - - - - - BELLSIALAAGWQLTYAPYASDYSRYLPVD - - - - - TKGSAW
 PA5099 LLAIR--HFSWA - - - - - SELLSAMLAASWQIAFGPYVADYSRYLPAS - - - - - TSSWKT
 Pden4351 LNIWTFRAEGQAS - - - - - LIVLFIANMSEWSTMAIDIEHLNREHVHTR - - - - - TGTRSPFL
 Mhp1 SEVMSMGGENPG - - - - - MPESSTALMIFMGGWIAVVVSIHDTVRECKVDPNASREGQT
 Pden0678 GANLINAEGTWG - - - - - APFWGGTMLFLGVYSTMMLNVSDEVREGE - - - - - VGRMRQV
 SCO0572 LIHSPAPGETIP - - - - - LAVATAIAGGFMTGAIVSEBMRVYRN - - - - - RKSHVIF
 SCO5579 VFDRSAGSVAAM - - - - - ITGMGLIAGGVSVSWIPSADEVREGE - - - - - AASKGIV
 PPA0619 ATSLPAASFPAFVG - - - - - TLLTLMTGTGLGWNTIAADWSRYQSRT - - - - - SPGFAIA
 Ip_3374 LLTWQAPTHLHMT - - - - - AGAGMDTLAARNLSWVTAGADEVREGE - - - - - RKKSGAT

0 370 380 390 400 410 420

Consensus Identity
 LGQXLGLPGLMFLFAXLGVVLAALAA-X-SADGEGIWDPVXIITAXID---GVXLLVLAALMLV

PA0438 VVCLVAFETLGNLSLMFIFGAAGAAA-VGKSD-----ISDVM---IAQGLLIPAILLV
 b0336 LVAMVAFETLGNLSLMFIFGAAGAAA-LGMAD-----ISDVM---IAQGLLIPAILLV
 SCO6417 RQALGLPTMTLFAFLSMMMTSG-SQAVYGEPIWDPVQLAAKTD---NVGGLLFAALVTV
 Pden1111 LGQLIGLPLPMALFAFISAAMTSA-TIVYIYGEAIWDPVQLTERMG---GVGLLVALMALV
 PA0443 LGQIFGLPLTMFLFAALGVVMTAA-SASLVGETVSDPVSLTGHIP---SPGWAVAMALIV
 b0511 LGQTLGLVVAAYILFAVAGVCIITAG-ASIHYGADTWNVLDIVQRWD---SLFASFFAVVVI
 EF3000 VGGTTGLVVGYGIFAFSSVVLILG-GSLYFGIQEWNILNIDRLD---NVAVVVLAMSVFI
 Bsu3645 KGFYGLPGTEALFAFASITVTSAG-SQVAFGEPIWDPVLDIAREP---NPYVIVLISVITL
 PA0476 VGNFVGLPMNILLFGLFAMVLAGA-QFRIDGQLIQSPDTIMAAIP---DTLFLVLASIALV
 SCO7500 RGNVIGIPMNMLFEAVIVAVLSGA-QFTLDGRVITSPTDIMRTIP---NMFLAVASACL
 PA2073 FWTYLGAAAGSALWMLSLGAFLLASA-LPAPDAIGSVREMGNRLYPG---FGTFTVLIAPVPA
 Bsu3867 WYSYAGTSSVSSIWMLLGCALTTS-LPDTFANS-----GSQIVQL---FGPFSFIMLIVM
 PPA2211 FGFAGSVLGAATISMTLGVFLAAA-GGQAFIGDQVGFMGNAQDN---RVLAILLVFSSIV
 PA5099 LAVGLGSVVGAAQSMVLGVFSAAL-ANGQFAGREVAYIVGLGSS---GTVAALLYFSTIV
 Pden4351 ERNRSIFLAQLVALPATQAMITAGIGAISFIATGNWNPVEVITQGD---QGLALLLALLVLMV
 Mhp1 KADARYATAQWLGMVPAISIFGFIGAASMVLVGEWNPVIAITTEVVGGVSIIPMAILLVQVFI
 Pden0678 VIYANSILPATLFMGLIGLIMVSSA-----TGEVDPKVFSSAVD---NKVLLVVTTLVFI
 SCO0572 LQSASSMILSEYIVGLTGVLLGHL-----VGSQVSVQIMVLSST---GVGGLVVMMS
 SCO5579 GVTVGGAGIVVLPVLMGAVMAVS-----TPDLASAADPMSFAGEILP---TWIAMPYLLIAI
 PPA0619 FWNVFGASLPLVILITGLLLAAS---SKNLSDAIGADPIGALATILP---TWFLVPLVLAAL
 Ip_3374 GMPFLGASTPLVFWFAFIGLAAATIS---IAITSGTYDPNNSDPSSTIAS---RLGLGLVALMLV

0 430 440 450 460 470 480

Consensus Identity
 IVAATISSTNKAANLVA PAYALT NJFPXR---ITFKRGVLLXGVGLIXPWXLNSP---

PA0438 IGLNIVTTNDMIALYASGLGFANITG-----LSSRFLAMINGAIGTAAALWLYNHVFG--
 b0336 IGLNIVTTNDMIALYASGLGFANITG-----MSSKTLVSVNGIGTICALWLYNNFVFG--
 SCO6417 IVATISVNIANLMSPADEFNSVAPRR---VSRAGALNTSVIVAVIFPDKVYSDP---
 Pden1111 IVAATITNLAANVVA PAHGFSNIIAPSR---INLRGGYITAAIGLAMPFWLLIN---
 PA0443 IVAATITNLAANVMSPTNDFQNIAPKL---IGRTRAVWITGFEIGLALMAHEKLLKLGWI
 b0511 IIMTITSTNATGNIIIPAGYQIAAIIAPTK---LTYKNGVLAASITSLICPKIMENQ---
 EF3000 IIMTITSTNATGNIIIPAGYQIAAIIAPTK---MTYKKGVMASVSPFIMPWKIMENA---
 Bsu3645 ICTATITSVNVAANVMSPAYDIANALPKY---INFKRGSFTALVIALEFVTPWKIMESA---
 PA0476 IIVTVAVNIMANVVA PAYVLTNIIAPNL---LNFRRAGLISATIAVILPWHIYNSP---
 SCO7500 IIALTVAVNLLANVVAPIYALVNIIPFHR---LDFRRAGLVSAAIGLITPNIYNSP---
 PA2073 IIVGIMAVNICY---GAMLTSLSAIDAEFR-KVTPTLNLRVTGIGVIALVVEGVAISIPES---
 Bsu3867 IIFGQIAINVFNLYGAFMSTTTTTEPFL-KLKVTPKVRILMILGLVTLVGTVLSLQGSN---
 PPA2211 IIVGKITINTLNAYGGVMALSTAGSFTGKSSVSPKTRFGFVIGENVVVLTAIAAASD---
 PA5099 IAFGKVTIS TLNSYGSFMCIATIISSVRGQLEITRVQRLFVLAIVGAS TLIAALGQHS---
 Pden4351 IIAAQWSTNNSANLIPAAALTFTVNIAPRV---ISYRGGVAIVAGVGTICPEVQLNDN---
 Mhp1 IIAAQWSTNNSANLIPAAALTFTVNIAPRV---FTFKTGVIVSAGVGLIMMPPQFAGVNLN
 Pden0678 IIAFAQTITNLLNVMPPAYALMDVFKIN---FKTSAVVGLGIAEAFVTPFVWKVDES---
 SCO0572 TAKINDWNLVYGSSTLGVVNFQVVFGRK---VHRGAVTIVLGIAGTFLS AVGLMHTFT---
 SCO5579 IIVGMLINISMMSMYSAGFTAQTGLGKVP---RHWA VSNVAVISLVEGGVLSLVTS---
 PPA0619 IIVSLIAGALNGIYSSGLTLLTLGIRVP---RPAASLVADETLLTIGLTVLVVFPVAPN---
 Ip_3374 IIVLTSMTANAVNLAAGSALSNIIEPRIR---LRPALWAVTIIATVTVTLIPVIMGSF---

0 400 500 510 520 530 540

Consensus Identity
 -----XF-XRGLLGAFLGEXAGVMMADYYLVRRRRXXDIPALVYXPXGRYGYWX---G

PA0438 -----WTFSSAANPPIGGVIIADMLSRRRR-----VRQHPCADIQA-----
 b0336 -----WTFSSAANPPVGGVIIADMLMNRRR-----VEHFATTRMMS-----
 SCO6417 -----QGYIFTVLGLVGGLLGTIVAGLIIADYYLIRRRARLHVVDVIRRRGRYVYEG---G
 Pden1111 -----HIVDLVAVFSALGGIAGVMMADYYLIRRGTRLSVDPIDRRGGIYEGSG---G
 PA0443 -----VSDLSLESVYSNVLGYSSLGGPIAGIMVADYYLVRRQRLDIAGVYR-DDVYV---A
 b0511 -----DSIYLELDIIGGMGPIVGVMMAHYEVVMRGOINDEHYRTPAGDYKYD---NG
 EF3000 -----DSIFIEINAVGAVIGEVAGVMMAHYEVVCKQCIDINAVYVDKHKKEANPFYG
 Bsu3645 -----TSVYALVGLIGGMIGEVAGVMMADDFIIRKRELSVDDIYSETGRVYVYK---G
 PA0476 -----GVILYRGGGALIGPLVYGLIIMVADYYLVRRGRVNVNPEIYETESRAGAYHYAR-G
 SCO7500 -----VVVNYRGGGALIGPLVGVVIMADYYLVRRSRVNVNPAIYTEDAGAEYHYRR-G
 PA2073 -----YLGSENTVLLMLLYFVVEWTAIVNIVDYYFVVRNNGHYAISEEINPEGIGYHWG---
 Bsu3867 -----FMELFLNRIFFNFSYFVEIWTAINIVDYYFVVRNNGHYAISEEINPEGIPYKVN---
 PPA2211 -----FLQVFKSRLTTLMLFFLEWSSAINIINLYLRCRCRIDPAIYTPRGRYGLKLH---
 PA5099 -----FLAAFKSRLTFLTFTTSAVAVNIVDYYFVITRERYDPAIYADPDGRYGRWN---
 Pden4351 -----LMEIRGGYGAFFSAGIGVMMADYYHAIIRRRRVNVPALIDHGGYRYFES---G
 Mhp1 -----TEINLIASAAGGIIAGIMISDYYELVRRRIRISIHDIYRTRKGIYTYWR---G
 Pden0678 -----AAGLSLEIRTYSAFIIIFALLIVDYYVIRKQQLNIEKIDPECGPYKGVN---
 SCO0572 -----DYSVGVVAVIIPVGGIIVAEVWVVRMRAPDADRDAETRLPATSP---
 SCO5579 -----FMGSEIARVSLVAVAFSAVWGVFGADMLRGRREYDGEALADTGRTSAYWYH---G
 PPA0619 -----FISPEQSRVTVGVPRISGWTGIMMADIT-IRRRRYPYDEADIDNGSRGHPDPISI
 Ip_3374 -----LAAFTAIRVDYIGMVIGPMISVMMVADYYWRHRHQHYDINNEIASSKGGQYVYHHG---

0 550 560 570 580 590 600

Consensus Identity
 -NWXATAAFALGXAJSLXGAXLPP-----XLGXADISVIFVGLAXAGXLYYVIMKRL

PA0438 INWAAITLAVACG---TAAGHWLPG-----IVPVNAVIGAAIAYLLNPLLSRS
 b0336 VNWVAITLAVACG---IAAGHWLPG-----IVPVNAVIGGAAISYLLNPLLNLRK
 SCO6417 WNWRAVLAFLVGGVLA MGGA DFHPLVDGRPVPPFLEPLADYGA VGLATS YLLNPLMLLP
 Pden1111 NNWAGTIGALIGLIPNLIGFMVAIGLSSSAPALFMSIYTYAVNEFVGLFVAGIAYLVSKLL
 PA0443 WNWPGFAAFAMPALTMALGNRQ-----FHWFYDYGNEFVGLVGGALYVGCGRR
 b0511 FNLTAFSVTLMAVILSLGGKFIHF-----MEPLSRVSEFVGVGAFMAAYALVKKRT
 EF3000 LNKPAYVATILALVLSLGGQFIPQ-----VKIIADIISFVGVATGFAVYLLVKKWT
 Bsu3645 YNYRAFAATMIGALISLIGMYVVP-----LKSLYDISFVGVLSLSEFYVIMRVH
 PA0476 VNPRATAAFVPAATIS TLLALLPA-----FAELSPFSFVFGAGLGLIHYAAKRR
 SCO7500 YNPRATAAFVPAATIAMVVALVP-----FHAAAGSFVFGATISAAVLYALVADRA
 PA2073 RRGESAYLAGLILMPPFMSLSFYSG--PFAVALGGADVAVVGLLVAGAVYAFMCRAL
 Bsu3867 -WITIAFVLSLLEIPFINTS FYIG--PLAKMFGGDVIAIITGLAVPSLVLYVMPKR
 PPA2211 -WPTIVVYLLGLVVMQIPFISQAFYVG--PVAAAMGGADISLIVGLVLTALVYVFPGRR--
 PA5099 -LPGIAVYTLGVLVQMPFLATLTYTG--PMVEHLGGVDISMLIIGLVLPALVYLLVARRK
 Pden4351 VNPAGMVAWLMAGGIIA WWSAYAFVGLFPLGFLLYFLMRITVIVPQYRQEEADRTRGDDE
 Mhp1 VNWVAIAVYAMALAMSFLT PDLMFVTGLIAALLLHIPAMRVIAKTFPLRESAEARNEDYL
 Pden0678 -MAAVIAMGMGIVFALIFS-----GISIYASLLPSGLLYLYLMOQHL
 SCO0572 -TWVPMSLVIWAAAFCKMGKFDYDGG-----IPALNSLITAVIVYCVVGLAGWIRPYGT
 SCO5579 FSPAAVVAWAGLAAGLMFTTSDWFTGPLATDNFVGEYGLGVAVTVVLSATLVAVMPKPA
 PPA0619 IITFVIMTVIGWGLVNTYIEGVNWNWDWQG-----FLGPIGLGGRREGDWAHANIGVFGALV
 Ip_3374 VNWVIALLCWVIGMAIFLLKHHVV-----IATVTAATFDMAIGMIVYVAMRIF

Consensus Identity	0	610	620	625
	AXAPXXRAXVX			
PA0438	PNKPEEVLHAD			
b0336	TTAAMTHVEANSVE			
SCO6417	ATRPGTDAAGSRNRSA			
Pden1111	NR			
PA0443	ARSPAALAKPLP			
b0511	TAE----KTGEQKTIG			
EF3000	WDS----KVKKETAYQEGK			
Bsu3645	PPASLAIETVEHAQVRQAE			
PA0476	TAYREVSGEAIAVDSVQH			
SCO7500	.SPFRDVDGESIAVAEE			
PA2073	DLEAERRLATLGERLLEEGGA			
Bsu3867	----LKKRAGYQEKLSLL			
PPA2211	-AFDYPLEIVYPKDVELDPSIVSR			
PA5099	ATRQAPARMILPADPGAFEHPAHGA			
Pden4351	LAASVGMDDWVHVAGGRFVRKPCGAA			
Mhp1	RPIGPVAPADESATANTKEQNQR			
Pden0678	PSAQREREN			
SCO0572	ALDDTDEPTPAATGTATVVGAP			
SCO0579	.LIVPDARAERAEQPEAETVTV			
PPA0619	GYVVTLIARRGTVRRQESR			
Ip_3374	YPLPAKRIV			

Appendix 3 DNA sequences of the recombinant NCS1 clones. The N-terminal MNSH (blue) and C-terminal His₆-tag (red). Data for EF3000, Pden1111 and Pden4351 are not shown.

b0336 (Note: silent mutation in b0336 which did not affect encoded amino acid sequence, highlighted in yellow)

ATGAATTCGCATATGTCGCAAGATAACAACCTTTAGCCAGGGGCCAGTCCCAGTCCGGCGCGGAAAGGGGTATTGGCA
TTGACGTTTCGTCATGCTGGGATTAACCTTCTTTCCGCCAGTATGTGGACCGGCGGCACTCTCGGAACCGGTCTTAGC
TATCATGATTTCTTCCTCGCAGTTCTCATCGGTAATCTTCTCCTCGGTATTTACACTTCATTTCTCGGTACATTGGC
GCAAAAACCGGCTGACCACTCATCTTCTTGCTCGCTTCTCGTTTGGTGTAAAGGCTCATGGCTGCCTTCACTGCTA
CTGGCGGAACCTCAGGTTGGCTGGTTTGGCGTCGTTGGCGCATGTGGCATTTGCCATTCCGGTGGTAAGCAACCGGCTG
GATATTAATTTGCTGATTGCCGTTTCCGGTTTACCGGTTTACCGGTTGATGACCGTACCGTCTTTTTTGGCATTTCGGCCTGACGGTT
CTTTCCGGTGATTGGCGTTCCGGCTATCGCCTGCCTGGCGGTTATTCGGTGTGGCTGGCTGTAAACGGCATGGGCGGC
CTGGACGCATTAAGCGGTCGTTCCCGCACACCGTTAGATTTCAATGTCCGCTGGCGCTGGTTGTGGGGTCATTT
ATCAGTGCGGGTACGCTCACCGTGCATTTGTCCGGTTTGGTCCGAATGCCAAACTGGCGGTGCTGGTGGCGATGGTG
GCCTTTTCTCCTCGGCAACTCGTTGATGTTTATTTTCGGTGCAGCGGGCGCTGCGGCACTGGGCATGGCGGATATCTCT
GATGTGATGATGCTCAGGGCTGCTGCTGCCTGCATTTGTGGTGTGGGGCTGAATATCTGGACCACCAACGATAAC
GCACTCTATGCGTCCGGTTTAGGTTTCGCCAACATTTACCGGGATGTCGAGCAAAAACCTTTCCGGTAATCAACGGTATT
ATCGGTACGGTCTGCGCATTATGGCTGTATAACAATTTTGTCCGGTGGTTAACCTTCTTTCCGGCAGCTATTCTCTCCA
GTGGGTGGCGTGATCATCGCCGACTATCTGATGAACCGTCCCGCTATGAGCACTTTGCGACCACGCGTATGATGAGT
GTCAATTGGGTGGCGATTCTGGCGTCCGCTTGGGGATTGCTGCAGCCACTGGTTACCGGGAATTGTTCCGGTCAAC
GCGGTATTAGGTGGCGCGCTGAGCTATCTGATCCTTAACCCGATTTTGAATCGTAAAACGACAGCACAATGACGCAT
GTGGAGGCTAACAGTGTGCAACATCATCATCATCATCAT

b0511

ATGAATTCGCATATGGAACATCAGAGAAAACCTATTCCAGCAACGCGGCTATAGCGAAGATCTATTGCCAAAACGCAA
AGCCAGCGGACCTGGAACAATTTAACTATTTTACCTTATGGATGGGTTCCGGTTCATAACGTTCCCAATTATGTGATG
GTCGGCGGCTTTTTTATTCTCGGCTTGTCTACCTTTAGTATTATGCTGGCAATTATCCTCAGCGCCTTTTTTATTGCC
GCGGTAATGGTATTAACCGGTGCTGCGGGCAGTAAATACGGTGTGCCTTTTGCCATGATCCTGCGTGCCTTCTTACGGT
GTACGTGGTGCATGTTTCCCGGATTAATAAGGGGCGGAATTCGCCCATCATGTGGTGTGGTTTGAATGTTACCGG
GGTCACTGGCCTGCTTGAATCTGATTGGCAAAAACCTGCGCCGGATTTTAACTCTCGGTGGTGAATTTCACTCTGTTA
GGCCTTTCTCTACCGGGCTTAATTACTTTCTTAATCTTCTGGCTGGTCAACGTTGGTATAGGTTTTTGGCGGTGGCAA
GTTTTAAATAAATTAACCTGCACTTCTTAACCGGTGCATCTATATCGTTTTTCCGGCGGATGGCGATTTGGGCGATTTCA
CTGGTCCGGATCGGTCCAATCTTTGACTACATTCAGAGCGGATTCAGAAAGCAGAAAACGGTGGCTTCTGTTCCTG
GTGGTGATTAACGCGGTAGTTGCGGTCTGGGCGGCACCGCGGTGAGCGCATCCGACTTTACGCAAAACGCCCACCTCG
TTTCGTGAGCAGGCGCTGGGGCAAACGCTGGGTTAGTTGTGGCTATATTCTGTTTGGCGTCCCGGGGTATGTATT
ATTGCCGGACCGATTTCACTACGGCGTGATACCTGGAACGCTGGATATTGTTACGCTTGGGACAGCCTGTTC
GCCTCGTTCTTTGCGGTACTGGTTATTCTGATGACAACATCTCCACTAACCGCAGCGGTAATATTATTCCAGCGGTT
TATCAGATTGCCGCCATTGCACCGACAAAACGACCTATAAAAACGGCGTACTGATTGCCAGTATTATCAGCTTGCTG
ATCTGCCCGTGGAAATTAATGGAAAATCAGGACAGCATTTATCTTTTCTCGATATTATCGGCGGAATGCTTGGTCCG
GTAATTGGTGTGATGATGGCGCATATTTTGTGGTGTGCGCGGCAAAATTAATCTTGATGAACGTGATACCGCACCT
GGCGATTATAAATATTACGATAACGGTTTTAACCTCACTGCGTTTTTCAGTAACCTGGTGGCGGTTATTTTATCTCTT
GGCGGTAAGTTTATCACTTTATGGAACCGTTATCGCGTGTTCATGGTTTGTCCGGCTCATCGTCGCTTTGCGGCC
TACGCTTATTAAGAAACGTACAACAGCAGAAAAACAGGAGAGCAAAAAACCATAGGTGCTGCGAGCGGTCGTGGC
AGCCACCATCACCATCACCAT

Bsu36470

ATGAATTCGCATATGAAATTTAAAAGAGAGTCAGCAGCAATCCAACAGGCTGAGCAATGAAGATCTGGTGCCTTTGGGA
CAGGAAAACGGACGTGGAAAGCAATGAACTTTGCCTCCATTTGGATGGGATGTATACATAACATACCGACCTACGCG
ACCGTGGGAGGATTAATTGCGATCGGCCCTTTCCGCTTGGCAGGTGCTGGCGATCATTATTACGGCATCACTTATCCTA
TTTGGCGCTCTTGCCTTAAACGGCCATGCGGGGACAAAATACGGGCTGCCGTTTCCAGTGATCATTCCGGCTTCTTAC
GGGATATACGGCGCGAATATCCCCGCGCTTCTAAGGGCGTTTACAGCTATCATGTGGCTTGGCATCCAGACCTTTGCG
GGTAGCACGGCACTGAACATTTTGCCTTTTGAATATGTGGCCAGGTTGGGGAGAAATGGCGGGCAGTGGAAACATTCTC
GGCATTCACTTGCCGGTTTGCCTTTGTATTCTTTGGGCCATTCAATTTACTCGTATTGCATCACGGCATGGAG
TCGATTAACGGTTTTGAGGTGTGGCAGGGCCTTAGTGTATCTGGTATTTGGCGGCATGGTATGGTGGCCGTTGAT
ATTGCTGGAGGATGGGCCCGATATACTCTCAACCAGGAAAGTTTCATACGTTTTTCAGAAACATTTCTGGCCGTTTGC
GCCGGAGTTACTGGCATCATCGGCATCTGGGCGACATTGATCTTAAATATACCTGATTTTACACGGTTTGGTAAACA
CAAAAAGAGCAAAATCAAAGGACAATTTTACGGTTTGGCGGAAACGTTTGCCTGTTCGCATTCCGAAGTATTACGGTG
ACCTCCGGCTCGCAGGTCGCGTTCGGAGAGCCGATTTGGGATGTGTCGATATTTGGCGCGGTTTGATAACCCCTTAT
GTCATCGTCTTGTCCGTGATCACACTCTGCATCGCCACGATCTCTGTAATGTGCGGGCGAATATCGTATCACCCGCT
TATGATATAGCGAATGCCCTGCCGAAATATATCAATTTCAAACGTGGCAGCTTTATCACAGCGTTGCTCGCTTTATTT
ACGGTTCCGTGGAGCTGATGGAGAGCGCGACAAGCGTGTATGCGTTTCTTGGCTTAATAGGCGGCATGCTTGGACCG
GTGGCAGGCGTGATGATGGCCGACTATTTTATCATTTCCGAAACGTGAGCTCTCGGTAGATGACCTGTATTCCGAAACA
GGACGGTATGTGATTGGAAGGGCTACAATTACCGTGCCTTTCAGCCACAATGTTAGGAGCGCTGATTTCCGCTGATT
GGCATGTATGTTCTGTATTGAAGAGCTTATACGATATTTCTTGGTTTGTAGGTGTGCTGATTTCTTTCTTTTCTAC
ATTGTTCTGATGCGTGTTCACCCGCTGCATCATTTGGCAATTGAAACAGTTGAACATGCGCAGGTCGCCAGGCTGAA
GCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT

lp3374 (Note: silent mutation in lp3374 which did not affect encoded amino acid sequence, highlighted in yellow)

ATGAATTCGCATGTGCGCAAGAGTCAAAGTATCAAATAGAAGTGGTGGCGCCTAAGCACCGACACATGTCTAATTGG
GACATGTTTGCACCTTGGATTGGCGCCAATGCCAATAATGGCACGTGGTACGTTGGTGGTGTATTGGCTGCGTGTGGT
TTATTGACGGCGCTTAAAGTGATTGTAGCATCGTCCACGCTATCGTATTTATGTTTATCGTTAGTCGGCTTCATGGGG
TATAAGACCGGGCGGCGCAGCATGAGCTTGATTCTGGTTCCTTTGGTGTGCGGGGACAGTTACGTTCCGCTCCTTTGTG
AACTTGACCCAGTATATTGGCTGGACAGCGGTAACACAGTTCATTGCAGCAACTTCTGTGAGTATTATTATACAGAT
CTAGTTGGCTGGCCAGTCTATGGTAAACCCGGTGGGGCTAAAGGACTGATCCTCGGAATTATTGTGATGAGTGTGTTG
CATTTATTAGTGTTCGGTCCGTTCAACGGTCCGTTTCCAGATGATTGAGCGTTTAGGTATTATTCTTGTGATCTTATTT
GTTTTATGGGAGACGGTCCGTCGTGTTCCAGACGGTCTCAGTGCATGATTTGTTAACGTGGCAGGCGCCGACACATTTA
CATATGACGGCGGGTGTGGGATGGATACGCTAGCGGGCTTCAATTTGTCTTGGGTGACAGCGGGTGTGATTTTACG
CGGTTTACGCGCAAAAAGAGTGGCGCGACTGGCATGCCATTTTTAGGGGCCTTTACCGGCTTGTTCGGTTCCGCTTT
ATCGGGCTGCCGCGACGATCAGTATCGCGATTACTTCTGGAACCTATGATCCTAATAACTCTGACCCGAGTACCATT
GCGATGACGACTAGGCTTGGGGTATTAGCACTACTGGTATTGTTCTTAACGAGTATGACGGCGAACCGGGTAAACTTA
TTGGCGGCTGGTTCCGGCGCTGTGCAATATCTTCCACGGATTCCGGTTACGACCAGCACTGTGGGCCGTGACAATCATT
GCCACGTTAGTCACTGATTTCCATTAATCATGGGAGTTTTTTAGCTGCATTACGGCGTTTCTAGATTATATTGGG
ATGGTGTGGACCGATGATCAGCGTGTGTTAGTCGATTACTATTGGCGGCATCGTCAGCACTATGATATCAACGAA
CTGGCGAGCTCTAAGGGGACGATTTGGTATCATCATGGGGTGAACGTGGATTGCGCTATTATGTTGGTACTAGGTGTG
GCCATCTTCTGTTATTGAAGCACGTTGTTTGGATTGCGACGGTACTGGTGAACCTTTTATTGACATGGCGCTTATC
GGGTCAATTTACGTGGCTTTAATGCGAATTTTCTACCCGTTACCTGCCAAACGTGTC**GCTGCAGGCGGTCTGGCAGC**
CACCATCACCATCACCAT

PA0438

ATGAATTCGCATATGTCCGAGACAGCAACTACAGCCAGCAACCGTCCCGCTGGCGGCGCGCAAGAGCGCGCTGGCC
CTGAGCGTGGTGTGCTCGGCCGACCTTCTTTCCGCCAGCATGTGGACCGGGCCACCCTCGGCACCGGCCTGAGC
TACGACGACTTCTTCTCGCCGTGCTGCTGGCAACCTGATGCTCGGCCTGTACACCGCGATCCTCGGCTTCATCGGC
GCGCGCACCGGCTTTCCACCACCTGCTGGCGCATTACTCGTTCCGCGAGCAAGGGCTCCTGGCTGCCCTCGGCGTTG
CTCGGCGGCACCCAGGTCGGCTGGTTCCGGCGTCCGCGTAGCGATGTTCCGCATCCCGGTGGCCAAGGCCACCGGCCTC
GACGTCAACCTGCTGATCCTGGTTTCCGGAGCGCTTATGACCCACAGGTGTTCTTCCGCATCGCCGGTCTGACCCCTG
CTGTCCGCCATCGCGGTCGCCGGCATCGTCTGTTCTCGGCAGCTACTCGGTATGGCTCGCGGTGCGCGACGCCGGCGGC
CTGGCGGCGCTGCAGCAGGTGACGCCGAGCGCGCGCTGAGCCTCTCGACCCGATCGCCCTGGTGGTTCGGCTCGTT
GTCAGCGCCGGCACCCCTCACCGCCGACTTCTGCTCCGCTTCCGGCGCAGCGCGCGTACCGCCGTGGTGGTGTGCCTGGTG
GCGTTCTTCTCGGCAACTCGCTGATGTTTCTTCCGCGCGGCGCGCGCGGCGGGTGGCAAGTCGGACATCTCC
GACGTGATGATGCCCGAGGCGCTGCTGATCCCGCGCATCTGGTCTCGGCCTGAACATCTGGACCACCAACGCAAC
GCGCTGTATGCCCTCCGGCTCGGCTTCCGCAACATCACCGGCCCTCCAGCCGCTTCTGGCGATGCTCAACGGCGCC
CTCGGCACCCCTCGCCGCGCTCTGGCTGTACAACCACTTCTGCTGGCTGGCTGACCTTCTCTCCGCCCATCCCGCCC
ATCGGCGGGGTGATCATCGCCGACTACCTGAGTCGCCGCGGCGCTACCGCCAGCACCCCTGCGCGGACATCCAGGCG
ATCAACTGGGCGGCGATCCTCGCCGCTGCTTGGCGCACCGCGCGGCCACTGGCTGCCGGGGATAGTGCCGGTCAAC
GCGGTGCTCGGCGCCGCATCGCTACCTGTTGCTCAATCCCTGCTGAGCCGACGCCGAACAACAGAGAGGT
CTCCATGCTGAT**CATCATCATCATCAT**

PA0443

ATGAATTCGCATATGCAACAGAGCAGATCGGAAGTGACCGAGCGGACGGCTTGGTCGAGCTGTCCGCCGGCAGCGAC
GTACTCGACAGCCCCGCTACAACCACGACATCGCCCCGACCCGGTGGAACAGCGTACCTGGAACAGGTGGCACATC
ACCGCGCTGTGGATCGGCATGTCCATCTGCGTGCCACCTACACCCTCGGCGGGCTCCTCACCGCTACTTCGGCCTC
AGCGTCGGCGAAGCGCTGCTGGCGATCCTGCTGGCCAACCTGGTGGTGCTGATCCCGCTGACCCCAATGCCTTCGCC
GGCACCCGCTACGGCATTCCCTTCCCGGTGCTGCTGGCGCTTCGTTTCGGCATCATCGGCTCCAACGTGCCATGCCTG
ATCCGCGCCGTGGTCGCGTGCGGCTGGTTTCGGCATCCAGACCGTGTTCGGCGGGCTGGCGATCCATCTCTTCCCTCGGT
TCGATCTTTCCCGCTGGAAAGCCCTCGGCGGCACCGGCGAGGTGATCGGTTTCTCCTGTTCTGGGGGTGAACCTG
TGGGTGGTGCTGCGCGGCGGAGTCCGATCAAATGGCTGGAGACACTCTCGGCACCGCTGCTGGTGGTGGTGGCCTC
GGCCTGCTGGTGTGGGCGCTGCCGCACATATCGGTCAGCGAACTGCTGGCGCAGCCGCCGAAGCGCCGGAAGCGCC
AGCGTCACCGGCTACTTCATGGCCGGGTGACGGCGATGGTTCGGCTTCTGGGCCACGCTGTGCTGAAACATCCCCGAC
TTCAGCCGCTATGCGCGTAGCCAGAAGGACCAGATACTCGGGCAGATCTTCGGCTGCCGCTGACCATGTTCTGTTT
GCCGCCCTCGGCTGGTGTGATGACCGCCGCTCGGCTTCGCTGGTGGGCGAGACGGTTTCCGACCCGGTCAGCCTGATC
GGCACATCCCCAGCCCCGGTGGGTGGCGGTGGCGATGGCGTTGATCATCGTCCGACCCCTCTCGACCAACCCGCG
GCGAACATCGTCTCGCCGACCAACGACTTCCAGAACATCGCACCGAAACTCATCGGCCGACCCCGCGCTGCTGGCTC
ACCGTTTTTCATCGGCTGGCGCTGATGGCCACGAGTTGCTGAAGAAGCTCGGCTGGATCGTTTTCCGACCTCGGCTG
GAAAGCGTCTACTCCAACCTGGCTGCTCGGCTACTCCAGCCTGCTCGGGCCGATCGCGGGGATCATGGTGGTGGACTAC
TTCCTGGTGGCAGGCAACGCCCTGATCTCGCCGACTCTACCGCGACGACGCTTACCCGGCTGGAACCTGGCCCGGC
TTCGCCGCTTCGCGGTGCCGGTGGCGCTGACCGTATGGCCCTCGGCAACCGCCAGTTCCACTGGTCTACGACTAC
GCTGGTTACCGGCTCGCTGCTCGGCGCGCGCTCTACTACGGCTCTCGGCCGCCGTGCGCGGAGCCCGCGGCA
TTGGCCAAACCGTGTCC**GCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT**

PA2073

ATGAATTCGCAGATGAGCAATAACAACGAGGCGTCACGCCCGAAGATCGAAGTGCGTTTCGATCGATTACGTTCCCCGC
CACGAGCGTCACGGCAAGGTCTGGCACAGGCGCCGTTCTGGTTACCGGCAATTTCTGACTCACCCACATGGTCACC
GGCTTACCGGGCCGGCCCTCGGCTGGGGGCGTGTACTCGATCCTGGCGATCGTCTCGGCGTCTGCTTCGGCACCC
TTCTTCATGGCCTTCCATGCCAACAGGGTCCGCGCATGGGCTTCCCGCAGATGATCCAGTCGCGTGGCCAGTTCCGGC
CTGCGCGGGGCCATCGTGCCGTTTCGCGCGGTGGTCTTCGCTATATCGGCTTCAACGTGTTCAACGTGATCTCGCC
ACCGACGCGATCAACACCGTGTGTCGGGACACCGCGCCCTGGTACGTCCTGCTGATCGGCGTGGCGGTGGTGGT
GCCATCGTCCGGCACGACCTGCTGCATACCGTGCAGCGTGGCTGACCTACGTGATGATCGCGGTGTTTCGGAGTCCCTC
ACCGTCAGCGCGTGCAGCCTCGAAGCGGACCGCGGCTGGGCGGTACGCCGGCGTTCCTCGGAGCGGTTCCCTG
GTCCAACCTGTGCGCGCCGCGCGGTACCAGATCAGCTATGCGGTCTACGTCTCCGACTACTCCCGTACCTGCCGCAC
GAAACGCCGACCCGCAAGGTGATCTTCTGGACTACCTGGGCGCCGCCGTTCCGGCGCTCTGGCTGATGTGCTGGGC
GCATTCCTGGCCTCGGCATTGCCGGCGCGGACGCCATCGGCAGCGTGCGCGAGGTGGGCAACCCGGCTGTATCCCGGC
TTCGGCACCTTACCGTGTGATCGCCGTCGCCGCTGGTGGGATCATGGCGGTCAACTGCTATGGCGCCATGCTC
ACCGCTCAGCGCCATCGATGCCCTTCCGCAAGGTCACTCCGACCCCTCAACCTGCGGTCACCGGGATCGGAGTGCATC
GCCCTGGTGGTATTCGGCGTGGCCCTGTCAATCCCGGAAAGCTACCTGGGCGACTTCAACACCTTCGTCCTGCTGATG
CTGACTTCTGGTGCCCTGGACCGGTCACCTGGTGGACTTCTACTTCTGCGCAACGGTCACTACGCGATCAGC
GAAATCTTCAATCCCGAAGGCATCTACGGTCACTGGGCGGACCGGGTTGAGCGCTACCTGGCCGGGTTGCTGGCG
ATGGTCCCGTTTATGTCCCTGAGCTTCTACAGCGGTCCTTTCGCCGTGGCGCTGGGCGGCGCCGACGTCGCTTCGTC
GTCGGCCTGCTGGTGGCCGGCGCGGTCTATGCCCTCATGTGCCGGGCCCTGGACCTGGAGGCGGAACGACGGCTCGCC
ACCTCGGCGAGCGCTGCTGGAAGAGGGGGCGCC**GCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT**

PPA0619

ATGAATTCGCATATGACGCAGACCGCCACCAAAACGAATGCGACGACGCACAAGGCATCGAGACAACCGGAATCGAG
ATCGTCAAAGAATCACAGCGCACCGCACAGCCCCAGGACCTTTTCTTCCCTGGTTCGCCTCAAATGTCTCGGCTTTT
GGCATGAGCTATGGAGCCTTATGTTGGGATTCGGGGTGTCTTCTGGCAGGCGATAGCCGCCACCCCTCGTCGGCGTC
ATCGTCTCTTTCCGGTCTGCGGCATTATCGCGATCGCCGTAACCGGGCTCCGCACCAACGATGGTGTGTCTCGG
GCAGCCTTTGGTACCCAGGGCAACAAGATCCCCGGGTCACTCTCTGGATGACCTCGATTGGTTGGGAAACATCTCTC
GCTATTACAGCCGCTTAGCGACGACGACGATCTTCCGCCGACTGGGCTGGAGCTCCGGGAACCTCGGTCAAGATCTGC
GCCACTATCATCGTGGCCTTTCTCATTGTGCGCGGTGCTGTGGCTGGTTATCACATCATCATGAAGTTGCAGGCCGTT
TTGACCTGGATTACCGGTATCCTTACCGTGATTTATCTAGTGATGGCGGTGAGCCATATCGACTGGCACGCCGCAACA
TCTTTGCTGCGGCATCCTTCCAGCCTTTCGTCGGCACGCTCACCTCATCATGACTGGCACCGGGCTGGGCTGGACG
AATATCGCTGCCGACTGGTCACGTTACCAGTCGCGCACATCTCCCGGGTTCGCCATCGCCTTCTGGAATGTCTTTGGG
GCCTCTTTGCCGCTCGTCATCTCATTACCGGTGGATTGCTCTTAGCTGCATCCTCAAAAATCTATCCGACGCCATT
GGGCGCATCCCATTTGGCGCGCTAGCAACCATCTTCCGACGTTGCTTCTCATCCCGTTCTTGTGGCTGCGATCCTC
TCCCTGTTGGCCGGTGCATCAACGGCATCTATTCGTCGACTCACCTTGTCCACCTGGGAATCCGAGTCCCACGC
CCAGCAGCTTCTCTCATTGACGAGACGATTCTCACCATCGGGACCTGTACGTCGCTTTCGTTGGTCCGAACTTTATC
TCCCCCTTCCAGTCTTCTGGTGACGCTGGGTGTGCCGCTCTCAGGATGGACCGGCATCATGATGGCCGATATCACC
CTGCGCCGCGCTCCCTACGACGAGCCGATCTCTCAATGGCTCTGGGCGCTACGGTCATTTTCGACCCCATCTCAATC
ATCACCTTCGTCATCGTGACCGTCACTGGGTGGGATGGTTCGTCACACCTACGAGGGCGTCAATTTGGAACGACTGG
CAGGGATTTCTGCTTGGGCCCATTTGGATTGGGTGGCGTGAAGGTGATTGGGCCACGCAACCTGGAGTATTCGGG
GCCCTCGTCTCGGCTATGTTGTTACCTCATCGCTCGGCGCGGACCGTGCCTGTCAGGAGTACGAG**GCTGCAGGC
GGTCTGGCAGCCACCATCACCATCACCAT**

PPA2211

ATGAATTCGCATATGGTGTCCGCGCCAAATAGTGCTCAGGATGAAGTTCGTGCATGAAAATACGGCTCCAACCGCTTCT
GGGGCTGGTTCGCGGTGTTGAGAAGCGAACGATTGATATGGTTCCCTGACGGCGAAAGGCACGGCACTCCGAAGAGTCAG
TTCACACTGTGGTTCGGAGTGAATATGCAGATTACGGCTGTCTGTAATGGTGCCTTTGGCGATTGTTTTCGGGCCGAT
GTTGTATGGGCGATAATTGGCCTTTTTATTGGGCAAGTCATCGGCGGCGCTGTTATGGCGCTCCATTCGCTCAGGGG
CCCGCGCTGGGGCTCCCACAAATGATCAGTTCGCGGGCACAGTTTGGCGTTTATGGTGTCTGCGCTCCGCTGTTTTCTC
GTACTGCTTCTCTACTTTGGCTTCGCTGCTACTGGAAGTGTCTTGGCTGGACAAGCTATCAACAATGCATTTGGTGTCT
CAGTCCAGGTGGATCGGAATACTGGTTTTTGCAGCTCTGACGGCAATTATTGCCATGTTTGGGTATGACCTAATCCAC
AAGCTGGTGAAGTGTATCGACAATTGGCCCTCTGGGGCTGCTTATGTCGCGGTTTCGCTGTGTTCCAAGTGAGTTC
TTGGGCCGACGCTGACCGATGTGCACTTCTCGTGGTGCCTTCTCCTCTGTCGATGCTTTAGCGGCAGGATGGCAG
TTGACTTACGCACCGTATGCCAGTGATTACTCACGCTATCTTCCCCGTGACACGAAGTTTGGGAGTGGTGGTTCGGT
CCCTTCGCGGGATCAGTACTTGGAGTACGATTTTCGATGACGCTGGGTGTGTTTTCATGCTGCTGCTGGTGGTCAAGCT
TTCATTGGCGATCAGGTTGGGTTTATGGGGAACCTTGCCAGAACTCGTGTGTTTGGCGATCTGGTTTTTCATCTCCATT
GTGATTGGCAAGCTGACCATCAACACCTTGAATGCGTATGGCGGAGTCATGGCGTTGTCGACAGCTCTTTTCGGGTTTC
ACCGGTAAGAGTAGTGTCTCTCCGAAGACTCGTTTCGGTTTCGTGATTGGTTTTCAATGTTGTGGTTGCTCTGACGGCG
CTCGCGGCGTCTCGGATTTTCTGCAAGTTTTCAAGTCTTTCATCTTGACTCTGCTGATGTTTTTCTGCCATGGTTCG
GCGATCAACTTGTGAACTATTACCTTATATGTCGTCGCTGTATTGACATTCTGCGCTGTACTCCGAGGGGGCCGT
TATGGCAAGCTCCACTGGCCGACCATTGTTGTTTATTTGCTAGGCGTGGTGTGTCAGATTCCGTTTCATCTCTCAGGCC
TTTTATGTGGGGCCGGTCCGCCGACGCGATGGGTGGAGCGGACATCTCGTGGCTCGTGGCCCTTGTGTTACAGCGTTG
GTATATTTTCCGCTTGGCAGACGAGCGTTTGATTACCCTTGGAGATAGTCTATCCGAAGGATGTGGAGCTGGATCCA
AGCATTGTTTTCTCGGG**GCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT**

SCO5579

ATGAATTCGCATATGAGCAAGACCGCCGAGAACGAAGGCGCCCTGGAAAACCGCGGCATCGAGCAGGTGCCCGACCAC
GAGCGCACCGCCCGACCCGCGAGCTGTTCCCGACCTGGGTTCGGCGCAACATCAGCGTGTGTTGCTCACGATGGGC
GGAGCCTCGTCTGGCGTACCGGCTCAACTTCTGGCAGGCCCTGGTTCGTCGCGGGGGCCGCGCCGGTTCGTGTCGTAC
GGCTGGTTCGGGTCGATCGGCATCGCCGGCAAGCGGGGCGGGGCTCCCGGGATGGCGCTGTCCCGTGGCGTGTTCGGG
CAGCGGGGCAATCTGCTGCCCGGCTCGCTGATCTGGGTTCGCCCGCTGGGGCTGGGAGACGATCAACCGGTTGACCGGT
CCATATGCGCTGCTCACCGTGTGGACATCGTCTTCGGCATAACCGCCCAACAGCGTGTGGACATGGTGACGCTGCTC
GCCCTCGTGGTGGCGACGTTTCGGGATCTCGGGGCTGGGCATCAACGCCGTGCAGAAGTGCAACAAGTACGCGGCGTAC
CTCTTCGGTGCCTTCTCGGTGCTGGTGTCTCGGCTATCTGGTTCGTCGACACCGACTGGTCCGCGGTGTTTCGACCGGTCC
GCCGTTTCGGTGGCCGCGATGATCACGGGTGTGGTCTGATCGCGCGGGCGGCGTCACTGGATCCCGTCCGCGCCC
GACTTCACCCGTTACCTGCCGCGCACGGCTCCTCGAAGGGGATCGTCCGGGTCACGGTTCGGCGGTGCGGGCATCGTC
GTACTGCCGATGGTGTGATGGGCGCGGTTCATGGCGGTCTCCACGCCGGATCTGGCTCCGCGCCGACCCGGTCTCC
TTCCTGGCGAGATCTCGCCGACCTGGATCGCGGTGCCGTACTGCTGATCGCGCTGATCGGCATGCTGCTGATCAAC
TCGATGTCGATGCTACTCGCCGGGTTACC CGCGCAGACCTGGGCTTCAAGGTCCCGCGCACCTGGCGGCTCGGTC
AACGCCGTGATCTCGCTGGTCTTCGGCGGCGTGTGATGCTGGTGGCGACCAGCTTCATGGGGTCTTTCATCGCGTTC
CTGTGCTGCTCGCGGTTCGCTTCTCCGCTGGGTTCGGGCTCTTCGGCGCGGACATGCTGCGCGGCGGGAGTACGAC
GGCGAGGCCCTGGCCGACACCGGCGCACAGCGCTACTGGTACC CGCGCGGCTTCTCCCGCGCGGCTCGTCCG
TGGGCGGTGGGCTGGCCGCGGCTGATGTTTACCACGTCGACTGGTTACC CGGTCGCGTCCGACGGACAACCTC
GTCGGCGAGTACGGACTGGGCTGGGTGGCCACGGTCTGATCTCCGCCCTGCTGTACCGGTCGTCGCGAAGCCGGCG
CTGATGTCGCGGATGCGCGCGGAGCGGGCCGAGCAGCCGAGCGGACCGTGACCGT**GCTGCAGGCGGTCTGT**
GGCAGCCACCATCACCATCACCAT

Appendix 4 Amino acid sequences of the recombinant NCS1 clones. The N-terminal MNSH (blue) and C-terminal His₆-tag (red). Data for EF3000, Pden1111 and Pden4351 are not shown.

b0336 (Note: silent mutation in protein, see DNA sequence)

MNSHMSQDNNFSQGFVPSARKGVLAALTFVMLGLTFFSASMWTGGTLGTGLSYHDFFLAVLIGNLLGIYTSFLGYIG
AKTGLTTHLLARFSFGVKGSWLP SLLGGTQVWGFVGVAMFAIPVKATGLDINLLIAVSGLLMTVTVFFGISALTV
LSVIAVPAIACLGYSVWLVAVNGMGLDALKAVVPAQPLDFNVALALVVGSFISAGTLTADFVRFGRNAKLAVLVAMV
AFFLLGNSLMFIFGAAGAAALGMADISDVMI AQGLLLPAIVVLGLNIWTTNDNALYASGLGFANITGMSKTL SVINGI
IGTVCALWLYNNFVWLTFLSAAIPVGGV I IADYLMNRRRYEHFATTRMMSVNWVA I LAVALGIAAGHWLPGIVPVN
AVLGGALS YLILNPI LNKRKTTAAMTHVEANSVE**HHHHHH**

b0511

MNSHMEHQKRLFQQRGYSEDLLPKTQSQRWTKTFNYFTLWMGSVHNVPNYVMVGGFFILGLSTFSIMLAILLS
AFFIAAVMVLNGAAGSKYGVPFAMILRASVGRGALFPGLLRGGIAAIMWFGLQCYAGSLACLILIGKIWPGFLLTGG
DFTLLGLSLPGLITFLIFWLVNVI GFGGGKVLNKFTAILNPCIYIVFGGMAIWAISLVGIGPIFDYIPSGIQKAENG
GFLFLVVINAVAVWAAPAVSASDFTQNAHSFREQALGQTLGLVVAYILFAVAVGVC I IAGASIHYGADTWNVLDIVQR
WDSLFAFFAVLVILMTTISTNATGNI I PAGYQIAAIAPTCLTYKNGVLIAS I ISLLICPWKLMENQDSIYFLFDIIG
GMLGPVIGVMMAHYFVVMRQINLDELYTAPGDYKYDNGFNLTAF SVTLVAVILSLGKFIHFMEPLSRVSWFVGV I
VAFAYALLKRRKTTAEKTGEQKTIG**AAGGRGSHHHHHH**

Bsu36470

MNSHMKLKESQQQSNRLSNEDLVPLGQEKRTWKAMNFASIWGCIHNIPTYATVGGGLIAGLSPWQVLAIITASLIL
FGALALNGHAGTKYGLPPFVIRASYGIYGANIPALLRAFTAIMWLGIQTFAGSTALNILLNMWPGWGEIGGEWNIL
GIHLSGLLSFVFFWAIHLLVLHHGMESIKRFEVWAGPLVYLVFGGMVWVAVDIAGGLGPIYSQPGKFHTFSETFWPFA
AGVTGIIGIWATLILNIPDFTRFAETQKEQIKGQFYGLPGTFALFAFASITVTSQSQVAFGEPIWVDVILARFDNPY
VIVLSVITLCIATISVNVAANIVSPAYDIANALPKYINFKRGSFITALLALFTVPWKLMEASATSVYAFGLGIGMGLP
VAGVMMADYFIIRKRELSVDDLYSETGRYVYWKGYNYRAFAATMLGALISLIGMYVPLKSLYDISWVFGVLSIFLFY
IVLMRVHPPASLAJETVEHAQVRQAE**AAGGRGSHHHHHH**

lp3374 (Note: silent mutation in protein, see DNA sequence)

MNSHMQQESKYQIEVVAPKHRHMSNWDMFATWIGANANGTWYVGGVLAACGLLTALKVIVASSTLSYLCLSLVGFMG
YKTGAATMSLIRGSFGVRGSYVPSFVNLQYIGWTAVENTFIAATSVSYLLHDLVWGPVYKPGGAKGLILGIIVMSVL
HLLSVSVGQRSVQMIERLGIILVILFVLWETVVVFQTVSVHDLLTWQAPTHLHMTAGAGMDTLAAFNLSWVTAGADFT
RFRTRKKSAGTGMPLGAFTGLFWFAFIFGLAATISIAITSGTYDPNNSDPSTIASRLGLGLVALLVIVLTSMTANAVNL
LAAGSALSNIFFRIRLRPALWAVTIATLVTLIPLIMGSFLAAFTAFLDYIGMVLGPMISVMLVDYWRHRQHYDINE
LASSKGQYWHHGWNWIALLCWVLGVAIFLLLKHVWVIATVGTGATFIDMALIGVIYVALMRIFYPLPAKRVA**AAGGRG**
HHHHHH

PA0438

MNSHMSADSNYSQQFPVLAARKSALALSVMMLGLTFFSASMWTGGTLGTGLSYDDFFLAVLLGNMLGLYTAIILGFIG
ARTGLSTHLLAHYSFGSKGSWLPALLGGTQVWGFVGVAMFAIPVAKATGLDVLNLLILVSGALMTLVVFFGIAGLTL
LSAIVPAIVVLGSYSVWLAVRDAGGLAALQQVTPSAPLSLSTAIALVVGVSFVSAGTLTADFVFRGRSARTAVVCLV
AFFLGNLSMFIIFGAAGAAAVGKSDISDVMIAQQGLLIPAILVLGLNIWTTNDNALYASGLGFANITGLSSRFLAMLNGA
LGTALALWLYNHVFWLTLFLSAAIPPIGGVIAIDYLSRRRRYRQHPCADIQAINWAAI LAVACGTAAGHWLPGIVPN
AVLGAAIAYLLNPLLSRSPNKPEEVLHAD**HHHHHH**

PA0443

MNSHMQQSRSEVTERDGLVELSAGSDVLDSPRYNNDIAPTRVEQRTWNRWHITALWIGMSICVPTYTLGGVLTAYFGL
SVGEALLAILLANLVLIPLTLNFAFAGTRYGIPFPVLLRASFGIIGSNVPCILIRAVVACGWFGIQTFLGGLAIHFLG
SIFPAWKALGGTGEVIGFLLFWGLNLWVVLRGAESIKWLETLSAPLLVLVGLGLLVWALPHISVSELLAQPPKRPEGA
SVTGYFMAGLTAMVGFWATLSLNI PDFSRYARSQKQDQILGQIFGLPLTMFLFAALGVVMTAASASLVGETVSDPVS
LI GHI P SPGWAVAMALIIVATLSTNTAANIVSPTNDFQNIAPKLIGRTRAVWLTGFIGLALMAHELKLLGWIVSDLSL
ESVYSNWLLGYSLLGPIAGIMVVDYFLVRRQRDLAAGLYRDDVYPAWNWPGFAAFVAVPVALTVMALGNRQFHWFYDY
GWFTGSLGGALYYGLCRRRARS PAALAKPLP**AAGGRGSHHHHHH**

PA2073 (Note: Cloned protein have a N-terminal MNSQ not MNSH)

MNSQMSNNNEASRPKIEVRSIDYVPRHERHGKVVHQAPFWFTGNFVLTMTVGTGTPALGLGALYSILAIIVLGVCFGT
FFMAFHANQGPRMGLPQMIQSRAQFGLRGAI VPF AAVV FVYIGFNVFNVI LATDAINTVLPGRHAPWYVLLIGVAVLL
AIVGHDLHTVQRWLTYVMIAVFVLTVSALRSLEADAALGGT PAFSWSAFLVQLSAAAGYQISYAVVYSDYSRYLPH
ETPTRKVI FWTYLGAAAGSALWMLSLGAFLASALPAPDAIGSVREVGNRLYPGFGFTFVLI AVPALVGMVAVNCY GAML
TSLSAIDAFRKTPTLNLRVGTGIGVIALVVFVVALSIPESYLGSFNTFVLLMLYFLVPWTAVNLVDFYFVRNGHYAIS
EIFNPEGIYGHWRPGLSAYLAGLLAMVPFMSLSFYS GPFAVALGGADVA FVVGLLVAGAVYAFMCRALDLEAERRLA
TLGERLLEEGGA**AAGGRGSHHHHHH**

PPA0619

MNSHMTQTATKTNATTHKGIETTIEIVKESQRQAQPQDLFLPWFASNVSVFGMSYGAFMLGFGVSFWQAI AATLVGV
IVSFGFCGIIAIAIGKRSAPT MVSRAAFGTQGNKIPGVISWMTSIGWETSLAITAVLATTTIFRRLGWSSGNSVKIC
ATIIVAVLIVGGAVAGYHIIMKLQAVLTWITGILTIVILVMVAVSHIDWHAATSLPAASFPAFVGTTLTLMTGTGLGWT
NIAADWSRYQSRTPGFAIAFWNVFGASLPLVLLITGGLLLAASSKNLSDAIGADPIGALATILPTWFLIPFLAAIL
SLLAGAINGIYSSGLTLLTLGIRVPRPAASLIDETILTIGTLYVVFVAPNFISPFQSFVLTLVGVPVLSGWTGIMMADIT
LRRRPYDEADLNFNGSGRYGHFDPISITFVIVTVIGWGLVVNTYEGVNWNDWQGFLLGPIGLGGREGDWAHANLVGF
ALVVLGYVVTLIARRGTVRRQESR**AAGGRGSHHHHHH**

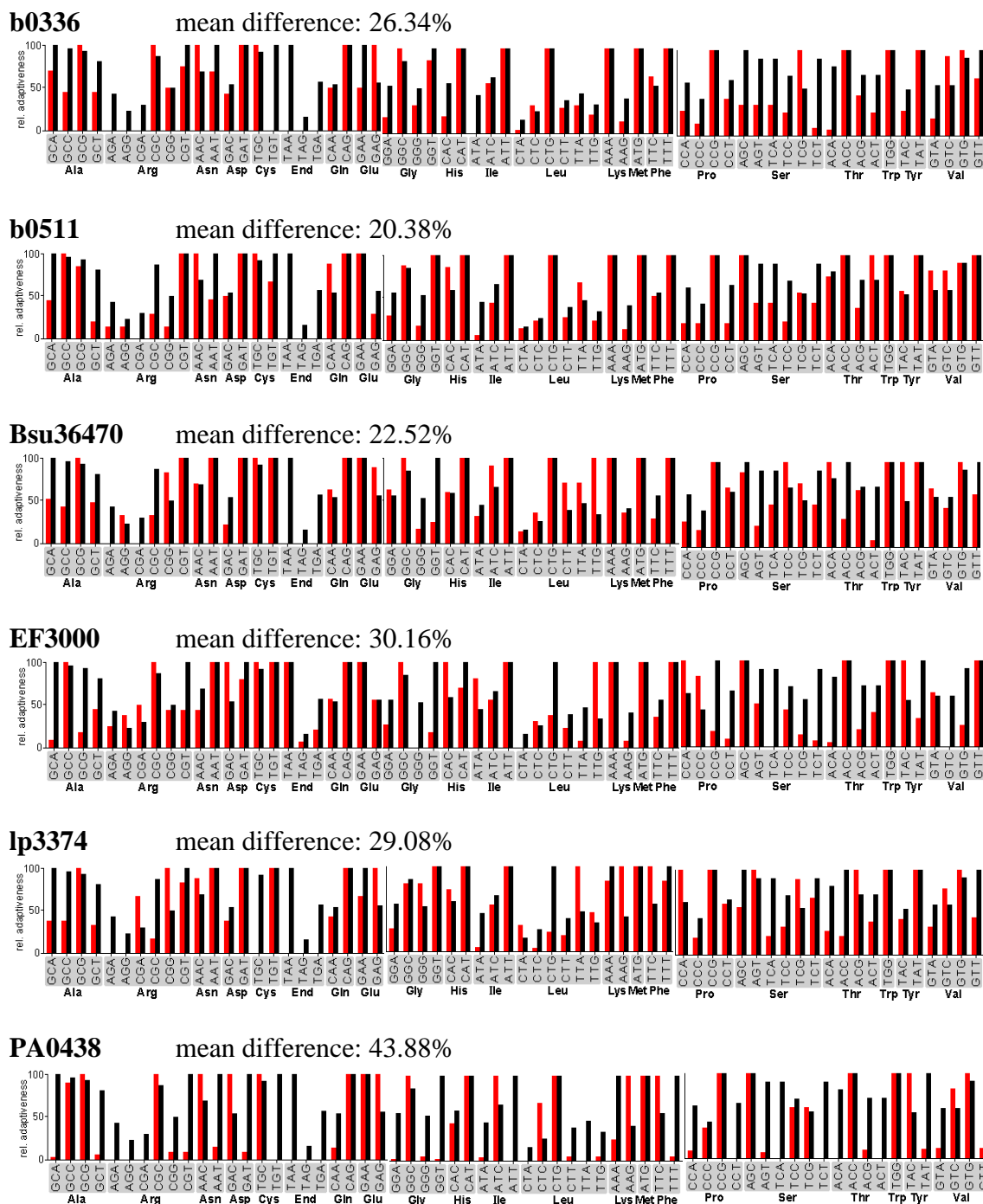
PPA2211

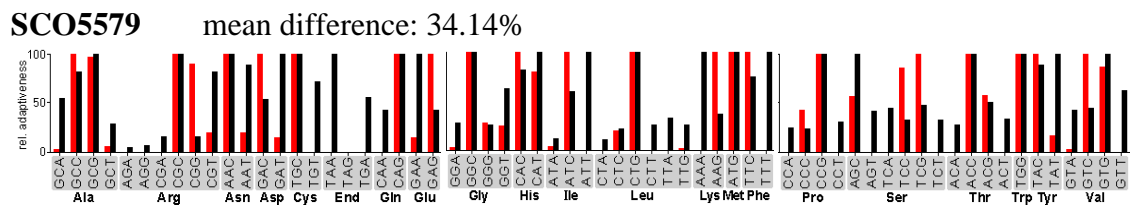
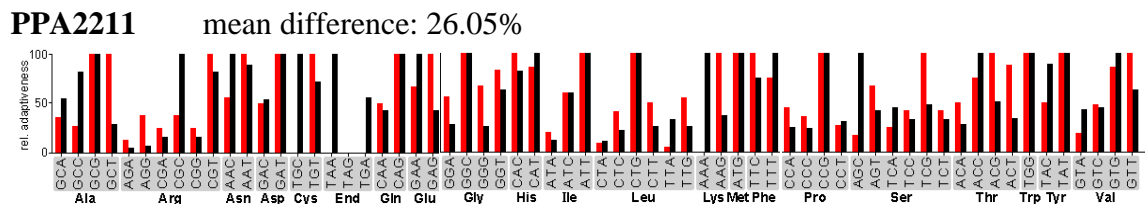
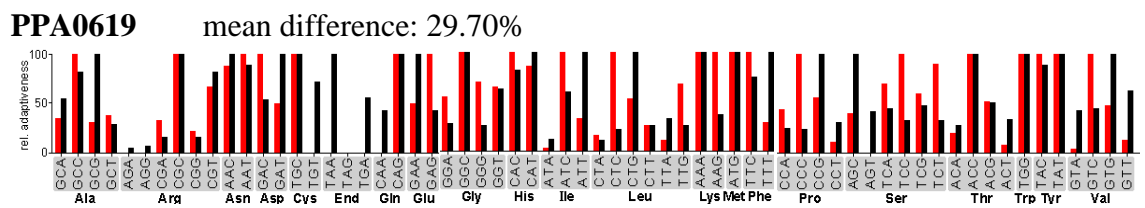
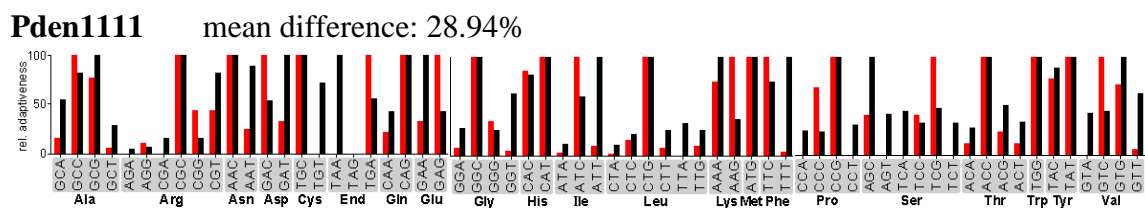
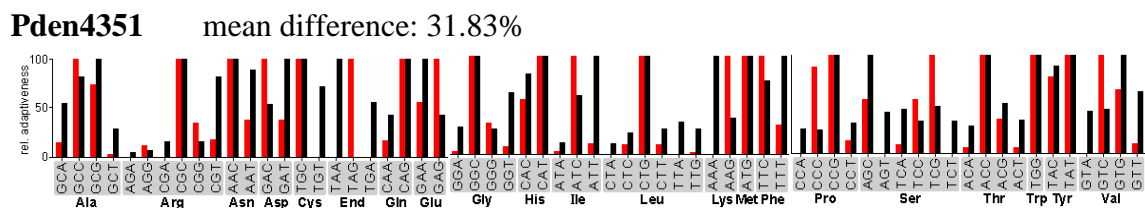
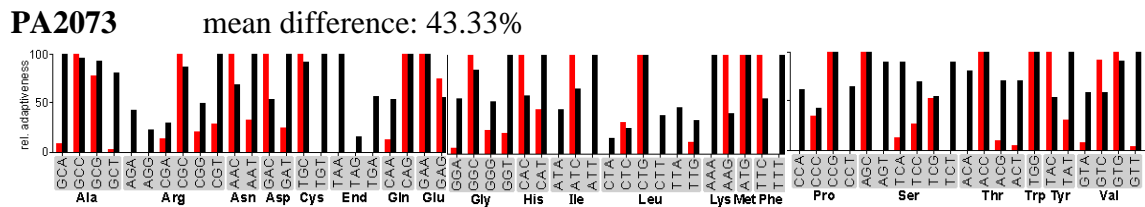
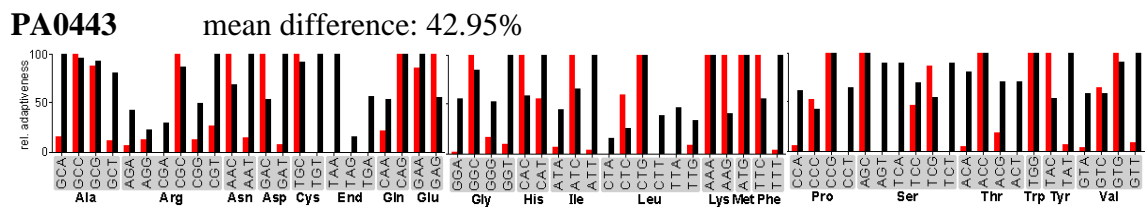
MNSHMVSAPNSAQDEVHRGNTAPTASGAGRVEKRTIDMVPDGERHGTPKSFQTLWFGVNMQITAVVNGALAIIVFGAD
VVWAIIGLFIGQVIGGAVMALHSAQGPALGLPQMISSRAQFGVYGAALPLFLVLLLYFGFAATGTVLAGQAINNAFGV
QSRWIGILVFAALTAIAMFGYDLIHKLVKVSSTIGLLGLLYAVRLFQVSSFGPQLTDVHFSVVPFLLSIALAAGWQ
LTYAPYASDYSRYLPRDTKFGSAWFGPFAGSVLGATISMTLGVFIAAAGGQAFIGDQVGFMGNLQNRVLAIVL FISI
VIGKLTINTLNAYGGVMALSTALSGFTGKSSVSPKTRFGFVIGFNVVVVLTALAASSDFLQVFKSILTLMLFFLPWS
AINLLNYLICRRCIDIPALYTPRGRYGLKHWPTIVVYLLGVVVQIPFISQAFYVGPVAAAMGGADISWLVLVVTAL
VYFPLGRRAFVDYPLEIVYPKDVELDPSIVSR**AAGGRGSHHHHHH**

SCO5579

MNSHMSKTAENEGALETRGIEQVPDHERTARTRELFPWVGANISVLLLTMGASLVVAYRLNFWQALVVAGAAPVVS
YGLVGLIGIAGKRGAPGMALSRAVFGQRGNLLPGSLIWARWGWETINAVTGAYALLTVLDIVFIRANSVLDVMTLL
AFVAVTFAISGLGINAVQKCNKYAAYLFGAFSVLVGLYLVVDTWSAVFDRAAGSVAAMITGVGLIAAGGVSWIPSA
P DFTRYLPR TASSKGI VGVTVGGAGIVVLPMLMGAVMAVSTPDLASAADPVSFLGEILPTWIAVPYLLIALIGMLLIN
SMSMYSAGFTAQTLGFKVPRHWAVSVNAVILVFGGVLMLVATSFMGSFIAFLSLLAVAFSAWVGVFGADMLRGREYD
GEALADTGRTSAYWYRGGFS PAAVVAWAVGLAAGLMTTSDWFTGPLATDNFVGEYGLGWVATVVISALLYAVLPKPA
LIVPDARAERAEQPEAETVTV**AAGGRGSHHHHHH**

Appendix 5 Comparison of the codon usage of the NCS1 family of transporters and the *E. coli* expression host. The fraction of usage of each codon in the protein was computed and plotted against the fraction of usage of the codon in the *E. coli* expression host (relative adaptiveness), and a mean difference value (mean relative adaptiveness) of codon usage calculated. The analysis were computed and plotted by ‘Graphical codon usage analyser’ (Fuhrmann *et al.*, 2004). Black bars represent the *E. coli* codon usage, and red bars represent the codon usage of the selected NCS1 protein.





Appendix 6 GC contents and rare codons in the NCS1 family of transporters. The number of rare codons in each NCS1 protein was calculated using 'Rare codon caltor' (Ng, E., 2001).

	Rare codon											GC content (%)	
	Arginine				Glycine		Isoleucine	Leucine	Proline	Threonine	No. of rare codon (No. of consecutive rare codon)		Consecutive rare codons (No. of times it appears in the protein sequence)
	CGA	CGG	AGG	AGA	GGA	GGG	AUA	CUA	CCC	ACG			
b0336	0	2	0	0	4	7	0	1	1	7	22 (0)	None	53
b0511	0	1	1	1	7	4	2	3	2	3	24 (1)	CCC,GGA (x1)	47
Bsu36470	0	5	2	0	15	4	7	2	2	8	45 (7)	CGG,AGG (x1) GGA,GGA (x2) GGG,AUA (x1) CUA,AGG (x1) GGA,AGG (x1) GGA,CGG (x1)	49
EF3000	8	7	6	4	3	0	13	0	9	4	54 (4)	AUA,AUA (x1) CCC,AUA (x1) AUA,GGA (x1) CGG,CGA(x1)	42
Ip3374	4	6	0	0	4	12	1	8	1	18	54 (4)	ACG,CUA (x2) ACG,CGG (x1) CGA,CUA (x1)	49
PA0438	0	1	0	0	1	2	1	0	4	2	11 (1)	GGG,AUA (x1)	68
PA0443	0	2	2	1	1	7	2	0	9	4	28 (0)	None	67
PA2073	2	3	0	0	2	8	0	0	6	2	22 (1)	CGA,CGG (x1)	66
Pden4351	0	6	2	0	1	12	1	0	7	6	35 (3)	GGG,ACG (x2) GGA,CGG (x1)	65
Pden1111	0	4	1	0	3	13	1	1	11	4	38 (1)	GGA,ACG (x1)	65
PPA0619	3	2	0	0	11	14	1	3	9	13	56 (5)	ACG,ACG (x1) CCC,GGG (x2) ACG,ACG,ACG (x1) GGG,CCC (x1)	58
PPA2211	2	2	3	1	10	12	4	2	4	8	48 (4)	CGA,ACG (x1) GGG,CCC (x1) GGA,AUA (x1) AGA,CGA (x1)	53
SCO5579	0	9	0	0	1	10	1	0	6	11	38 (1)	CCC,GGG (x1)	70

Appendix 7 DNA sequences of the recombinant histidine kinases clones. The N-terminal MNSH (blue) and C-terminal His₆-tag (red). Only the clones that are made by myself are shown.

EF1051

ATGAATTTCGCATATGAAAAGAACGATTAAAAAGAAGCTGGAAGGCCCATCTTTAACTATAAAAGTGGGCTTTCGCAAGT
TCTTTCTTTATATTTGTAGTTTTTACCATTTTTGCGGTGATTACCTATAAGTCTTCTGTGAGTCTTATTTGTTGCCAAA
GAAAAGGAAAATGTCGAAGCAACGATTGCAGAAGTAACAAATCGATTAGCTAATGCCAATGAAAATTTAACGGTAACA
GACGTTTTTGGACTATTTAAAAACGCCGAGTGAAAAGAGATGAAAACACTATAAATAAGCATAACAGCAGTAGAAGGTTCCG
TTTATGAAAATGGACAGTTTCATATCTGAACTAGGACAACCAGAAGCTTTATTTATCTGTGTATGACACAAAACAAAAG
CTAGTTTTTAAAAACAAAATGAATATGATAAGTTATTACAATTTGGATCGACAATTACCTGTTGTTGCGAACCGTTTTT
GATAAAACCGGTTTTTATTCCGTGGAACCTATTTTTTCAAAGAACACAGAGAAAAGATTGGCTATATTCAAGCGTTT
TATGAACTTTCTCTTTTTATGAAATTCGTAATCATTTATTATTAAACACTTGTAGTTTTAGAAGTGATTTCTCTGATT
GTAAGTAGTGTCTTAGGCTTTATCCTCTCTCTTATTTCTTAAACCATTTAAAGTATTGAGAGACACGATGGATACA
ATCCGCAAAGACCCGCAATCAGATGTCCACATGCCGAGATTAAATACGAGAGATGAGTTAGCAGATATCTCGGAAATC
TTTAATGAAATGTTAGATCGTATGAGACGTATATCGAACAACAAGAGCAGTTTTGTTGAAGATGTTTCCCATGAATTA
AGAACGCCCGTTGCGATTATGGAAGGCCATTTAAACCTTTTAAATCGTTGGGGGAAAAGACGATCCTGAAAATTTTAGAT
GAATCATTAAAGGCCAGTTTACAAGAAATTAGTCGTATGAAGAGTTTGGTCCAAGAAATGCTTGACCTTTCACGCGCT
GAAACAGTGGACACCCAATATGCAAATGAACGAACAGATGCTAAACAAGTAGTCTACCAAGTATTTAATAACTTCCAA
TTGGTTTATCCTGAGTTTCATATCACATTAGATGACGATTTACCAACCGAAGTTGAGCTGAAAATCTATCGTAATCAC
TTTGAACAATGCTAATTATCTATAGATAATGCGATTAAATATTTCAACCGATCGAAAAGAAGTGCATATTTCCATT
TCTCGAACGATGAATGAATTTGAAATAGCTGTGCAAGATTTTGGTGAAGGAATCACTGAAGAAGACTTGGAGAAAATA
TTTGATCGATTTCTATCGAGTAGATAAAGCCCGTCAAGAAATAAAGGTGGTAATGGCTTAGGCTATCGATTGCAAAA
CAATTAGTCGAAAACATAAAGGACGTATTGATGCAGAAAGCGTATTCATCAAGGAACGATCTTTAGAATTTTTATT
CCAATCGCTGGAATAAAGAAGAGAGCAACGTACAAAGCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT

EF1820

ATGAATTTCGCATATGATTTTTGTCGTTATTAGCTACTAACGTTTTGCTTGTATCTAGCTTTATCGTTTTTGTCTTTTTA
CGAGTAACATTGATCAAGATTGAATGTAATAACCGTTACTTTTCATTGCTTATTGTCATTAATCTTTGTTGTTTTGCG
GCATTAATGTTGGGCTACTCTTGGTTGATTTATGCGCTGACAGTCGTTATTTTTACAGGATTTTTACTCATTACAAA
AAGAGGTTCTCAATTTTTAAAGCGATATTTTTGCTGTGTTTTTACATTGCTTATGGTTTCGTTTTCAATTACACGGAG
CAAACGATTTTTAGTGTTTTTTTTCAACAGATTTATCAAAAATAAATATTATGGATTGCCTCAAATGTTCTTCTGTTG
CTTATAAATATCTGGATTGCTTTAAAAATTTCCCAATAGTGTTTTTTTAAGATTAATCGTGTGTTAGAAAATAGCCGA
ATTTTTTTTTGGTTGTTACTTTTTATTGTTGATTCTGTTGTTACTTTTTGTGTTTTTGAATTCGCCAGAGATTTACCT
GACTTTTATGCGAGGATTTGTCACGGTAAATAGTCTAAATTTGGAGTTATTAATAAGTGTAGGTTTTTTTTAATTTCTG
ATTGGCTTAGTCATTGAAGCTTATTTGGAAGAACAACGTATCAACACTCAATTATTGAATAATTTAACGATTTATACT
GAAAAAATAGAATCCATTAACGAAGAGCTAGCGATGTTTCGTATGATTATAAAAAATTTATTGTATAGTTTACAAAT
GCTATTTTCATACGAAGATATTTCTGAAATCAAAAAGATTTATGAAGAAACGATTGCACCAACCAAAAAAATTTATTGAT
AATGAAGAATTTGAACCTATGAAGTTAAATCGTTTGAAAAATATGGAACCTAAGGCACTCATAAGTATGAAAATTAAT
ACCGCAAAGCAAGCAAAACTAAAAGTGATTGTTGATGTGCCAGAAGTATTTATTTGATACATCAATCGACTTAGTG
GTTGTGATTCGTTTGTGGCTATTCTATTAGATAACGCGATTGAAAACAGTGCAAAATCAGAATGAAAATGTTTCGCT
ATCTCAATTTTCAATAAAAAATGAAACACAAGATTTGTGATAACAATAAGTGTCCAAAGCCGAATTTGATTTTAAAGTT
ATGAAGAAAACCAAAATTTAGTTCTAAAAGTAATCCAGAAGAGCACGGTTGGGGATTGTTATATGTAAGAAAATTTGTT
GATTTTTTCAGATCAATTTGATTTGCAAACGTCCTTCAATGAAGGAGCGGTCACTCAACATTTAATTATGAAAAAAT
CATAACAGTAAAAAGTTGTTAACGAAAGCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT

EF2912

ATGAATTTCGCATATGACCGATCGGATTTCAAGACGCATGATTTTCATATATGCGTCCCTTAGCACCTTTATTGTTATC
TTAATTACATTTGTTTCATATTTTCATTTCGATTAACAAAACCGGTGGTTATTAGAGCTTCTTCAGAGAAAAGTCTTT
TATTTACCCTAATTTGTGCACATTTGTTCTCATATCCTTACTAATAGGCTTATTGACCTTTTTACTGATTTTCATTGGTT
CAAAAAGGGCAATATGGACGGATTGAAGAAAACCTTCGGTTATTGGCCAACGGTAATTTAGAAAGTCCAGTCTTAAAC
AAACCAACGACAGTGAATAATCAAGACCATTATCTAACGAAGTCGAACAAGATATTTGGTCGATTAATAAATAAATTA
TTAGAGATGTCTAAAAGAAATGCAATTTAAACAGTCGACCGCAATTAATGGATGGGCAAAACAAAAGAAGAAATTTTA
GAGAACGAGCGGCATCGTTTGGCGCGGAGTTGCATGATTCAGTCAGTCAACAACTTTTTGCAGCCATGATGATGTTG
TCTGCATTAATGAACAAGCACAAACGAACAAGAAACCCCGGAACCATATCGTAAACAACACTAGCCATGGTGGCAGAAATC
ATTAATGCCTCCCAATCGGAAATGCGCGCGCTACTATTGCACCTGCGTCTATCAGTCTAGAAGGAAAAAGTTTGGCGT
AAAGGTATTGAACAATTAAGAAACTACAAACAAAATTTAAATTTGAATTTGATTTGGGATGTTGAAGATGTTTCAT
TTAAATAGCAGCATTGAGGATCATTTTTCCGAATTTGTGAGGAGTTACTTTCAAATACCTTAAGACATGCCAAAGCA
AAGGAATTAGAGGTATACCTTACACCAAGTCGATAAAAACGTTGTTATTGCGTATTGTTGATGATGGTGTGCGCTTTGAT
ATGAAGGAACAAGTAATAAAGCCGGTAGTTATGGCTTAAATAATATTCGAGAACGTTGTTGCGGCATGGGCGGTACA
GTTAAAATTTATTGTTTTAAAGGCGAGGAACCGGTTGAAATTTAAAGTTCCGTGTCATAAAGGAGGAAACTGCAAGT
GATCAAAGTAATGTTAGTGGAGCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT

EF3197

ATGAATTCGCATATGGTTCGAATTATTTATTTTAAATGATGGAACGGGTGGGCTTAATTATTTTACTGGCCTTTCTTTTATGTAATGTGCCTTATTTTAAAGCGTGTATTTAAGTCGTGAAAAATGTCTTCAAAGTTTCAGTTGATTTTGATTTTGGTTTATTTGCTATTATTTCTAATTTTACAGGTATCGAAATCGTAAAGAAATCAGATTGTTCCCAATAATTTACTAACC TATTTAAGCAGCAATGCTTCCATTGCGAATACGCGTACGTTAGTAATTGGCGTTTCTGGATTAGTGGGTGGTCCGATTGTCGGCAGTGCAGTTGGTTAATTGCTGGCTTTCATCGGGTTATTCAAGGCGGGGGGCATAGTTTCTTTTATGTGCCT GCTTCTTTAATTGTTGGCTTAATTGCCGGTTTTTATAGGGAGCCGCATGGCAAAGCAAAGTGTCTTCCGTCTGCCGGC TTTTCTGCCATTGTCGGAGCGTGCATGGAATGATTCAGATTGATTTTATTTTCTTTTTCAGCGGTGATCTATCGGAT GGGGCCACGTTGGTTCGTTTTATTGCGCTGCCGATGATTTTGTTAAATAGCGTAGGAACGTTTCATTTTATGTGCGATT CTAACGACGACATTTGAAACAAGAAAGAGCAGGCCAAAGCGGTTTCAGACGCACGATGTTTTGGAATTAGCAGCGGAAACA CTGCCATATTTTCGAGAAGGCCTCAATAAAAAATTCAGCAAAAAAGTTGCGGAGATTATTAAGCATTATACAAAAGTC AGTGGGATTAGCATGACAAACAGTCATCAAATTTAGCTCATGTTGGCGCAGGAAGCGATCATCATATCCAGAGCTA GAAGTCATTACCGAGCTTTCTCGGGAAGTGTACGGACGGGCCGAATGACCATCGCGCATGCCAAAAGAAAGTTGGT TGTTTCGGACCCAAATGTCGGTTCGAAAGCGGGCATTTGTGATTCCATTATTTTCCCATCAGCAAATTTGTGGCACCTTA AAAATGTACTTTACAGATCCAGCGCAGTTAACGCATGTCGAGGAGCAGTTAGCAGAAGGTCTTGGGACAATCTTTCTC TCACAAATTTGAATTAGGTGAAGCGGAAGTACAATCAAATTTACAAAGAGCTGAAATTAATTCCTTGCAGCGCAA GTGAACCCGCATTTCTTTCTTTAATGCCATCAATACTATTTCCGCATTAATGCGGAAAGACAGTGAAAAAGCGCGGAAA CTATTGCTCCAATTAAGCAAATACTTTTCGAGGCAACTTGCAGGGAGCCGTGCAAACACTACGATCCAGTTTACAAGAA CTAGAACAAGTAAAAGCCTATTTGTCATTAGAACAAGCAGGTTCCGAATCGTTATCAAGTAACATTTGACGTGGAA GAAGCGGTTGCTTCAGAAAAAGTACCGCCGATGCGATTACGGTGTAGTGGAAAAATACCATCAAGCATGCATTTGGC AGCCGGAAGGAGAACAATCAAGTCCGGGTAGTGGTCAAAGCAAAACAACAATTAATCATGTGGATGTGTATGACAAAT GGGCAAGGGATTCTGAAGAACGCCGCTTGTATTAGGAAAGCAGCAGTGACTTCGAAAAAGGCCAGCGCACGGCA CTGGAAAACTTCCCGCAGAATGGCGAACCTTTATGGATCAGAAGGTTGTTTTTAAATTGAAAAATTTAGAGACAGGC GGGAGCCATGTTCAATTTAGAAATCCCAGTGAACAGGAGGAATTAGATGCACGTATTAATCGTCTGTCAGGCGGTCGT GGCAGCCACCATCACCATCACCAT

EF1209 (mutation highlighted in yellow, from C to T, giving rise to amino acid change H to Y)

ATGAATTCGCATATGAAGCGTGGCGGGAAATATGGTGTTCCTTATCAGTGATTTTTATTGCAACATTAATTATTATG ATTTACTTTCTTTTACGGAGTAACGACGGTTCAAACAATTAAGAAGTTCGTAAAAACCAAGAGCAAGCACTTTTAGCC GTCGGGGAACAGTTGGCTATTGAACCGAATGTTATTGAAGCCTTAAAGAATGATCACTATTCAGATGAATTGGAAGCG TATACGGTTCGTTTAGGTGAAATCCATCAGTTGGATTTTATCGTGATTATGAATATGCAGGGGATTCGTTAACACAT CCTGATCGCCAAAAAATAGGTAAGCATTTTGAAGTGGCGATGAAGTGCAGGATTAAGGGGAGGAACATTTGTCT GTTAGTCAAGGGAGTTTAGGAGAATCATTACGAGGTTTGTCCAGTCTATGATCAAGGAAAACAATTTGGGGTTGTG GCAATGGGCATTAATAATGACCTCGCTTTCACAATTTGATTGAACGAACAAAAATGATTACACAGTGAGTGTCTTTTA AGTGTGGTTTTGGCTTTATATTAGCCATCGTGCTTTCTTATTATTTAAAAAACCACTTCACGATTCAGAGCCAAAGA GAAATGCTCGTCTATTAGAAGAGCGAAACGCCATGCTTGGAGAAACGAAAGACGCAATCCTTGTGATTGATACCGAT CAAAATATTTTATTAGCAAATATTGAGGCGACTAAAATGTATCATAATAACAATAGTGAAGAAAAATTTACTAGGA AAGAACTTTTCAGATTGGTTTTATCTCCTGAAAAATTAGTCGTACATTCAAAAACAGAGCAATTCATCGTCAAAAT GGCCAGGATTATTTGCTCTCAATCGCCCCGATTAATGTCCGTAAAAAAAACAATTTGGGCATGTTATCTTTTGAAAAAT GCGACAGAAACGTTTATTGTGCGCAGAACAGCTTGTGACGACAAACAATATGCTTCGGCATTAACAAGTCAATCACAC GAATTTATGAATAAGATGCATGTCATTTATGGCTTAGTCGATTTAGAAGATTATGAGGCGTAAAACATTTATCTAGCC GATTTATTGAAACAGAAAAGGAATTTGCACAGCGTTTAGTATTATTTTAAAGAAACCACTTCACGATTCAGAGCCAAAGA AGTGGGAAAGAAATCAAGTTTGCAGAAATTAACAACAATTAGCTATTGAAATTTATCCAGAAATTCACCAAATAAAA CGAGATGAAGATACGAAAATTTGATTGCTATCTACCGTTATATTATCGTTTCTGATGGAGCAAACGTTGCCAGAA GAGATCATTGAAACCATTGATTATCAGCCAGGAAGTTTGCAGCAGACTTATTCGTTTGTCTATCCAAAAGAAACAAC TGGAACGCTTTGAGCAAGAGTTTTCACTTCTTATTGGCAAGATTATTAGAGAATGCGGAAGCGACATTTGACTTGGGAA AACCAACAACAATTTGGCTTGTCTTCGCATAAATGTTCACTTACGAAGGAGCAGAGGAAAAATGAACCTATTGATTAT GCTGTCAGGCGGTCGTGGCAGCCACCATCACCATCACCAT

EF0373

ATGAATTCGCATATGACCATTAAACGCGCTTTTTTATTTCTTATATTAGTGCCATCATTATTACTTTAGCCTCTGTTT TAGCTGTTT TATCCTTAGCTTCTTACATCACATTAGGAACGTCCCCAGTTTGGCCCAAGCTTACCGAATGATGAATAAGCAACGACCATTGACCGC TAACGAAGAAGAAAGTTATTTGGCCTTAGATCAACTGCTAAAAAATCACCAGAAATTAAGTATACACCTTTGAGTAAAGAAATTAAG GAAACGATTCAAACAATTAAGCAAAAAGGATTAAGTGTGATTATTCGAAAGAAATGCACGGTTTCTTATTTTCGGATAATTTAGTAG AAAAATCTTTGCTGTACATATTCAAAACAGCAAAATGAATAATATTATGCCAACTGGCACACTGGATAATGCAGGGCGCTATATCA TTATGTCAAATCCGATTTTCATTTTATAGTGGTCTAACGGGAGCTTTATTGTCTTGAACCGGAAAGCAGCCTTTTGAATCTTTC ACACGCTGGATTATCTGGATGATTTTGTAAATATCCTTGTAGCAATTTGCGGCAGCTTGGCTCATTAATAAGCGCTTAACCAAACCA CTATCGAGCCATTAGAAGCATTAGAAAAAGGACTAAAACATTTGGGTAAGGATGCCAAAAAGAAAAATCCATTTACGCCAACTAGCCA CCAAACCGTCTCAACAGAAGTCAAACAGCTCCAACCTAGTTTTGAACAAATGTGGCAAGATTAGAAGAAGCAATGCGGAACGCGAA AAATATGAAGCTAATCGGAAAGAAATTAATGCAAAATTTCTCATGACTTAAAAACCAATTACTTCTATTATTGGTTATGTTGAAG GTTTGATGGATTGGCGTTGCCAATACTGAAGAAAAAAGACCGCTATTTAACGGTCATTCATGAAAAAGTTTAGCCCTCAATGATTT AATTGAAGAATTTTTTGTATTCTAAATTAGACCTAGACCGTGCTGTCTTTACAATGGAAAAGACCAATTTCACTCGTTTTATTGCA CATATTTGAAAGAAATACCGTTTGGAAACAAGAACTCGTAATCACTTCTGTACTTCCACCGAAGCTTTATACGTTCAAATGGATCCCA CGCAGATGAATCGCGTGATTACCAATCTAATTCAAAAATAGCATTAATTTGGGGATCCCACTAAAGAACAACCTAGCCTTTACGATTTCT TTTAACGCATAATCAAACAGACTTAGTTTTAAACCATTACAGATAACCGCATTTGGTATCGATAAAAAAGAAATTAACCTTATTATTTCGAA CGCTTCTATCGTGTAGACAAATCTCGAACCCCAACAGTTAAAGGAAGCGGCTTGGGATTAAGTATTGTCAAACAATTTATTGACTATC ATCAGGGAACAATCACTGTCAAGTAAAAAGGAGACGGCACCACCGTATAATTACGCTGCCGTTGTTGGAGGATGAAAAAGCTGC AGGGGTCGTGGCAGCCACCATCACCATCACCAT

Appendix 8 Amino acid sequences of the recombinant histidine kinase clones. The N-terminal MNSH (blue) and C-terminal His₆-tag (red). Only the clones that are made by myself are shown.

EF1051

MNSHMKRTIKKELEGPSLTIKWAFASSFFIFVVFVTFIFAVITYKSSVSLIVAKEKENVEATIAEVTNRLAN
ANENLTVTDVDFDYLKTPSERDENYYNKHTAVEGSGFMEMDSFISELGPPELYLSVYDTNQLVFKTQNEYD
KLLQLDRQLPVVTRVFDKTFGYSVPEPIFSKETREKIGYIQAFYELSSFYEIRNHLLLTLLVLEVISLIVS
SVLGFILSSYFLKPLKVLDRDMDTIRKDPQSDVHMPEINTRDELADISEIFNEMLDRMRRYIEQQEQFVE
DVSHELRTPVAIMEGHLNLLNRWGKDDPEILDES LKASLQEISRMSLVQEMLDLSRAEQVDTQYANERT
DAKQVVYQVFNFLVYPEFHITLDDDLPTVELKIRYRNHFQQLLILLDINAICYSTRKEVHISISRTM
NEFEIADVDFGEGITEEDLEKIFDRFYRVDKARARNKGGNGLGLSIAKQLVENYKGRIDAESVLHQGTIF
RIFIPIAGIKEESNVQAAGGRGSHHHHHH

EF1820

MNSHMILSLLATNVLLVSSFVVFVFLRVTLIKIECKIPLLSLLIVINLCSFAALMLGYSWLIYALTVVVIF
TGFLLIHKRFSIFKAIFLSVFTLLMVSFINYTEQTIILSVFFQYIQNKLLWIASNVLLLLINIWIALKI
PNSVFLRLNRVLENSRIFFGCLLLLLILLLLFVPLISPEISPDFMRGFVTNSSKLELLISVGLFLLIG
LVIEAYLEEQRINTQLLNNLTITYTEKIESINEELAMFRHDYKNLLYSLQIAISYEDILEIKRIYEETIAP
TKKIIDNEEFELMKNLRLKNMELKALISMKINTAKQAKLKVIVDVPEVFILDTSIDLVVVIRLLAILLDN
AIENSAKSELKMFASIFNKNETQEFVITNSVQAEFDFKVMKTKFSSKSNPEEHGWGLLYVKEIVDFSD
QFDLQTSFNEGAVTQHLLIEKNHNSKQVVNEAAGGRGSHHHHHH

EF2192

MNSHMTDRISRRMISLYASLSTFIVILITLFSYFHSIKQNRWLELLQKRVFYLPPIVHIVLISLLIGLL
TFLLISLVQKQYGRIEEKLRLLANGNYESPVLNKPPTSENQDHYLVEVEQDIWSIKNKLEMSKELQLL
NSRPQLMDGQTKEEILENERHRLARELHDSVSQQLFAAMMMLSALEQAQRTETPEPYRKQLAMVAEIIIN
ASQSEMRAALLHLRPIISLEGKSLRKGIEQLLKEQLTKIKIELIWDVEDVHLNSSIEDHLFRIVQELLSNT
LRHAKAKELEVYLHQVDKNVLLRIVDDGVGDFMKEQSNKAGSYGLNIRERVVGMGGTVKIIISFKGQGTS
VEIKVPVIKEETASDQSNVSGAAGGRGSHHHHHH

EF3197

MNSHMVELFILMMERVGLIILLAFLLVNPYFKRVLLSREKMSKVQLILIFGLFAIISNFTGIEIAKNQ
IVPNNLLTYLSSNASIANTRTLVIGVSGLVGGPIVGSVAVGLIAGFHRVIQGGGHSFFYVVPASLIVGLIAG
FLGSRMAKQTVFSPAGFSAIVGACMEMIQMIFIFFSGDLSDGATLVRFIALPMLLNSVGTFFIMSILT
TTLKQEEQAKAVQTHDVLEAAETLPYFREGLNKNSKVAEIIKHVTKVSAISMTNSHQILAHVAGSD
HHIPELEVITELSRVLRTRMTIAHAKEEVGCSDFNCPQLQAAIIVPLFSSHQIVGTLKMYFTDPAQLTH
VEEQLAEGLGTIFSSQIELGEAEVQSKLLKEAEIKSLQAQVNPFFFNAINTISALMRKDSEKARKLLQ
LSKYFRGNLQGAQVTTIPVSQELEQVKAYLSLEQARFPNRYQVTFDVEEAVASEKVPPYAIQVLVENTIK
HAFGSRKENNQVRVVVKAKQKHLHVDVYDNGQGIPEERRLLLKKTVTSEKGTGTALENLSRRMANLYGS
EGCFLIENLETGGSHVHLEIPMEQEELDARINRAAGGRGSHHHHHH

EF1209

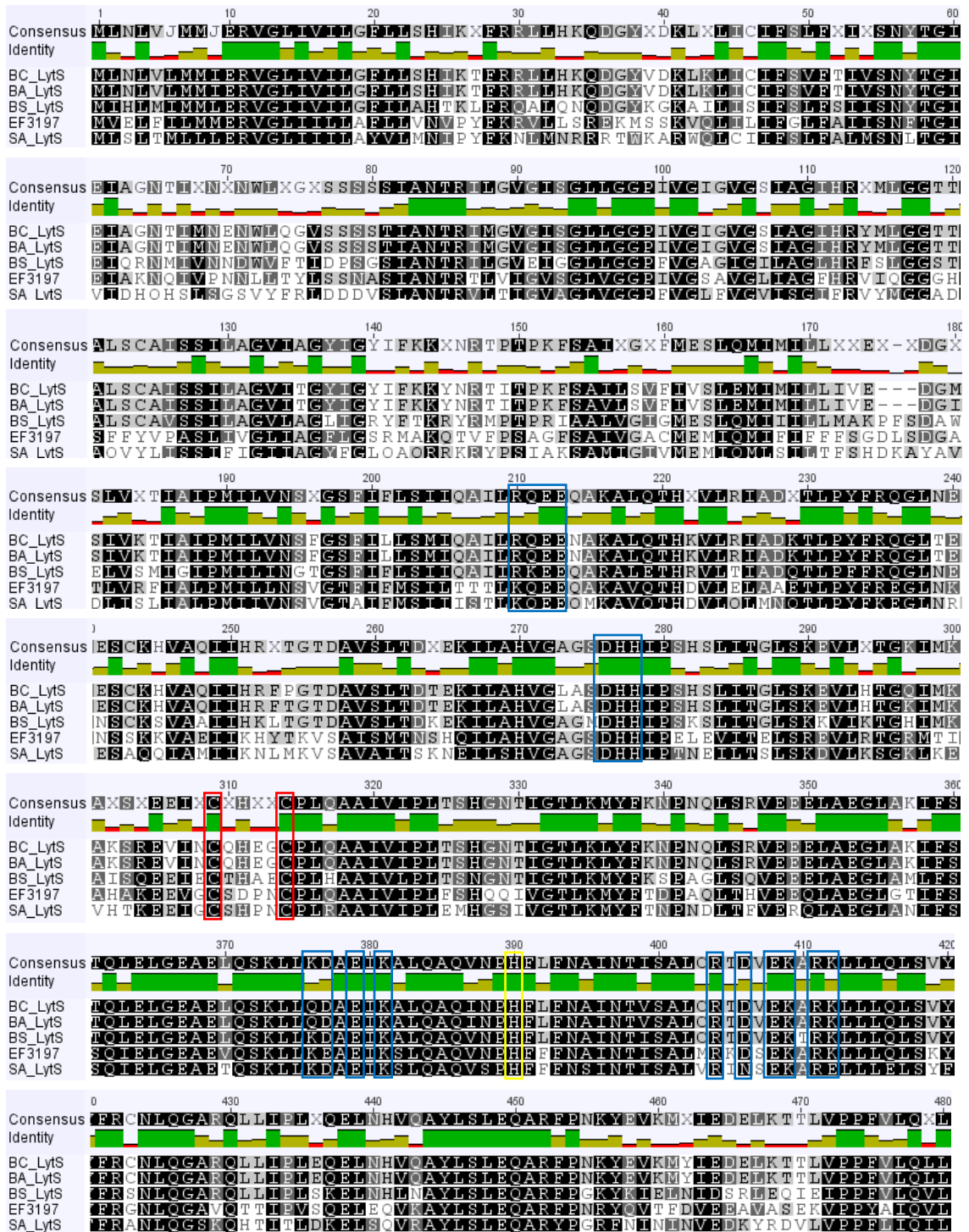
(NOTE: mutation highlighted in yellow, H to Y)

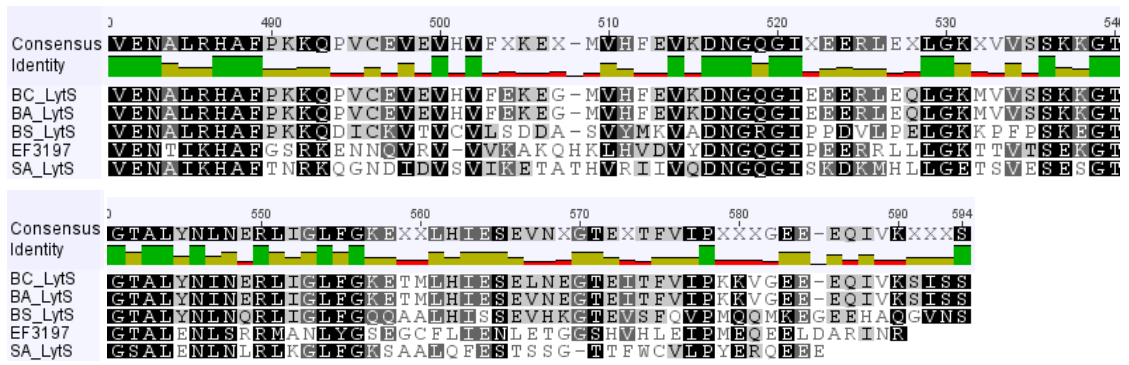
MNSHMKRGGKLCFLSVIFFIATLIIMITFFYGVTTVQTIKEVRKNQEQAALLAVGEQLAIEPNVIEALKND
HYSDELEAYTVRLGEIHQLDFIVIMMQGIRLTHPDRQKIGKHFEGGDEVRAKGEHLSVSQGSIGESL
RGFVVPYDQGGQIGVVAMGIKMTLSQLIERTKNDYTVSVLLSVGFGFILAIVVSYLKKQLHDLEPREI
ARLLEERNAMLEETKDAIILVIDTDQNILLANIEATKMYHNI TNSEENLLGKLSALVLSPEKLVVHSKTE
QFYRQNGQDYFVSIAPINVRKKTIGHVIFLKNATETFIVAEQLVSTTTYASALQSQSHEFMNMHVYIGL
VDLEDYEAALKHYLADLLKPEKEFAQRLAILVRNPILAGFLSGERIKFAEIKTQLAIEIYPEIPPNKRED
TQNLIAIYRYIYRFLMEQTLPEEIIETIDYQPGSLTTTYSFAYPKEQLERFEQEFFTSYLARLLENAEAT
LTWENQONNWLVRINVHYEGAEENEPIDYAAGGRGSHHHHHH

EF0373

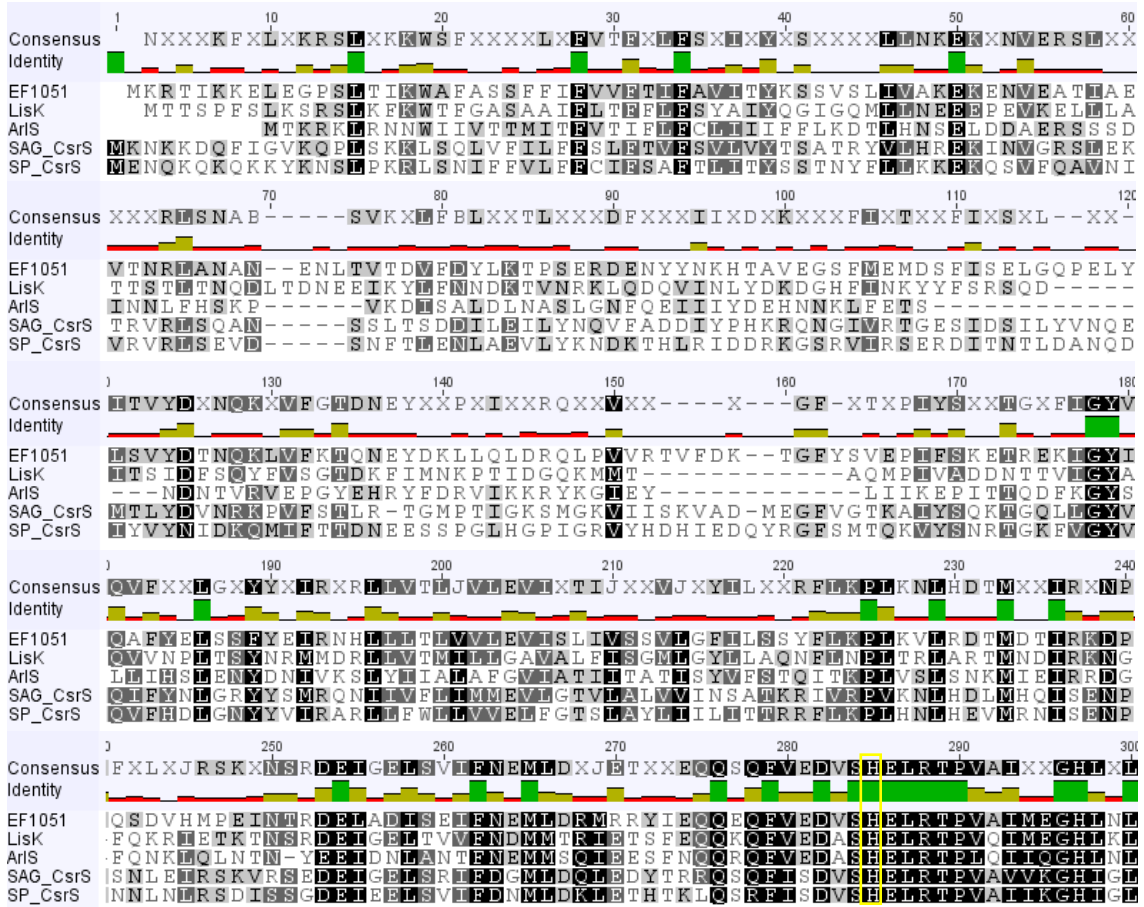
MNSHMTIKRRFFISYISAIITLASVLAVLSLASYITLGTVPSPQAYRMMNKQRPLTANEEESYLALDQ
LLKKSPLKLDTPLSKELKETIQTIEAKGLSVIRKNARFPYYS DN LVEKLSLVHIPPYEMNNIMPTGTLD
NAGRLYHYVKSDFHYLDGSNGSFIVLKRESSLFEFFTRWIIWMILLIILVAIAAAWLINKRLTKTTIEPL
EALAKATKTLGKDAQKENPFTPTSHQTVSTEVKQLQLSFEQMWQDLEANAEREKYEANRKELIANISHD
LKPITSIIGYVEGLMDGVANTEKKQRYLTVIHEKSLGLNDLIEELFLYSKLDLDRAVFTMEKTNFTRF
IAHILEEYRLEQELVITSVLPTEALYVQMDPTQMNRVITNLIQNSIKFADPTKEQLAFTISLTHNQTDLV
LTTDNGIGIDKKELPYLFRFYRVDKSRPTPTVKSGSLGLSIVKQIIDYHQGTITVTSKKGDTNVIITL
P LLEDEKAAGGRGSHHHHHH

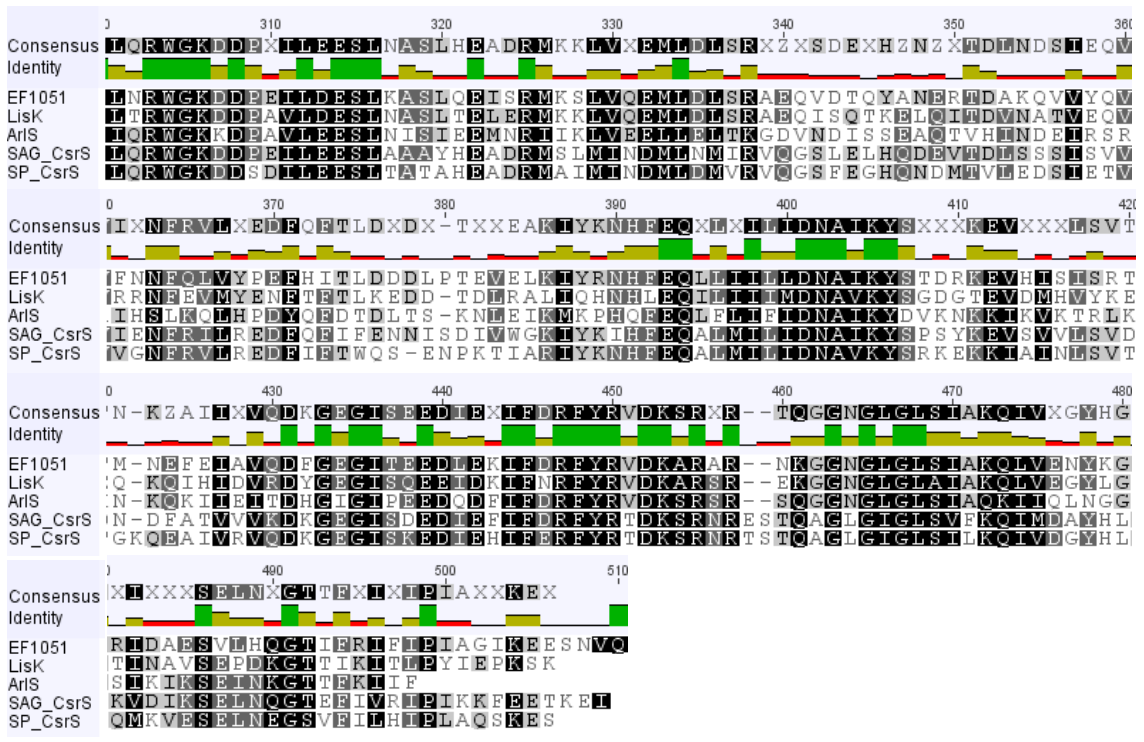
Appendix 9 Sequence alignment of EF3197 and LytS. Amino acid sequence of the proteins were aligned using ClustalW (Thompson *et al.*, 2002; Larkin *et al.*, 2007) and similarity between the sequences were analysed using Geneious Pro 4.6.5 (Drummond *et al.*, 2009). ■ 100% similar; ■ 80-100% similar; ■ 60-80% similar; □ <60% similar. LytS of: *Bacillus cereus* (BC_LytS); *Bacillus anthracis* (BA_LytS); *Bacillus subtilis* (BS_LytS); *Staphylococcus aureus* (SA_LytS). Conserved cysteine residues in red box (■); conserved group of charged amino acids in blue box (■); conserved histidine, putative in autophosphorylation in yellow box (■).



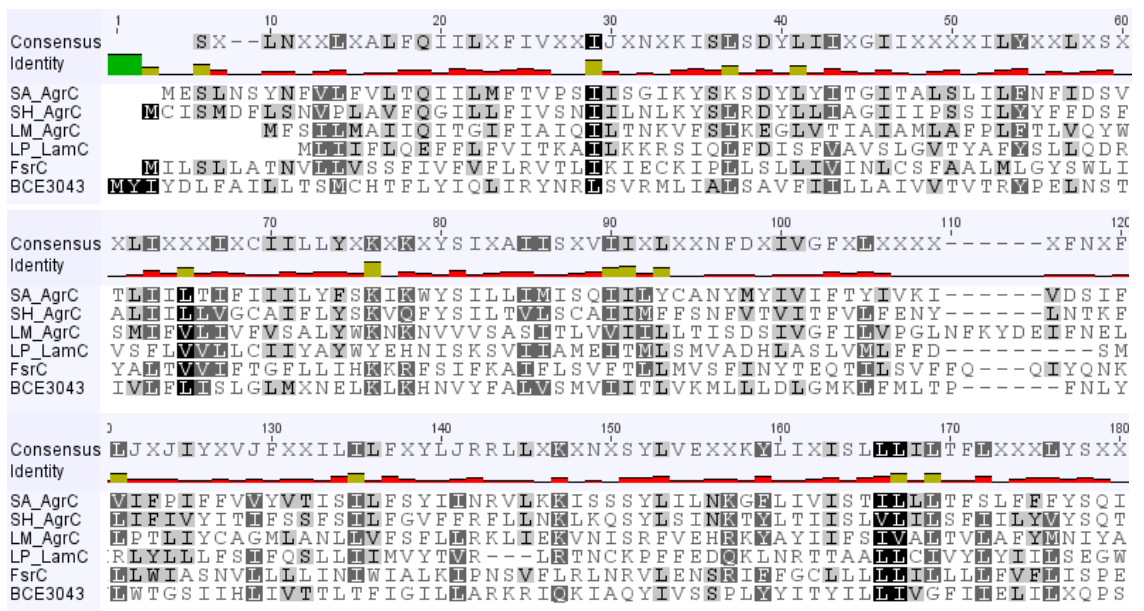


Appendix 10 Sequence alignment of EF1051, LisK, CsrS and ArlS. Amino acid sequence of the proteins were aligned using ClustalW (Thompson *et al.*, 2002; Larkin *et al.*, 2007) and similarity between the sequences were analysed using Geneious Pro 4.6.5 (Drummond *et al.*, 2009). ■ 100% similar; ■ 80-100% similar; □ 60-80% similar; □ <60% similar. EF1051 (sensor kinase of *E. faecalis*); LisK (LisK of *Listeria monocytogenes*); SAG_CsrS (CsrS of *Streptococcus agalactiae*); SP_CsrS (CsrS of *Streptococcus pyogenes*); ArlS (ArlS of *Staphylococcus aureus*). Phosphorylatable histidine in yellow box (□).





Appendix 11 Sequence alignment of EF1820 (FsrC) and AgrCs. Amino acid sequence of the proteins were aligned using ClustalW (Thompson *et al.*, 2002; Larkin *et al.*, 2007) and similarity between the sequences were analysed using Geneious Pro 4.6.5 (Drummond *et al.*, 2009). ■ 100% similar; ■ 80-100% similar; □ 60-80% similar; □ <60% similar. FsrC (EF1820 of *E. faecalis*), BCE3043 (putative AgrC of *B. cereus*), LP_LamC (LamC of *L. plantarum*), LM_AgrC (AgrC of *L. monocytogenes*), SA_AgrC (AgrC of *S. aureus*), SH_AgrC (AgrC of *S. haemolyticus*).



	190	200	210	220	230	240
Consensus	X S X X - - - - X S V J K S X X L I F F I G I T X L L X L L I J X X V X S N X X X X B L R X X R X K E X I E X L X X Y T X K I					
Identity						
SA_AgrC	N S D E - - - - A K V I R Q Y S F I F I G I T I F L S I I T F V I S Q F L L K E M K Y K R N Q E E I E T Y Y E Y T L K I					
SH_AgrC	P K L D - - - - T S S I K M Y G L I F F I G I I L F F T V I I I F I S N Y M I K B L R Y K R N M E E I E T Y Y E Y T L Q I					
LM_AgrC	G S I A G E D - G S V I K I N T L I F T G Y T I L L I V I V T V V I N T A T N B L K V Q N Q K E Q I E Q L Q D Y V T T I					
LP_LamC	K N Q A N - - - - V I F S N L I I L I L V M I G L F V I I Y N E Y L N N V K S K Y E V Q K Q R I Q I Q N D T R Y M N E I					
FsrC	I S P D E M R G F V T V N S S K L E L I S V G L F L I I I G L V I E A Y L E E Q R I N - - T Q L I N N L T I Y T E K I					
BCE3043	T E F L A K I N Q Q X S E V S Y N S A I I I F L I L L I I V L I S S H L S K E K L R E E H E K R L D K E L L D D Y V E K I					
	250	260	270	280	290	300
Consensus	E S I N N E X R K F E R H D Y V N I I T X I S E Y I R X D D M P G I X X Y X X X N I X E P X K K X I Q M X E X K J N X I X X					
Identity						
SA_AgrC	E A I N N E M R K F E R H D Y V N I I T I S E Y I R E D D M P G I R D Y E N K N I V P M K D N I Q M N A I K I N G I E N					
SH_AgrC	E S I N N E M R K F E R H D Y V N I I T I S E Y I R E D D M P G I R Q Y E N E N I V P M K D N I Q M K S I K I N G T E N					
LM_AgrC	E S I H R E M R V E R H D Y V N I I T I S D V G Y I D N N D M P G I R Y E E N N I V P I N K T I E S N N Y K I S L I Q N					
LP_LamC	E A H Y N E M R F E R H D Y Q N I I I S I D E Y I K T D D I E G I Q E Y Y Q K N I A P M S N R V L K E K Y N I E D I S R					
FsrC	E S I N N E E I A M F E R H D Y K N I I S I Q I A I S Y E D I L E I R I Y E E T I A P T K K I I D N E E F E I M K I N R					
BCE3043	E N M H D E I A S F E R H D Y M N V I S I E G I R T K N V K E I E K V Y Y D V I A P T F K T I N D H E L D I A K I S R					
	310	320	330	340	350	360
Consensus	I K M V R E I R K G I I T A K I L X A Q E K I X V X J V P E P - I E X I S M N X I D I X R I I G I I D N A T E A S V T					
Identity						
SA_AgrC	I K M V R E I R K G I I T A K I L R A Q E M S I P I S I E I P D E - V T H I N L N M I D I S R S I G I I D N A T E A S T E					
SH_AgrC	I K M V R A I R K G I I T K I L Q A Q E K N I P I S I E V P E L - V E H I E M N I D I S R I I G I I D N A T E A S E T					
LM_AgrC	I H V I A I R K G I I V A V K I I R A Q E L R I D A I I E V V E P - I D R I E S M D S I D I C K V G I I D N A V B A A L T					
LP_LamC	V K M K S I R S I I F S K I S Y A Q S Q E I E V H F D I K E P - I I D I T V N E I D I I A I G I I D N A T E A S V G					
FsrC	I K N M E I K A I I S M K I N T A K Q A K I K V I D V P E V F I L D T S I D L V V M I R L I A I I D N A T E N S I A K					
BCE3043	I H I E P E V R S V I R A K V G T A Q C Q Q I K V M E I D I P E N - I E S V S M A I I S F I R I I S V I V D N A T E E A V Q					
	370	380	390	400	410	420
Consensus	S E D X I I R I A F F X K D X S X X E I M M N S X K E X M P X I H K I I E E G E S T K G - - E X R C L G I S S I K E I X					
Identity						
SA_AgrC	I D D P I I R V A F I E S E N S V T E I M M N K C A D D I P R I H E I F Q E S E S T K G - - E G R C L G I S S I K E I A					
SH_AgrC	L E D A I I R I A F I N T D S V M E I M M N K C K E D M P R I H E I F Q E R E S T K G - - E N R C L G I S S I K E I T					
LM_AgrC	C E N P V I R I A F V K K G D S I I I V F A N S I L P V N M P P I Y K I I E E G E S T K G - - E G R C L G I A S I R E I M					
LP_LamC	H A D G E I M S A I F I E K N S T V S I I Q N N V F E Q L P P I W K I I K E A G E S T K G - - I N R G I G I S S I K E I V					
FsrC	S E L K M F A S I I F N K N E Q E E V I T N S V Q A E F D - F K V M K K T R E S R S N P E H E G W G I L Y V E E I V					
BCE3043	S E E K I I Q I A F F K M D S R Q Y E I M R N S S K S E G I D I Q R I V E R N H S S K E - - G T R G Y G I E S I R I M					
	430	440	450	455		
Consensus	D K T X N V L I D D T X E N G Y E I Q K E I X N N					
Identity						
SA_AgrC	D N A D N V L I D D T I I E N G E F I Q K V E I I N N					
SH_AgrC	D S T E N V L I D D T I I E N G Y E I Q K V E I I N N					
LM_AgrC	I K K Y S H V A I D D K V T D R E V I Q E E I M					
LP_LamC	N R N E M I I E R V L G A V E L Q R I T V K R V S Q N D					
FsrC	D F S D Q E D Q Q S F N E G A V T Q H I I E K N H N S K K V V N E					
BCE3043	I N K T N N A T E E T T Y V S P Y E T Q T F I I K E M K					

Appendix 12 DNA and amino acid sequences of recombinant *fsrB*. The N-terminal MNSH (blue) and C-terminal His-tag (red).

EF1821 (*FsrB*)

ATGAATTCCTAATCGATTGGATTCTAAAAAATATATGGATATGGATCAGGAAGATCAATCAGGAAAAACACAATGG
 ACAAAGTATTATCTAACCCTTTATTTTCTGGCTATTTAATCTTCTGATGATTCTGATTTTATCAGTTTATTTGGG
 ACGTTAAGCGAAACCTTTATTGTATACGTCGACTGATTTTTTTACGGCCTGTCCGAGGTGGCTGGCATGCAAAAAC
 AAATGGCTCTGCTCGTCTAGAAAGCATTGTTATCTATGTCGCCATACCATTTGTATTGAAAAATCTTCTGTGAGCTTA
 CCGTTATTATAAAATCTATTGATGTGCCCTTAGTCGTAATATTTTATTGGTATGCGCCACAAGGAACAGCGATT
 GAACCTGTTACAGCATCTGATTTAAACGTGCTCAAAAAGCAAAAGCCTTATAAGGGTGTGTTTACTTATTTTATGTAGT
 CTGTTTGTCAAAGAAAAGATTGCTTCAGTAATACTCTACGGTCTCGTCATCCAAGGTCTGATGATACTCCCTGTAA
 AAAATTTAATGAAGGAAGTGTTTTGCA

GCTGCAGGCGGTCTGGCAGCCACCATCACCATCACCAT

MNSLI DWILKNIMDMQEDQSGKTQWTKYYLTVYFSGLFNLLMILILSVLFGTLSETFIV
 YVVLIFLRPVAGGWHARKWLRLLESIVYVAIPFVLKNSSVSLPFIYKILMCLLVVLF
 YWYAPQGTAEIPVQPSDLNVLKQSLIRVCLLLILCSLFVKEKIASVILYGLVIQGLMILP
 VTKNLIEGVSFAAGGRGSHHHHHH