

**Functional characterisation of
bacterial tripartite ATP-
independent periplasmic (TRAP)
transporters**

Christopher Mulligan

A thesis submitted for the degree of
Doctor of Philosophy

University of York

Department of Biology

July 2008

Abstract

Tripartite ATP-independent periplasmic (TRAP) transporters are extracytoplasmic solute receptor (ESR)-dependent secondary transporters that are widespread in bacteria, but not eukaryotes. TRAP transporters are composed of an ESR component and 2 unequally sized integral membrane components. The large membrane component has 12 predicted transmembrane helices (TMHs) and is thought to form the translocation channel. The small membrane component is composed of 4 TMHs and is of unknown, but essential function. TRAP transporters combine an ESR, which is normally associated with ATP-binding cassette (ABC) transporters, with the utilisation of electrochemical gradients across the membrane, using an undefined mechanism.

Analysis of the data contained within the relational database, TRAPdb (www.trapdb.org), has revealed a number of new insights into TRAP transporter function, including the observation that marine-dwelling organisms have a propensity for high numbers of TRAP transporters – possibly indicating a role for Na^+ in the transport cycle.

Structural and functional analysis of TRAP transporters has been limited due to the recalcitrant nature of integral membrane proteins. The work presented here details the first overexpression and purification of the integral membrane proteins of a TRAP transporter. Using *Escherichia coli* and *Lactococcus lactis* expression systems, the integral membrane proteins from the *E. coli* TRAP transporter, YiaMNO, and the *Haemophilus influenzae* TRAP transporter, SiaPQM, have been expressed and purified.

The entire sialic acid-specific TRAP transporter, SiaPQM, has been functionally reconstituted into liposomes. The energetic requirements of transport by SiaPQM have been elucidated revealing that a Na^+ gradient in combination with an applied membrane potential are required to achieve maximal transport rates. Under normal conditions, SiaPQM is a unidirectional transporter, unlike most other secondary transporters, however, the presence of excess unliganded ESR is able to induce efflux of substrate from the proteoliposomes.

List of contents

The work carried out during this investigation was performed in laboratories in York, Leeds and Groningen, The Netherlands. Certain techniques were executed differently according to the location they were performed. Unless otherwise indicated, the protocols in the Materials and Methods section were performed in York.

List of contents	3
List of figures	11
List of tables	14
Acknowledgements	15
List of abbreviations	16
Chapter 1: Introduction	18
Overview of introduction	19
1.1 Bacterial cell envelope	20
1.2 Passive versus active transport	22
1.3 Diversity of active transport systems in bacteria	24
1.3.1 Primary transporters	24
1.3.2 Secondary transporters	25
1.3.3 Group translocation systems	25
1.4 Chemiosmotic theory and energy transducing membranes	26
1.5 Secondary active transporters in bacteria	28
1.5.1 Uniport, symport and antiport	28
1.5.2 Substrate and counter ion considerations	28
1.5.3 Electroneutral and electrogenic transport mechanisms	29
1.5.4 Classical features of secondary active transporters	32
1.6 Major facilitator superfamily (MFS)	34
1.6.1 Lactose permease as a model secondary transporter	34
1.6.1.1 General features	34
1.6.1.2 High resolution structure of lactose permease	35
1.6.1.3 Proposed transport mechanism	38
1.7 ATP-binding cassette (ABC) transporters	41
1.7.1 ABC transporters: an introduction	41

1.7.2	A short history of ABC transporters functional determination	42
1.7.3	<i>E. coli</i> maltose transport by MalEFGK ₂ as a model ABC system	43
1.7.3.1	The ESR domain – MalE	44
1.7.3.2	ATPase domain (NBD) – MalK	47
1.7.3.3	Transmembrane domains (TMD) – MalFG	50
1.7.3.4	High resolution structure of complete MalEFGK ₂ complex	52
1.7.4	Proposed mechanism of transport by ABC importers	56
1.8	Tripartite ATP independent periplasmic (TRAP) transporters	58
1.8.1	Discovery of the first TRAP transporter: DctPQM from <i>R. capsulatus</i>	58
1.8.1.1	C ₄ -dicarboxylate transport in bacteria	58
1.8.1.2	C ₄ -dicarboxylate transport in <i>Rhodobacter capsulatus</i>	60
1.8.1.3	Transport by DctPQM is dependent on $\Delta\Psi$ and not ATP hydrolysis	63
1.8.2	Phylogenetic analysis of the DctP, DctQ and DctM families of proteins	65
1.8.2.1	The DctP family of ESRs	65
1.8.2.2	The DctM family of integral membrane proteins	66
1.8.2.3	The DctQ family of integral membrane proteins	67
1.8.3	Genetic arrangement of TRAP transporters	68
1.8.4	Known and predicted substrates of TRAP transporters	70
1.9	SiaPQM: an <i>N</i> -acetylneuraminic acid TRAP transporter from <i>H. influenzae</i>	72
1.9.1	An introduction to sialic acids	72
1.9.2	Acquisition of sialic acid	73
1.9.3	Utilisation of sialic acid in bacteria	73
1.9.4	Bacterial sialic acid (Neu5Ac) transporters	74
1.9.5	<i>Haemophilus influenzae</i> and <i>N</i> -acetyl neuraminic acid	76
1.9.6	Characterisation of SiaPQM from <i>H. influenzae</i>	77
1.9.6.1	Characterisation of SiaPQM <i>in vivo</i>	77
1.9.6.2	<i>In vitro</i> characterisation of SiaP	78
1.9.6.3	High resolution crystal structure of SiaP	79
1.9.6.4	Ligand binding site and specificity determinants of SiaP:Neu5Ac interaction	81
1.10	Other ESRs working with TRAP integral membrane components	85
1.10.1	TAXI-type TRAP transporters	85
1.10.2	ABC-type ESR working with TRAP integral membrane components	87
1.11	Other notable TRAP transporters	87

1.11.1	A dimeric DctP-type ESR	87
1.11.2	A glutamate specific ESR-dependent secondary transporter from <i>R. sphaeroides</i>	88
1.12	Tripartite tricarboxylate transporter (TTT) family	90
1.13	Expression and purification of bacterial integral membrane proteins	91
1.13.1	Choice of expression strain	92
1.13.2	Choice of expression vector	93
1.14	Aims of this investigation	95
Chapter 2: Materials and methods		96
2.1	Suppliers	97
2.2	Table of primers	97
2.3	Strains and plasmids	97
2.4	Media and antibiotics	97
2.4.1	M9 minimal medium	97
2.4.2	Enhanced M9 minimal medium	97
2.4.3	Luria-Bertani (LB) medium	97
2.4.4	M17 medium	97
2.4.5	Solid media	101
2.4.6	Antibiotics	101
2.5	Buffers and solutions	101
2.5.1	SDS-PAGE	101
2.5.2	Western blotting	101
2.5.3	Agarose gel electrophoresis	101
2.6	General cloning techniques	102
2.6.1	Agarose gel electrophoresis	102
2.6.2	Polymerase chain reaction (PCR)	102
2.6.3	Preparation of DNA fragments for cloning	102
2.6.3.1	Preparation of plasmid DNA	102
2.6.3.2	Gel extraction	102
2.6.3.3	PCR clean-up	103
2.6.4	Ligation-dependent cloning	103
2.6.5	Ligation-independent cloning (LIC)	103
2.6.6	Vector backbone exchange (VBEx) into <i>L. lactis</i> pNZ8048 vector	104
2.6.7	Transformations	104

2.6.7.1	Making chemically competent <i>E. coli</i>	104
2.6.7.2	Transformation of chemically competent cells: freeze-thaw method	105
2.6.7.3	Transformation of chemically competent cells: heat shock method	105
2.7	Growth of bacteria	106
2.7.1	Initial trial for expression from pTTQ18-derived vectors in <i>E. coli</i>	106
2.7.2	Initial trial for expression from pBAD-derived vectors in <i>E. coli</i>	106
2.7.3	Initial trial for expression from pNZ8048-derived vectors in <i>L. lactis</i>	106
2.7.4	Standard timecourse experiment for expression from pTTQ18 derivatives	106
2.8	Preparation of bacterial extracts	107
2.8.1	Preparation of total membrane extracts using the water lysis method	107
2.8.2	Preparation of total membrane extracts using the a FastPrep machine	107
2.8.3	Preparation of total membrane extracts from <i>E. coli</i> and <i>L. lactis</i>	108
2.8.4	Separation of <i>E. coli</i> inner and outer membranes using sucrose density gradient	109
2.8.5	Periplasmic preparation using lysozyme treatment	109
2.9	Protein concentration determination	110
2.9.1	Bradford assay	110
2.9.2	Bicinchoninic acid (BCA) protein assay	110
2.9.3	Schaffner-Weissmann protein assay	110
2.9.4	Using A_{280} /extinction coefficient	111
2.10	SDS-PAGE	111
2.10.1	Molecular weight markers	111
2.10.2	Sample preparation for SDS-PAGE	111
2.10.3	Mini-SDS PAGE gels	112
2.10.4	Large SDS-PAGE gels	112
2.10.5	Staining/destaining SDS-PAGE gels	112
2.11	Western Blotting	113
2.11.1	Wet transfer of protein	113
2.11.2	Semi-dry transfer of protein	113
2.11.3	Processing the membrane after protein transfer (York)	113
2.11.4	Processing the membrane after protein transfer (Groningen)	114
2.11.5	Processing the membrane after protein transfer (Leeds)	115
2.12	Tryptophan fluorescence spectroscopy	115
2.13	Integral membrane protein detergent solubility screening	115

2.14	Protein purification techniques	116
2.14.1	Extraction of integral membrane proteins for purification by DDM solubilisation	116
2.14.2	Column-based Ni ²⁺ affinity chromatography	116
2.14.3	Column-based Ni ²⁺ affinity chromatography	117
2.14.4	Ni ²⁺ affinity chromatography using disposable gravity-flow columns	117
2.14.5	Ni ²⁺ affinity chromatography using HisTrap HP columns	117
2.14.6	Anion exchange chromatography	118
2.14.7	Size exclusion chromatography (SEC)	118
2.15	Reconstitution of integral membrane protein into liposomes	118
2.15.1	Purification of <i>E. coli</i> lipid extract and production of liposomes	118
2.15.2	Reconstitution of SiaQM into liposomes by the rapid dilution method	119
2.15.3	Reconstitution by the destabilized liposome method	120
2.16	<i>In vitro</i> transport assays	120
2.16.1	Preparation of proteoliposomes for <i>in vitro</i> [¹⁴ C]-Neu5Ac transport assays	121
2.16.2	Standard <i>in vitro</i> [¹⁴ C]-Neu5Ac transport assay	121
2.16.3	External buffers for formation of electrochemical gradients	121
2.16.3.1	Na ⁺ gradient alone ($\Delta\mu\text{Na}^+$)	121
2.16.3.2	Na ⁺ gradient plus membrane potential ($\Delta\mu\text{Na}^+ + \Delta\Psi$)	121
2.16.3.3	Na ⁺ gradient plus pH gradient ($\Delta\mu\text{Na}^+ + \Delta\text{pH}$)	122
2.16.3.4	Membrane potential alone ($\Delta\Psi$)	122
2.16.3.5	pH gradient alone (ΔpH)	122
2.16.4	Solute counterflow assay	122
2.17	<i>In vivo</i> transport assays	123
2.17.1	Preparation of cells for <i>in vivo</i> [¹⁴ C]-Neu5Ac transport assay	123
2.17.2	<i>In vivo</i> [¹⁴ C]-Neu5Ac transport assay: application of PMF ($\Delta\mu\text{H}^+$ and $\Delta\mu\text{Li}^+$)	123
2.17.3	<i>In vivo</i> [¹⁴ C]-Neu5Ac transport assay: application of SMF ($\Delta\mu\text{Na}^+$)	123
2.17.4	<i>In vivo</i> [¹⁴ C]-Neu5Ac transport assay: application of $\Delta\mu\text{Li}^+$	124
Chapter 3: Genome-wide analysis of TRAP transporter gene frequency and organisation		125
3.1	The TRAPdb relational database	126
3.2	Specific enrichment of TRAP transporters in marine bacteria	128

3.3	Genetic arrangement of <i>dctQ</i> and <i>dctM</i> genes	131
3.4	TRAP transporter systems with multiple ESR components	132
3.5	Recruitment of TRAP ESR domains for non-transport functions	135
3.6	UspA gene linked to TAXI TRAP transporter genes	136
3.7	A TRAP ESR functioning with an ABC transporter?	137
3.8	A TRAP transporter for malonate	138
3.9	Chapter summary	140
Chapter 4: Expression and purification of YiaM and YiaN from <i>E. coli</i>		142
4.1	4.1 Cloning and expression trial of the <i>yiaMN</i> genes using pTTQ18	143
4.2	4.2 Optimisation of growth and induction conditions for YiaM and YiaMN expression	147
4.2.1	IPTG concentration optimisation	147
4.2.2	Growth medium, temperature of growth and harvest time	148
4.3	Comparative expression timecourse of YiaN from pCM22 in 3 different expression strains; BL21 (DE3), C41 (DE3) and C43 (DE3).	152
4.4	Determination of YiaM and YiaN subcellular location by separation of <i>E. coli</i> membrane fractions	155
4.5	Differential solubilisation of YiaM and YiaN by a variety of detergents	157
4.6	Purification of YiaM and YiaN by immobilized metal affinity chromatography (IMAC)	160
4.7	Treatment of purified YiaM with heat denaturation and reducing agent	164
4.8	Chapter summary	167
Chapter 5: Expression and purification of SiaQM from <i>H. influenzae</i>		168
5.1	Cloning and expression trial of the <i>siaQM</i> genes using pTTQ18	169
5.2	Optimisation of growth and induction conditions for SiaQM expression	171
5.3	Determination of SiaQM cellular sublocation by separation of <i>E. coli</i> membrane fractions	173
5.4	Differential solubilisation of SiaQM by a variety of detergents	175
5.5	Purification of SiaQM by immobilized metal affinity chromatography (IMAC)	177
5.6	Further purification of SiaQM from inner membrane using anion exchange chromatography followed by IMAC	180

5.7	Cloning and expression trials of the <i>siaQM</i> genes using the arabinose inducible pBAD vector in <i>E. coli</i> and the nisin inducible pNZ8048 vector in <i>L. lactis</i> .	184
5.8	Optimisation of growth and induction conditions for SiaQM expression from MC1061 pBADcQM and MC1061 pBADnQM	188
5.9	Optimisation of growth and induction conditions for SiaQM expression from NZ9000 pNZcQM and NZ9000 pNZnQM	191
5.10	Purification of N-terminally decahistidine tagged SiaQM by immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC)	193
5.11	Chapter summary	197
Chapter 6: Reconstitution of SiaPQM and characterisation of N-acetylneuraminic acid transport mechanism		199
6.1	<i>In vivo</i> N-acetylneuraminic acid transport by SiaPQM	200
6.2	Cloning, expression trials and expression optimisation of the <i>siaP</i> gene using the arabinose inducible pBAD vector in <i>E. coli</i>	202
6.3	Purification of ligand-free N-terminally decahistidine tagged SiaP by immobilised metal affinity chromatography (IMAC) and anion exchange chromatography	206
6.4	Initial reconstitution of SiaPQM using different reconstitution methods	210
6.5	Characterisation of the energetic requirements for N-acetylneuraminic acid transport by SiaPQM	213
6.6	Requirement of a Na ⁺ gradient over just the presence of Na ⁺ for the transport of N-acetylneuraminic acid via SiaPQM	216
6.7	Investigation into N-acetyl neuraminic acid transport by SiaQM and SiaP orthologue, VC1779	217
6.8	Directionality of N-acetyl neuraminic acid transport by SiaPQM	223
6.9	Effect of increasing concentrations of SiaP on N-acetylneuraminic acid transport by SiaPQM <i>in vitro</i>	229
6.10	Effects of excess SiaP chase on N-acetylneuraminic acid transport and accumulation <i>in vitro</i>	231
6.11	Comparison of effects of SiaP chase and VC1779 chase on N-acetylneuraminic acid transport and accumulation <i>in vitro</i>	234
6.12	Chapter summary	236

Chapter 7: Discussion	238
7.1 Distribution and diversity of TRAP transporters	239
7.2 Expression and purification of TRAP integral membrane proteins	241
7.3 Basic structural requirements of TRAP transporters	245
7.3.1 ESR component	245
7.3.2 DctQ and DctM components	248
7.4 Energetic requirements and characteristics of Neu5Ac transport via SiaPQM	252
7.5 Proposed model for TRAP transporters	256
List of references	261

List of figures

1.1	Components of a typical bacterial cell envelope	21
1.2	Composition and energetic requirements of different classes of transporter	23
1.3	Energetic requirements and unique transport characteristics of secondary transport.	31
1.4	High resolution crystal structures of secondary transporters.	36
1.5	Substrate binding site and H ⁺ translocation channel of lactose permease.	39
1.6	Conformational changes of lactose permease during the proposed alternating access model.	40
1.7	High resolution crystal structure of MalE and topological differences between type I and type II ESRs.	46
1.8	Crystal structure and conformational changes of the ATPase domain, MalK.	49
1.9	Crystal structure of the intact MalEFGK ₂ complex	53
1.10	Domain interactions in the MalEFGK ₂ complex.	55
1.11	Proposed mechanism for solute transport by ABC uptake system	57
1.12	Composition of archetypal ABC, secondary and TRAP transporters	59
1.13	Topology modeling of DctQ and DctM from <i>R. sphaeroides</i>	64
1.14	Organisation of genes encoding TRAP transporters containing DctP homologues	69
1.15	Known and predicted TRAP transporters substrates	71
1.16	Dual function of Neu5Ac in <i>H. influenzae</i> and genetic arrangement of sialometabolic operon in <i>H. influenzae</i> and related species	75
1.17	Crystal structure of SiaP in ligand-free and ligand-bound forms	80
1.18	Binding site interactions between Neu5Ac2en and SiaP	82
1.19	Neu5Ac orthologues used in the biophysical analysis of SiaP binding determinants	83
1.20	Other ESRs interacting with TRAP integral membrane components	86
1.21	Crystal structure of dimeric TRAP ESR in complex with sodium pyruvate	89
3.1	Screen captures of TRAPdb web interface	127
3.2	Genetic organization of selected TRAP transporters contained within TRAPDb	133

3.3	Organisation of genes predicted to encode a TRAP transporter for malonate and related malonate metabolism-related genes	139
4.1	Effects of induction with IPTG on the growth of BL21 (DE3) pCM14	144
4.2	Expression of YiaM and YiaN from pTTQ18 based expression constructs in the presence (+) and absence (-) of IPTG induction	146
4.3	Effect of varying IPTG concentration on the accumulation of YiaM and YiaN after 3 hours growth	149
4.4	Optimisation of time of harvest after induction, medium used and temperature of growth for the production of YiaM and YiaN	151
4.5	Comparative timecourses of YiaN expression from BL21 (DE3) pCM22, C41 (DE3) pCM22 and C43 (DE3) pCM22	154
4.6	Comparison YiaM and YiaN content of mixed, inner and outer membrane fractions prepared from BL21 (DE3) pCM14 or BL21 (DE3) pCM22 cultures	156
4.7	Solubilisation trial of YiaM with 10 different detergents	158
4.8	Solubilisation trial of YiaN with 10 different detergents	159
4.9	IMAC purification of YiaM and YiaN from inner membrane	162
4.10	Effect of β -mercaptoethanol treatment and heat denaturation on the apparent YiaM ladder after IMAC purification	166
5.1	Expression of SiaQM from pTTQ18 based expression constructs in the presence (+) and absence (-) of IPTG induction	170
5.2	Optimisation of time of harvest after induction, growth medium used and temperature of growth for the production of SiaQM	172
5.3	Comparison SiaQM content of mixed, inner and outer membrane fractions prepared from BL21 (DE3) pCM23.1	174
5.4	Solubilisation trial of SiaQM with 10 different detergents	176
5.5	IMAC purification of SiaQM from inner membrane	178
5.6	Anion exchange chromatography of SiaQM from inner membrane	182
5.7	Analysis of fractions from anion exchange chromatography of SiaQM from inner membrane	183
5.8	IMAC purification of SiaQM from anion exchange eluate	185
5.9	Expression of SiaQM from the <i>E. coli</i> pBAD expression vector and <i>L. lactis</i> pNZ8048 expression vector	187

5.10	Optimisation of temperature of growth, inducer concentration and optical density (OD_{660}) of induction for the production of SiaQM from MC1061 pBADcQM and MC1061 pBADnQM	190
5.11	Optimisation of inducer concentration for the production of SiaQM from NZ9000 pNZcQM and NZ9000 pNZnQM	192
5.12	IMAC followed by SEC of SiaQM from membrane preparation from MC1061 pBADnQM	195
5.13	IMAC followed by SEC of SiaQM from membrane preparation from NZ9000 pNZnQM	196
6.1	<i>In vivo</i> Na^+ -driven transport of [^{14}C]-Neu5Ac by SiaPQM expressed in <i>E. coli</i> strain BW25113 $\Delta nanT$	201
6.2	Optimisation of temperature of growth, inducer concentration and optical density (OD_{660}) of induction for the production of SiaP from MC1061 pBADcP and MC1061 pBADnP	204
6.3	IMAC purification of SiaP from clarified whole cell lysate from MC1061 pBADnP	208
6.4	Anion exchange chromatography purification of SiaPH ₁₀ from Ni-NTA eluate	209
6.5	Na^+ -driven transport of [^{14}C]-Neu5Ac by SiaPQM reconstituted by three different methods	212
6.6	Energetic requirement for [^{14}C]-Neu5Ac by SiaQM containing proteoliposomes	215
6.7	Transport of [^{14}C]-Neu5Ac in the presence and absence of a Na^+ gradient	218
6.8	Neu5Ac concentration-dependent fluorescence change in VC1779	221
6.9	Uptake of [^{14}C]-Neu5Ac by SiaQM in the presence of $\Delta\mu Na + \Delta\Psi$ and either SiaP or VC1779	222
6.10	Comparison of the tertiary structure and surface exposed electrostatic characteristics of SiaP and VC1779	224
6.11	Effect on the accumulated levels of [^{14}C]-Neu5Ac upon addition of excess cold Neu5Ac to assess the directionality of SiaQM	227
6.12	Counterflow assays in the presence and absence of SiaP	228
6.13	Transport activity of SiaQM in the presence of increasing concentrations of SiaP with a fixed [^{14}C]-Neu5Ac concentration	230
6.14	Effect of addition of excess SiaP on accumulated Neu5Ac	233
6.15	Effect of addition of excess SiaP and excess VC1779 on accumulated	233

7.1	Proposed models for the transport mechanism TRAP transporters	257
-----	---	-----

List of tables

2.1	Details of oligonucleotide primers used for amplification of DNA fragments.	98
2.2	Details of plasmids created or used during this study.	99
2.3	Details of bacterial strains used during this study.	100
3.1	Organisms containing high frequencies of TRAP transporters ordered by their TRAP density	130

Acknowledgements

I would firstly like to thank my supervisor Gavin Thomas for his support, advice and enthusiasm throughout the course of my PhD. I would also like to thank Gavin for arranging for my research visits to Leeds and Groningen. I would like to show my appreciation to past and present members of the Thomas lab for their help and support.

I would like to thank Prof. Peter Henderson for allowing me to work in his lab for 3 months, and members of his group for helping me with the pTTQ18 expression work. I would especially like to give thanks to Kim Bettany and Nick Rutherford for their advice and guidance.

My thanks also go to Prof. Bert Poolman and his entire group for hosting my research visit to his lab in Groningen. I would especially like to thank Eric Geertsma for introducing me to the wonder that is *Lactococcus lactis*, arabinose expression systems and the art of membrane protein reconstitution.

I would like to thank the members of my PhD training committee, Marjan van der Woude and Dale Sanders for advice and helpful discussions.

I would like to thank the Society for General Microbiology (SGM) for funding my visit to Groningen through the President's Fund and the BBSRC for funding my project.

I would like to thank my friends for all the good times.

I would especially like to thank Ally for her love, support and patience.

Last but by no means least; I would like to thank my parents, Kath and Martyn. Without their love and support I wouldn't be who or where I am today.

List of abbreviations

APS	Ammonium persulfate
β -OG	<i>n</i> -octyl- β -D-glucopyranoside
CBB	Coomassie Brilliant Blue
CMC	Critical micellar concentration
CMP-GlcNac	Cytidine monophosphate – <i>N</i> -acetylglucosamine
DDM	<i>n</i> -dodecyl- β -D-maltoside
DEER	Double electron-electron resonance
DM	<i>n</i> -decyl- β -D-maltoside
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ESR	Extracytoplasmic solute receptor
FPLC	Fast protein liquid chromatography
Fructose-6-P	Fructose-6-phosphate
GlcN-6-P	Glucosamine-6-phosphate
GlcNAc-6-P	<i>N</i> -acetylglucosamine-6-phosphate
HPLC	High performance liquid chromatography
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl- β -D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LB	Luria-Bertani medium
LIC	Ligation independent cloning
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
MALDI-TOF	Matrix assisted laser desorption/ionization - time of flight
ManNAc	<i>N</i> -acetylmannosamine
MCD	Malonyl-CoA decarboxylase
NBD	Nucleotide binding domain
Neu5Ac	<i>N</i> -acetylneuraminic acid
NTA	Nitrilotriacetic acid
PBST	Phosphate buffered saline + Tween 20

PEP	Phosphoenol pyruvate
PMF	Proton motive force
PTS	Phosphotransferase system
PVDF	Polyvinylidene fluoride
RMSD	Root mean square deviation
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMF	Sodium motive force
TBE	Tris-borate-EDTA buffer
TBST	Tris buffered saline + Tween 20
TBSTT	Tris buffered saline + Tween 20 + Triton X-100
TEMED	N, N, N', N'-tetramethylethylenediamine
TMD	Transmembrane domain
TMH	Transmembrane helix
VBEx	Vector backbone exchange

Chapter 1

Introduction

Overview of introduction

All living cells require the ability to transport a multitude of molecules across their cell envelope without compromising their integrity. Bacteria, in particular, are able to colonise a huge variety of environments with vastly different concentrations of various compounds, both beneficial and toxic, that need to be actively transported against a concentration gradient. Bacteria have therefore evolved a number of active transport systems that have different structural components, energetic requirements, specificities and substrate affinities.

One particularly interesting family of bacterial active transporters is the tripartite ATP-independent periplasmic (TRAP) transporters. These transporters are novel in that they combine certain characteristics of two other major families of transporters; secondary transporters and ATP-binding cassette (ABC) transporters. But in contrast to these two families there is very little structural and functional information on TRAP transporters.

To introduce TRAP transporters this chapter will be split into 4 sections. First, the bacterial cell envelope will be introduced along with passive versus active transport of solute across the membrane, diversity of active transporters and details of the chemiosmotic theory of energy transducing membranes. The second section will describe secondary transporter structure and mechanism in more detail. The model secondary transporter lactose permease from *Escherichia coli* will be used to illustrate the characteristics of this transporter family. Thirdly, the structure and mechanism of ABC transporters will be introduced. The maltose ABC transporter from *E. coli* will be used to exemplify their general characteristics. The final section will describe the known structural and functional details of TRAP transporters and how they relate to secondary and ABC transporters. The sialic acid TRAP transporter from *Haemophilus influenzae* and the C₄-dicarboxylate TRAP transporter from *Rhodobacter capsulatus* will be used to typify the characteristics of TRAP transporters.

1.1 Bacterial cell envelope

Bacteria can be divided into two large groupings depending on the nature of their cell envelope; Gram-positive or Gram-negative. The cell envelope of Gram-negative bacteria consists of a cytoplasmic (inner) membrane, a periplasm, a peptidoglycan layer and an outer membrane (Fig. 1.1a), whereas Gram-positive bacteria possess a cytoplasmic membrane and a peptidoglycan layer that is generally much thicker than that of Gram-negative bacteria (Fig. 1.1b). The primary function of the bacterial cell envelope is to separate the cytoplasm from the environmental milieu while still maintaining the ability to readily transport molecules into and out of the cell.

The outer membrane of Gram-negative bacteria is an asymmetric bilayer consisting of an inner leaflet of phospholipids and an outer leaflet composed of lipopolysaccharide (LPS) (Nikaido, 2003). LPS is a glycolipid composed of the glucosamine-based phospholipid, lipid A attached to an oligosaccharide core domain followed by O-antigen which consists of repeating oligosaccharide units (Bos *et al.*, 2007). The majority of the LPS biogenesis machinery has been elucidated and has revealed that the LPS component parts are synthesized on the inner leaflet of the inner membrane, transported across the inner membrane and transferred to the outer membrane (for review of LPS biogenesis, see Bos *et al.* (2007)). The primary function of the outer membrane is as a selective barrier and it is impermeable to hydrophilic compounds. However, movement of hydrophilic compounds of up to 600 Da across the outer membrane can be mediated by outer membrane proteins (OMPs) that can be non-specific, for example OmpF from *E. coli* (Nikaido & Rosenberg, 1983), or specific for a particular compound or group of compounds, for example maltose and maltodextrins via LamB from *E. coli* (Schirmer *et al.*, 1995). The outer membrane is not energized, therefore movement of compounds is generally via passive diffusion (for an excellent review of outer membrane permeability see Nikaido (2003)).

The cytoplasmic (inner) membrane of both Gram-positive and Gram-negative bacteria is primarily composed of phospholipids and embedded proteins. The

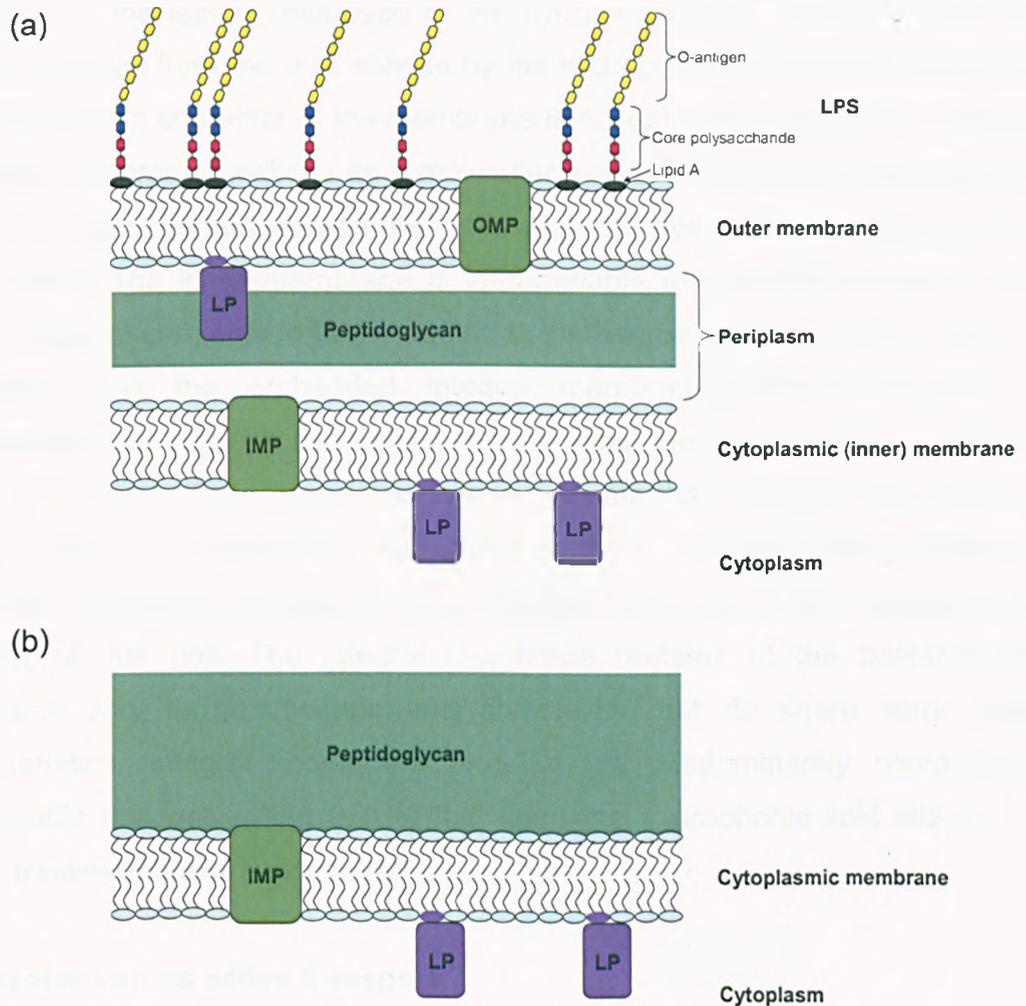


Figure 1.1 Components of a typical bacterial cell envelope. (a) Gram negative bacterial cell envelope and (b) Gram positive bacterial cell envelope indicating the relative locations of the cytoplasmic (inner) membrane, periplasm, peptidoglycan, outer membrane and lipopolysaccharide (LPS). LP represents lipoproteins, IMP stands for inner membrane protein, OMP is outer membrane protein.

phospholipid element of the membrane is composed of two leaflets of lipids with the centre of the leaflet composed of the hydrophobic fatty acid tails from each leaflet separated from the bulk solvent by the hydrophilic phosphatidyl headgroup. The hydrophobic character of the membrane is essential for its functions. Although the outer membrane functions as a molecular sieve, it is the inner membrane that forms the major barrier between the cytoplasm and hydrophilic compounds in the environment. The inner membrane is impermeable to electrolytes, which allows electrochemical gradients to be formed that can then be used to perform work. The membrane and the embedded integral membrane proteins regulate the concentrations of ions and hydrophilic solutes within the cell by either an uptake or efflux mechanism. The integral membrane proteins that perform these tasks are termed carriers or transporters and come in many varieties, which reflects the multitude and different abundance of substrates that need to be transported into, and out of, the cell. The integral membrane proteins of the bacterial inner membrane vary in size, shape and complexity, but do share some similar characteristics. Integral membrane proteins are predominantly composed of hydrophobic residues (often >70%) that span the hydrophobic lipid bilayer in α -helical transmembrane helices (TMHs).

1.2 Passive versus active transport

There are two general mechanisms by which molecules and ions can traverse the inner membrane; passive or active transport. Passive transport is diffusion of the substrate across the membrane down the concentration gradient, resulting ultimately in substrate equilibrium. Diffusion of most molecules and ions across the membrane is very slow, however, in most cases this process is catalyzed by carriers (carrier-facilitated diffusion) or pores (pore-facilitated diffusion). Pore-facilitated diffusion in bacteria is quite a rare occurrence. A well documented example is GlpF in *E. coli* (Sweet *et al.*, 1990). GlpF is a member of the aquaglyceroporin family of proteins for which there is a high resolution crystal structure available (Fu *et al.*, 2000). Transit of glycerol through GlpF is powered by the substrate gradient alone as glycerol kinase phosphorylates the transported glycerol as soon as it enters the cell, thus maintaining the inwardly-directed glycerol gradient (Fig. 1.2a). An example of carrier-facilitated diffusion is the K^+ -

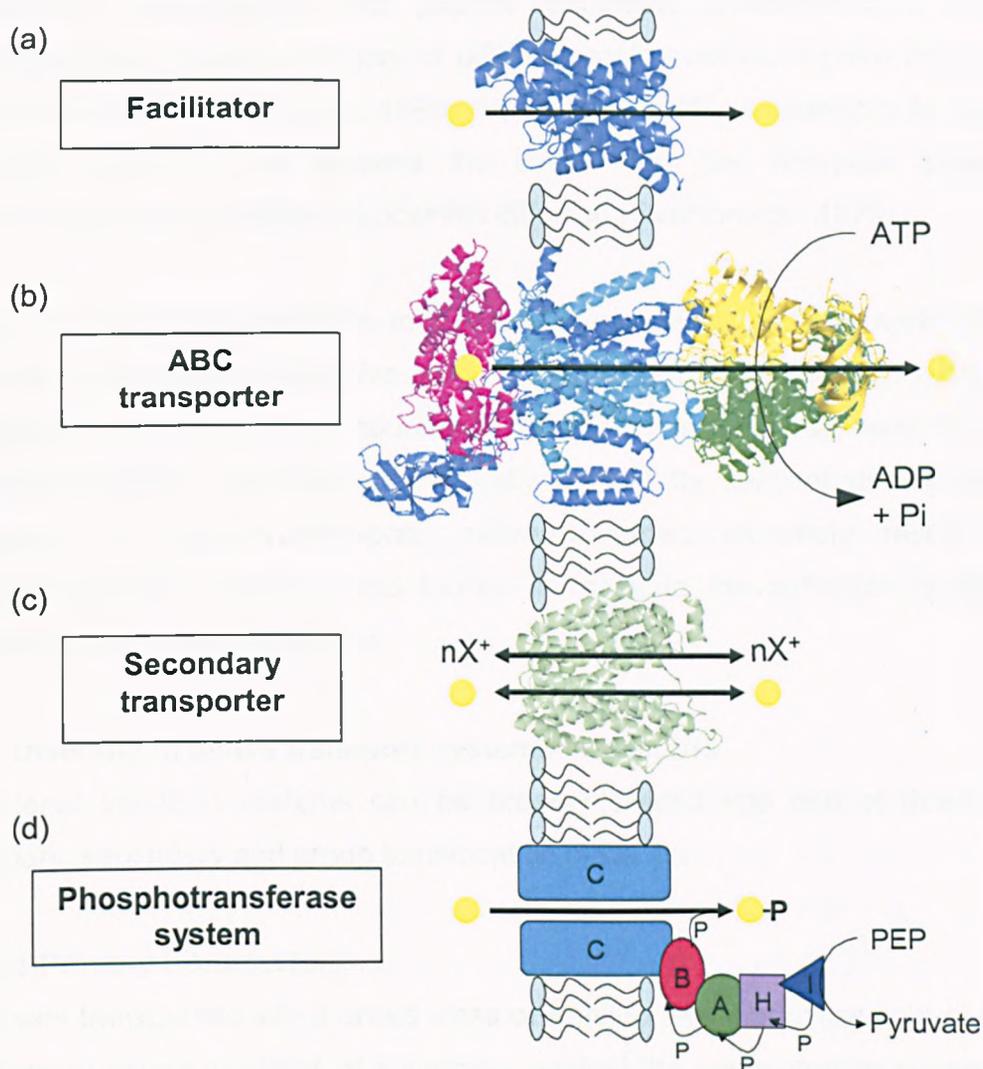


Figure 1.2 Composition and energetic requirements of different classes of transporter. (a) Facilitators represented by the aquaglyceroporin, GlpF, powered only the substrate gradient (Fu et al., 2000). (b) Primary transporters represented by the maltose ABC transporter, MalEFGK₂, powered by the binding and hydrolysis of ATP (Oldham et al., 2007). (c) Secondary transporters represented by lactose permease, utilizes existing electrochemical gradients (where X is usually H⁺ or Na⁺) (Abramson et al., 2003). (d) Schematic of components of a typical phosphotransferase (PTS or group translocator) system (Saier et al., 2005). Phosphoenol pyruvate (PEP) is used to start a phosphorelay between Enzyme I (I), HPr (H), Enzyme IIA (A) and Enzyme IIB (B) eventually phosphorylating the substrate that crossed the membrane through Enzyme IIC (C). P is the phosphate group. The yellow circle represents the substrate in each case.

ionophore, valinomycin. This peptide antibiotic, produced from strains of *Streptomyces*, adopts a number of different conformations in polar and non-polar environments (Perkins *et al.*, 1990). Upon binding K^+ , valinomycin forms a rigid, lipophilic structure that screens the cation from the non-polar phospholipid environment and facilitates its downhill diffusion (Ovchinnikov, 1979).

The majority of transporters in bacteria are active transport systems. Active transport, in contrast to passive transport systems, requires energy, which can be obtained from a variety of sources. This energy is used to power transport of substrate across the membrane usually against its concentration gradient. As opposed to passive transport, active transport ultimately results in the accumulation of substrate and further increase in the substrate concentration gradient across the membrane.

1.3 Diversity of active transport systems in bacteria

Bacterial transport systems can be broadly divided into one of three classes: primary, secondary and group translocation systems.

1.3.1 Primary transporters

Primary transporters are a broad class of transporters that utilize light or chemical energy to power transport of substrates against the concentration gradient, which can in turn result in the formation of an electrochemical gradient. Primary transport systems include bacteriorhodopsin, which is a small, 7 TMH containing integral membrane protein that upon illumination catalyses the translocation of a proton across the membrane (Lanyi, 1993). Bacteriorhodopsin is produced in halophilic archaea in the presence of low oxygen, and generates a proton motive force (pmf) when no respiration is taking place in order to power oxidative phosphorylation. An example of a primary transporter that extrudes H^+ across the membrane by a completely different mechanism are the proteins of the electron transport chain, which uses the redox potential as an energy source. A well studied example of a primary transporter that uses chemical energy are the P-type ATPases. These adenosine triphosphate (ATP)-driven ion pumps are multisubunit complexes found ubiquitously from bacteria to animal cells (Kuhlbrandt, 2004). P-type ATPase

complexes pump a number of different substrates including H^+ , K^+ , Na^+ and Ca^{2+} forming electrochemical gradients (Kuhlbrandt, 2004). The formation of these electrochemical gradients can power processes such as transport, or are used in a different capacity, for example, in inducing skeletal muscle contraction in animal systems (Siebers *et al.*, 1992, Kuhlbrandt, 2004). The primary transporters most pertinent to this project are the ATP-binding cassette (ABC) transporters that are found ubiquitously (Fig. 1.2b (Davidson *et al.*, 2008)). ABC transporters are multisubunit complexes that are able to transport a wide range of different substrates, for example, sugars (MalEFGK (Davidson & Nikaido, 1991)), amino acids (HisPQM (Ames & Nikaido, 1978)), vitamins (BtuCDF (Hvorup *et al.*, 2007)), oligopeptides (OppABCDF (Tynkkynen *et al.*, 1993)) and phospholipids (MsbA (Zhou *et al.*, 1998)). ABC transporters utilize the binding and hydrolysis of ATP to power transport of substrate across the lipid bilayer.

1.3.2 Secondary transporters

Secondary transporters (Fig. 1.2c) utilize the free energy present in existing ion gradients to power transport of a substrate across the membrane. The electrochemical gradient of the counter ion (usually H^+ or Na^+) is used to drive accumulation of a substrate against its concentration gradient. These transporters are involved in both the uptake and efflux of compounds from the cell and are known to transport a number of different compounds such as amino acids (almost all (Heyne *et al.*, 1991)), sugars (LacY (Abramson *et al.*, 2003)), antimicrobial compounds (tetracycline via TetA (Kaneko *et al.*, 1985)) and ions (NhaA (Padan & Schuldiner, 1994)).

1.3.3 Group translocation systems

Group translocation systems, otherwise known as phosphotransferase systems (PTS), couple the chemical modification of the substrate with transport into the cell (Fig. 1.2d). These structurally complex, multi-protein complexes are found ubiquitously in eubacteria, but not in archaea or eukaryotes. There is a wide variation in the number of PTS transporters in different bacteria, for example *E. coli* has 38, whereas, *Mycoplasma genitalium* has only two (Saier *et al.*, 2005). The substrate range of these transporters is limited to sugars, including glucose,

ascorbate, glucosides, maltose and *N*-acetylglucosamine (Saier et al., 2005). The substrate is transported into the cell by an integral membrane protein in the complex and then phosphorylated by a complex of cytoplasmic enzymes using phosphoenol pyruvate (PEP) as a substrate (Saier et al., 2005).

1.4 Chemiosmotic theory and energy transducing membranes

To fully appreciate the energetic requirements and physiological relevance of the different classes of transporter it is necessary introduce of the energy transducing membrane in more detail. In the case of bacteria, the cytoplasmic membrane is the energy transducing membrane.

Generally, in bacteria there are two forms of potential energy used to perform work; the energy potential in the hydrolysis of ATP and the electric/concentrative potential of the proton motive force (pmf) or sodium motive force (smf). These energy sources are intimately linked in the bacterial cell and this is underpinned by the chemiosmotic theory first postulated by Mitchell in the 1960s.

Bacteria produce this energy by a number of different mechanisms. Fermentative bacteria produce ATP by substrate level phosphorylation, this can be used to power a number of other reactions and mechanisms such as biosynthesis and transport of compounds across the cell envelope. Respiratory bacteria can also obtain ATP from substrate level phosphorylation, but also produce a pmf by way of the electron transport chain, or by light-driven ion pumps in photosynthetic bacteria. The cytoplasmic membrane-associated electron transport chain and light driven ion-pumps use energy (either redox potential or light) to extrude protons (H^+) across the membrane, thus forming the pmf (Nicholls, 2002). The pmf can power the production of ATP in the cytoplasm via the F_1F_0 ATP synthase and can also be directly coupled to the transport of compounds across the cytoplasmic membrane (Nicholls, 2002, Harold, 1986). The pmf can also drive the formation of other electrochemical gradients, most notably, the sodium gradient in *E. coli* is formed by the action of an H^+ - Na^+ exchanger, NhaA (Padan & Schuldiner, 1994).

The basis of the chemiosmotic theory is the formation of the pmf (also denoted as Δp), which is formed by a gradient of protons ($\Delta\mu_{H^+}$) across the bacterial cytoplasmic membrane that is inherently impermeable to ions. There are two components to the pmf that are important when considering its functions; the concentrative difference (potential) of protons across the membrane known as the pH gradient (ΔpH) and the electrical difference (potential) across the membrane, which is known as the membrane potential ($\Delta\Psi$). The sodium motive force (smf) is analogous to the pmf, but instead of a ΔpH , sodium forms a gradient across the membrane ($\Delta\mu_{Na^+}$) which acts similarly to the concentrative potential of protons. In general under physiological conditions, the cytoplasm is negatively charged and alkaline when compared to the outside of the cell.

Mitchell defined the pmf (with units of mV) as;

$$\Delta p = \Delta\Psi - (2.3RT/F)\Delta pH$$

Where R is the gas constant ($J.K^{-1}.mol^{-1}$), T is the absolute temperature and F is the Faraday constant. At 20°C, the term $2.3RT/F$ is ~58.8. A similar equation can be derived for the sodium motive force;

$$\Delta smf = \Delta\Psi - (2.3RT/F)\Delta pNa$$

The electrochemical potential can also be related to the change in Gibb's free energy (ΔG), the membrane potential and the concentrative gradient across the membrane using the following equation;

$$\Delta G (kJ mol^{-1}) = -mF \Delta\Psi + 2.3RT \log_{10}([X^{m+}]_B/[X^{m+}]_A)$$

where m is the valence of the ion, F is the Faraday constant, R is the gas constant and T is the absolute temperature. X^{m+} is a cation moving down an electrochemical potential from a concentration of $[X^{m+}]_B$ to $[X^{m+}]_A$. For the specific cases of the pmf and smf the following equations apply;

$$\Delta G \text{ (kJ mol}^{-1}\text{)} = \Delta\mu\text{H}^+ = -F \Delta\Psi + 2.3RT\Delta\text{pH}$$

$$\Delta G \text{ (kJ mol}^{-1}\text{)} = \Delta\mu\text{Na}^+ = -F \Delta\Psi + 2.3RT\Delta\text{pNa}$$

The free energy (ΔG) of ATP hydrolysis is negative, therefore it is termed energetically favourable (downhill) and can be used to power energetically unfavourable (uphill) reactions. This is the same situation as the pmf (and smf) because the free energy change as H^+ (or Na^+) moves back into the cell is negative, therefore can power unfavourable reactions such as the movement of a compound into the cell against its concentration gradient.

1.5 Secondary active transporters in bacteria

The defining feature of a secondary transporter is that it couples the energetically unfavourable transport of one substrate against its gradient with the energetically favourable transport of another substrate (the counter ion, usually H^+ or Na^+). This, however, is a simplified view, as secondary transport mechanisms are complex and variable depending on the substrate and counter ion involved.

1.5.1 Uniport, symport and antiport

Secondary transporters can be split into 3 broad categories based on the direction of flow of the substrate in relation to the counter ion. A substrate can undergo uniport where the transport of the substrate is independent of any counter ion. An example of a transporter using this mechanism is the lysine transporter from *Bacillus stearothermophilus* (Heyne et al., 1991). A substrate can undergo symport in which transport of the substrate is accompanied by at least one other solute in the same direction. An example of this is lactose permease from *E. coli* (Abramson et al., 2003). Substrate antiport is where transport of substrate is accompanied by transport of at least one other solute in the opposite direction. An example of this is the glycerol-3-phosphate:inorganic phosphate exchanger (GlpT) from *E. coli* (Huang et al., 2003).

1.5.2 Substrate and counter ion considerations

The mechanism of secondary transport is more complicated than simply uniport,

symport or antiport because the transport mechanism is also dictated by the charge of the substrate, the charge of the counter ion and the stoichiometry of the two during the transport cycle. The substrate specificities for secondary transporters are diverse and there are many examples of substrates that are neutral, positively and negatively charged. This means that each transporter has different energetic requirements, for example, different numbers of counter ions to allow transport. The nature of the counter ion further complicates matters. In the majority of cases the counter ion is H^+ , Na^+ , Li^+ or occasionally a transporter is capable of using all three (MelB (Pourcher *et al.*, 1995)). The counter ion used is often determined by the nature of the organism and its environment. In the case of alkalophilic bacteria, Na^+ is the preferential counter ion because in an environment of pH 10-11 the cytoplasm is maintained at \sim pH 9, therefore there is no inwardly directed proton gradient. To maintain this cytoplasmic pH, the Na^+ that is symported with the substrate in secondary active transport is used to exchange for H^+ via an H^+ - Na^+ antiporter (Harold, 1986). Marine bacteria also prefer Na^+ due the abundance of Na^+ in its environment and the fact that seawater has a slightly alkaline pH (Kogure, 1998, Tsuchiya & Shinoda, 1985).

There are examples of very similar proteins using different counter ions indicating that they have different mechanisms despite their similarity. A prime example of this is L-glutamate transport in *E. coli* and *B. stearotherophilus*. *E. coli* possesses two secondary transporters for L-glutamate; one that uses a $\Delta\mu H^+$ (GltP) and one that uses a $\Delta\mu Na^+$ (GltS) (Poolman & Konings, 1993, Tolner *et al.*, 1995a). *B. stearotherophilus* possesses only one L-glutamate transporter, GltT, which is \sim 60% identical to GltP, and has been shown to use gradients of both $\Delta\mu H^+$ and $\Delta\mu Na^+$ in combination (Tolner *et al.*, 1995b). Another interesting observation from this study was that GltT from *B. stearotherophilus* utilized a $\Delta\mu H^+$ when expressed in *E. coli*, but if homologously expressed was found to use a $\Delta\mu Na^+$ to power L-glutamate uptake. This indicates that it is not only the environment of the bacterium and the protein itself influencing the nature of the counter ion, but also the protein-lipid interactions that could have an effect on the transport mechanism.

1.5.3 Electroneutral and electrogenic transport mechanisms

There are a number of different combinations of substrate and counter ion with

different energetic consequences. For the purposes of this introduction, only the mechanisms of bacterial uptake systems (uniporters and symporters) will be described. However, it should be noted that antiporters are just as abundant. An excellent review of secondary transport mechanism is available by Poolman and Konings (1993).

Electroneutral transport is defined as transport of a substrate that results in no net charge transfer across the membrane. Conversely, electrogenic transport results in the net transfer of charge across the membrane and the creation of electrical potential. This has an impact on the energetic requirements of secondary transport, for example, the membrane potential may have a positive or negative effect on the uptake of substrate.

The simplest form of secondary transport is electroneutral substrate uniport, but this could really be argued to be facilitated diffusion. An example of this is glycerol transport by GlpF in *E. coli* (Fig 1.3a i, (Sweet et al., 1990)). In this situation, the driving force for transport is derived only from the substrate gradient itself.

Electrogenic substrate uniport occurs when the substrate is cationic (Fig. 1.3a ii). An example of this is lysine and arginine transport in *B. stearothermophilus* (Heyne et al., 1991). The driving forces for transport are the substrate gradient, the membrane potential and the valency of the substrate, i.e. the more positively charged the substrate is the stronger the effect of the membrane potential will be.

Electroneutral substrate-cation symport (Fig. 1.3a iii) occurs when an anionic substrate is in a 1:1 stoichiometry with a cationic counter ion. The membrane potential has no effect on transport, however, the counter ion and substrate gradients will have a positive effect on the transport rates.

There are four different mechanisms by which electrogenic substrate-cation symport can occur. The first mechanism is where a neutral substrate is in a 1:1 stoichiometry with the counter ion (Fig. 1.3a iv). Transport of the substrate will be positively influenced by the membrane potential and the substrate and counter ion gradients. A number of neutral amino acids were found to use this mechanism in *B. stearothermophilus* (Heyne et al., 1991). The second mechanism is where a cationic substrate is in a 1:1 stoichiometry with the cationic counter ion (Fig. 1.3a v). Transport of the substrate will be affected by the gradients as in the previous example, except the membrane potential will have a more pronounced effect due

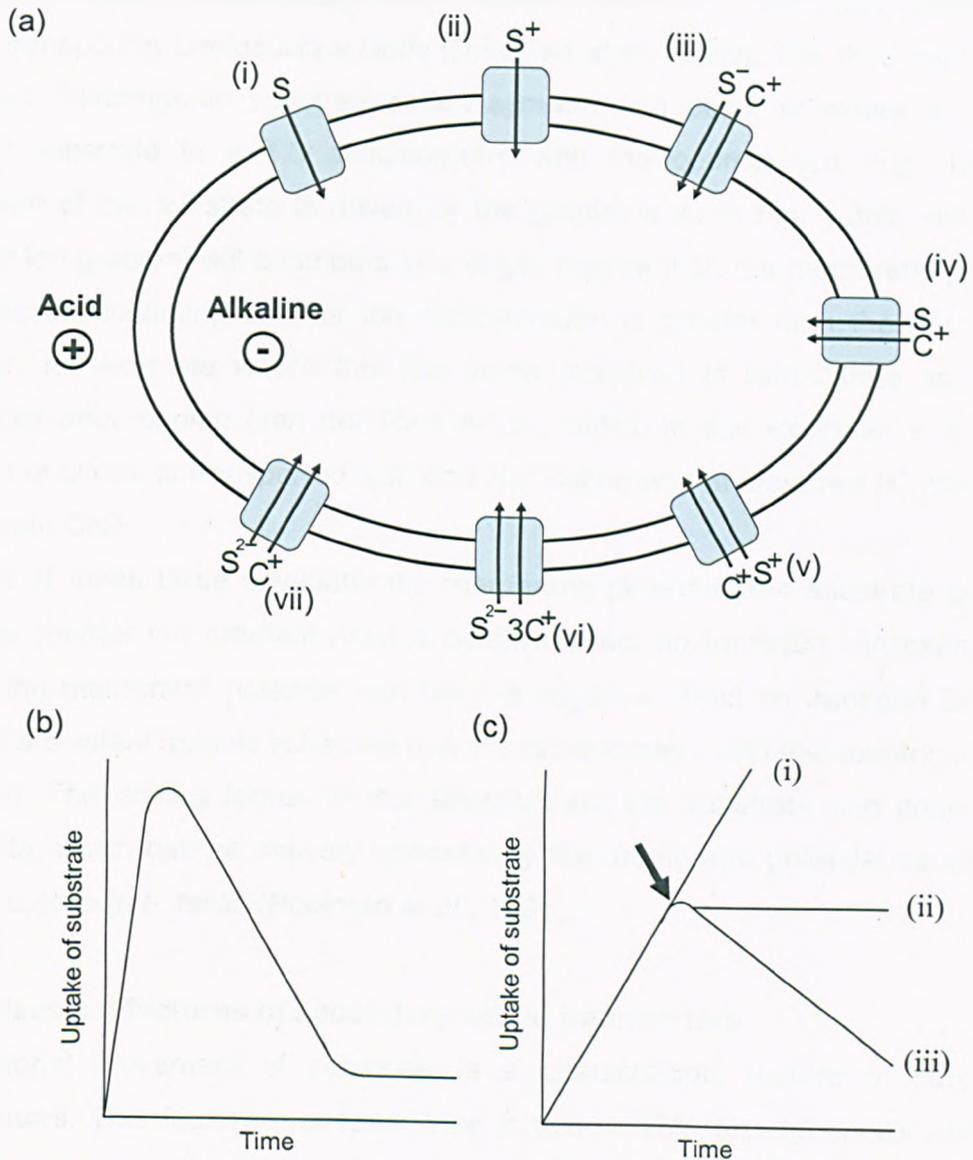


Figure 1.3 Energetic requirements and unique transport characteristics of secondary transport. (a) Examples of secondary symporters showing the relationship between solute (S) and cationic counter ion (C); (i) electroneutral solute uniport, (ii) electrogenic solute uniport, (iii) electroneutral solute-cation symport, (iv-vii) electrogenic solute-cation symport. (b) Schematic of model substrate counterflow data. (c) Schematic of model “cold chase” data. (i) No excess substrate added, (ii) presence of a unidirectional transporter, i.e. an ABC transporter, (iii) presence of a bidirectional transporter, i.e. secondary transporter. Black arrow indicates addition of excess non-radiolabelled substrate.

to the increased overall charge. An example where this has been observed is lysine transport in *Lactococcus lactis* (Driessen *et al.*, 1989). The third mechanism by which electrogenic substrate-cation symport can occur is where a divalent anionic substrate in a 1:3 stoichiometry with the counter ion (Fig. 1.3a vi). Transport of the substrate is driven by the gradients as in Fig. 1.3a ii, except the counter ion gradient will contribute to a larger degree than the membrane potential because the incoming counter ion concentration is greater than the net charge transfer. An example where this has been observed is with citrate uptake by *Klebsiella pneumoniae* (van der Rest *et al.*, 1992). In this example, a dianionic species of citrate is transported with one Na⁺ cation and at least two H⁺ cations by the protein CitS.

In each of these three situations the membrane potential, the substrate gradient and the counter ion gradient have a positive effect on transport. An example of where the membrane potential can have a negative effect on transport is in the case of a divalent anionic substrate in a 1:1 stoichiometry with the counter ion (Fig. 1.3a vii). The driving forces in this situation are the substrate and counter ion gradients, which can be actively opposed by the membrane potential as seen for malate uptake in *L. lactis* (Poolman *et al.*, 1991).

1.5.4 Classical features of secondary active transporters

Bidirectional movement of substrate is a characteristic feature of secondary transporters. This feature manifests itself in two notable experimentally observed phenomena; counterflow and exchange reactions. There are no equivalent phenomena in primary transporters, e.g. ABC transporters.

Substrate counterflow or counter transport was first theorized in 1952 (Widdas, 1952) and then experimentally determined in 1957 in erythrocytes (Rosenberg & Wilbrandt, 1957). It has since been observed using a number of different secondary transporters, including lactose permease (Wong & Wilson, 1970).

In a counterflow experiment, a high concentration of cells (or membrane vesicles or proteoliposomes) containing a secondary transporter are pre-loaded with a high concentration of substrate (Substrate A), for example, 10 mM. The cells are then diluted into buffer containing a low concentration of radiolabelled substrate A (Substrate B), for example, 0.4 mM. This results in an outwardly directed Substrate

A gradient which exchanges with Substrate B resulting in accumulation of Substrate B inside the cell. Eventually, the gradient of Substrate A dissipates because the Substrate A concentration across the membrane reaches equilibrium, which is followed by equilibration of Substrate B across the membrane meaning the counterflow phenomenon is transient (Schematic of "ideal" counterflow result in Fig. 1.3b).

Counterflow occurs because Substrate A and Substrate B compete for the same empty transporter. Inside the cell, where Substrate A is at a much higher concentration than Substrate B, Substrate A will out-compete Substrate B for the transporter more so than it would on the other side of the membrane. Thus there is an apparent net flow of Substrate B in the opposite direction of Substrate A. As soon as the outwardly directed gradient dissipates, Substrate A can once again compete for the empty transporter and reaches equilibrium across the membrane. Counterflow is only theoretically possible if there is a reorientation of the transporter when transitioning from facing one side of the membrane to the other, hence counterflow is not observed in rigid channels and pores. Counterflow is also not observed in ABC uptake systems as conformational changes in the transmembrane domains are coupled to other protein components and ATP hydrolysis and these systems are also inherently unidirectional due to the presence of an extracytoplasmic solute receptor (ESR).

Another characteristic of secondary transport is seen in exchange (or "cold chase") experiments. In a cold chase experiment, transport of radiolabelled substrate is allowed to proceed, leading to accumulation of radiolabel inside the cell (or vesicle or proteoliposome). At an appointed time, excess unlabelled ("cold") substrate is added to the outside of the cell leading to exchange of radiolabelled and unradiolabelled substrate, ultimately leading to apparent efflux of radiolabelled substrate (Fig. 1.3c). An example of this is observed for citrate uptake into *Leuconostoc mesenteroides* (Marty-Teyssset *et al.*, 1996). Excess cold substrate on the outside inhibits further uptake of radiolabelled substrate, but not efflux of radiolabelled substrate leading to a net efflux of radiolabelled substrate from the cell. In transport systems incapable of exchange, like ABC transporters, addition of excess non-radiolabelled substrate prevent further uptake of radiolabelled

substrate (presumably by dilution), but does not lead to efflux from the cell/vesicle/proteoliposome, as seen for the vitamin B₁₂ transporter from *E. coli* (Borths *et al.*, 2005).

1.6 Major facilitator superfamily (MFS)

The largest class of secondary transporters is the major facilitator superfamily (MFS) (Pao *et al.*, 1998). The MFS is a huge family of secondary transporters that have diverse substrate specificities despite sharing structural similarities and conserved sequence motifs (Pao *et al.*, 1998). Members of the MFS are composed of a single polypeptide, the majority of which are predicted to have 12 TMHs, although some are predicted to have 14. Hydropathy profile and sequence analysis revealed 2-fold pseudosymmetry in MFS members indicating that a gene duplication event of a 6 helical bundle resulted in the formation of the 12 TMHs observed. Structural analysis of three distinct MFS members (OxIT, GlpT and LacY) has confirmed this theory (Abramson *et al.*, 2003, Hirai *et al.*, 2003, Huang *et al.*, 2003).

1.6.1 Lactose permease as a model secondary transporter

1.6.1.1 General features

The most comprehensively studied secondary transporter is lactose permease (LacY) from *E. coli*. Studies on this protein have been performed, to some degree, for almost 50 years starting with the work of Jacob and Monod (Jacob & Monod, 1961) right up to modern day with the structural characterisation of the transporter in last decade (Abramson *et al.*, 2003, Guan *et al.*, 2007).

lacY is the second structural gene in the *lacZYA* operon and was the first gene encoding a transporter to be cloned, overexpressed (Teather *et al.*, 1980), sequenced (Buchel *et al.*, 1980), purified and functionally characterised *in vitro* (Newman *et al.*, 1981, Viitanen *et al.*, 1984). LacY is specific for a number of different α - and β -galactosides, is known to transport mono-, di- and tri-saccharides and is a member of the oligosaccharide:H⁺ symporter (OHS) family in the MFS based on sequence identity (Pao *et al.*, 1998).

The protein itself is 417 amino acids long, is ~85% alpha helical when measured using circular dichroism and is also predicted to have 12 TMHs using hydropathy

analysis (Foster *et al.*, 1983). The functional unit was experimentally determined to be a monomer (Guan *et al.*, 2002, Sahin-Toth *et al.*, 1992). All of these experimentally derived observations have been confirmed with the elucidation of high resolution crystal structures (Abramson *et al.*, 2003, Guan *et al.*, 2007).

Lactose permease utilises a downhill proton gradient ($\Delta\mu\text{H}^+$) to drive accumulation of galactosides against their concentration gradient. Conversely, it has been observed that LacY can use a downhill galactoside gradient to accumulate H^+ against its concentration gradient, thus forming a $\Delta\mu\text{H}^+$ (Kaback *et al.*, 2001). A vast number of biochemical studies have been performed to ascertain the mechanism by which the movement of protons through the protein facilitates accumulation of substrate. One of the more powerful and informative techniques that has been used is cysteine-scanning mutagenesis, which can give both mechanistic and structural information (Frillingos *et al.*, 1998, Kaback *et al.*, 2001). Using this technique, all 8 of the native cysteine residues in LacY were mutated and then amino acids 2-402 were individually mutated to cysteine producing 401 different mutants. One of the major findings from this study was that only 6 out of the 417 residues are absolutely required for transport of substrate (Frillingos *et al.*, 1998, Joshi *et al.*, 1989). Each of these essential residues was predicted to have a particular role in the transporter cycle; Glu126 and Arg144 are vital for substrate binding as they have been shown to form a charged pair and Glu269 is involved in the coupling of H^+ translocation to substrate transport (Sahin-Toth *et al.*, 1999, Weinglass *et al.*, 2003). Residues Arg302, His322 and Glu325 are all involved in H^+ translocation because it has been shown that mutations in these residues cannot catalyse $\Delta\mu\text{H}^+$ -mediated accumulation of substrate, but can catalyse exchange and counterflow (Kaback *et al.*, 2001). However, these data could not be explained without a high resolution structure of the transporter.

1.6.1.2 High resolution structure of lactose permease

Two high resolution crystal structures of LacY have been solved. The original structure published in 2003 was of a mutant that could bind substrate but was not functional (Fig. 1.4a (Abramson *et al.*, 2003)). Protein crystallography of LacY was difficult due to its inherent flexibility and so a conformationally constrained mutant was produced (Smirnova & Kaback, 2003). This structure was published

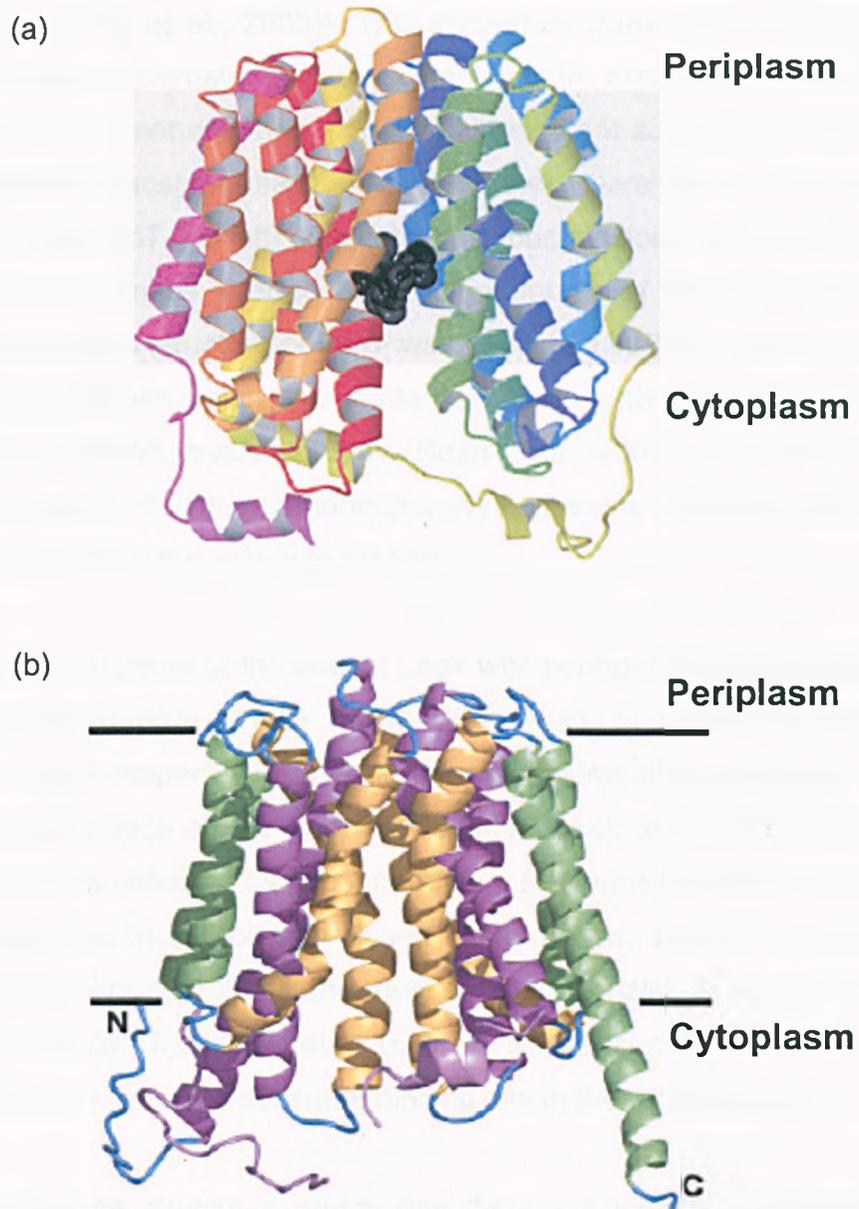


Figure 1.4 High resolution crystal structures of secondary transporters. (a) High resolution crystal structure of lactose permease (modified from Abramson (2003)) showing bound substrate (black). (b) The glycerol-3-phosphate transporter, GlpT (modified from Huang (2003)). The cytoplasm and periplasm are indicated in each case.

concomitantly with a high resolution crystal structure of another MFS member, GlpT (Fig. 1.4b (Huang et al., 2003)). This secondary transporter catalyses the uptake of glycerol-3-phosphate into the cytoplasm in exchange for inorganic phosphate, therefore, it not only has a completely different substrate to LacY, but also has a different transport mechanism. However, there are many parallels between LacY and GlpT, which suggests that observations of structure and mechanism in LacY can be inferred for other members of the major facilitator superfamily. The second structure of LacY was published in 2007 and was solved using the wild-type protein which was made possible by advances in the field of integral membrane protein crystallography (Guan et al., 2007). The major insight taken from this study is that the conformationally constrained mutant LacY is not significantly different from the wild-type protein.

The high resolution structure of the mutant LacY with bound β -D-galactopyranosyl-1-thio- β -D-galactopyranoside (TDG) reveals that LacY is a monomer and is reported to be "heart-shaped" with a large internal, hydrophilic cavity (Fig. 1.4a). This internal cavity is open on the cytoplasmic face, but closed on the periplasmic face, representing the outside-closed conformation. Both structures that have been published of LacY are in this outside-closed conformation, as is the structure of GlpT, possibly indicating that this is the most conformationally stable intermediate in the transport cycle. The substrate, TDG, is in the centre of the molecule indicating that there is only one substrate binding site in the molecule (Fig. 1.4a).

The lactose permease structures reveal that there are indeed 12 α -helices as predicted (Foster et al., 1983), which has 2-fold pseudosymmetry with the formation of 2 helical bundles containing 6 helices each connected by a loop between helices VI and VII. The N- and C-terminal domains are superimposable with a root mean square deviation (RMSD) value of 2.2 Å. These observations suggest that one of the domains in the functional unit emerged through gene duplication followed by divergence. A similar observation has been made for OxlT (Hirai et al., 2003) and for the crystal structure of GlpT (Huang et al., 2003) indicating that this may be a feature of all MFS members.

The substrate binding site is found in the hydrophilic cavity in the middle of the protein/membrane in the vicinity of the axis of pseudosymmetry. The binding site is composed of residues from both the N- and C-terminal domains which make a number of different contacts with the substrate including van der Waals interactions, hydrogen bonding, salt bridges and some possible interactions via water molecules (Fig. 1.5a). The essential residue Arg144 makes direct contacts with the ligand. Contrary to the biochemical data, Glu126 does not appear to interact with either the ligand or Arg144, however, it may be doing so via interactions with water molecules that are not discernable at this resolution (Fig. 1.5a). Analysis of the transport activities of mutations of essential residues revealed that neutral substitutions of Arg302, His322 and Glu325 can catalyse transport of substrate as long as it does not depend on H⁺ translocation (exchange reactions and counterflow (Kaback et al., 2001)). In the crystal structures, these residues are closely packed and form a complex salt bridge/hydrogen bonding network that is ~6 Å away from the substrate binding site (Fig. 1.5b and c). Even with the large amount of biochemical, biophysical and structural data, the exact nature of the communication between H⁺ translocation and substrate translocation cannot be inferred from this structure alone. More work is required to make definite assertions on the exact transport mechanism.

1.6.1.3 Proposed transport mechanism

The proposed mechanism of H⁺ gradient-dependent galactoside transport, and that of other members of the MFS, is termed the “rocker switch” or the alternating access model (Fig. 1.6). In this model, it is proposed that upon binding substrate and counter ion to the outside-open conformation of the protein, a conformational change occurs which closes the outer opening of the internal cavity and opens the inner opening, thus allowing release of the substrate into the cytoplasm. The protein is then thought to reconfigure to the outside-open conformation upon release of the substrate. Unfortunately, there are no crystal structures that support this model, as both crystal structures of LacY and the crystal structure of GlpT are in the inwardly-open conformation. However, a recent study revealed that the binding of substrate to LacY induces conformational changes on both the

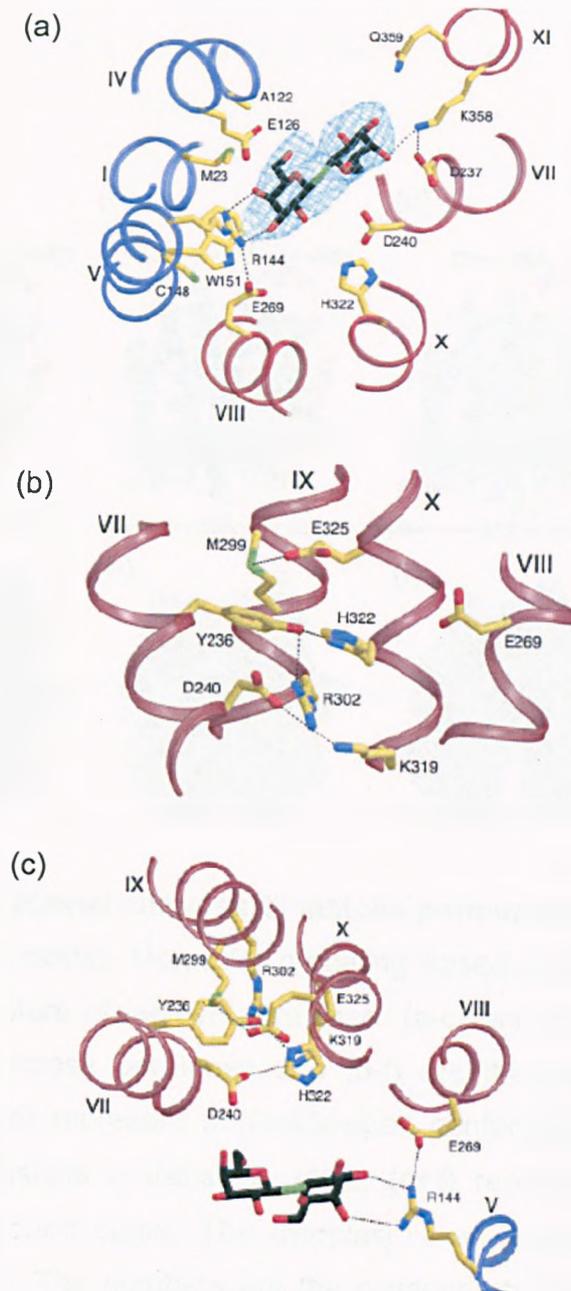


Figure 1.5 Substrate binding site and H⁺ translocation channel of lactose permease. (a) Substrate binding site of lactose permease showing interaction of protein residues with bound substrate, β -D-galactopyranosyl-1-thio- β -D-galactopyranoside (TDG). (b) Residues in lactose permease forming a complex salt bridge/hydrogen bonding network proposed to be forming the H⁺ channel. (c) Close positioning of substrate binding site and H⁺ translocation channel (6 Å) allows communication between the sites of substrate and H⁺ binding/translocation. Figure modified from Abramson (2003).

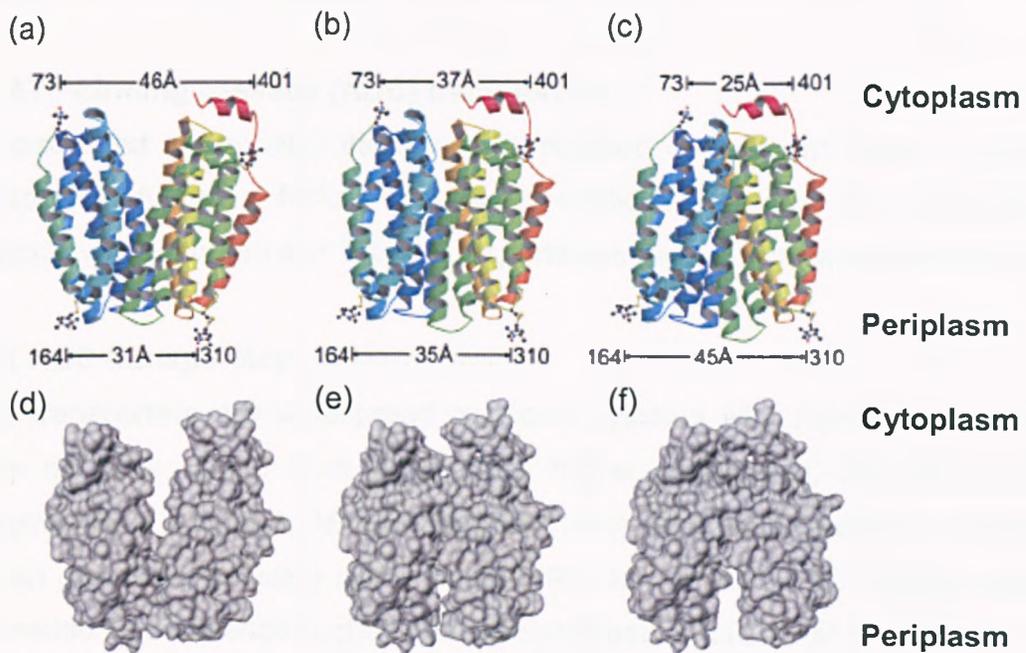


Figure 1.6 Conformational changes of lactose permease during the proposed alternating access model. Molecular modeling based on DEER measurements and the crystal structure of lactose permease. (a-c) are the secondary structure representations of lactose permease and (d-f) are the equivalent space filling representations. **(a+d)** represent an inside-open conformation prior to substrate binding, **(b+e)** represents a transition state, **(c+f)** represents an outside-open conformation with bound sugar. The cytoplasmic and periplasmic faces of the protein are indicated. The numbers are the distances between labelled residues measured by DEER. Figure modified from Smirnova (2007).

cytoplasmic and periplasmic sides of the internal cavity (Smirnova *et al.*, 2007). This study was performed using double electron-electron resonance (DEER) to measure the distances between specifically labelled helices allowing a model to be constructed which is in agreement with the alternating access model.

1.7 ATP-binding cassette (ABC) transporters

As discussed previously, there are a number of different types of primary transporters found in Nature. The most applicable system for comparison to secondary transporters and TRAP transporters is the ABC transporter superfamily.

1.7.1 ABC transporters: an introduction

ABC transporters are widespread transport systems with representatives in all three domains of life from bacteria to higher eukaryotes, including humans (Higgins, 2003). Indeed, ABC transporters have been implicated in a number of human diseases, including cystic fibrosis (Riordan *et al.*, 1989), Tangiers disease and multidrug resistance in chemotherapeutic treatment (Cole *et al.*, 1992).

ABC transporters perform active transport on a number of diverse substrates resulting in the formation of a high substrate gradient (up to $10^6:1$ (Dippel & Boos, 2005)). ABC transporters are primary transporters because it has been found that they couple the binding/hydrolysis of ATP forming ADP and P_i to the transport of the substrate across the membrane (Dean *et al.*, 1989). Although the composition of ABC transporter components is quite variable, they can be generally divided into two categories; ABC importers and ABC exporters. ABC importers catalyse the unidirectional uptake of a wide range of substrates including sugars (maltose (Dean *et al.*, 1989)), a number of amino acids (histidine (Yao *et al.*, 1994), polypeptides (Tame *et al.*, 1994), vitamins (vitamin B₁₂ (Borths *et al.*, 2005)) and metal-chelate complexes (Pinkett *et al.*, 2007). ABC exporters catalyse the extrusion of virulence factors, for example, Haemolysin (Dinh *et al.*, 1994) and hydrophobic compounds, the majority of which have anti-microbial action. The ABC importers are the most pertinent group to discuss in detail as they are most comparable to TRAP transporter, therefore, ABC exporters will not be described in any further detail.

The core structure of an ABC transporter contains two components; a hydrophobic integral membrane domain that forms the translocation channel, otherwise known as the transmembrane domain (TMD) and a soluble, cytoplasmic nucleotide binding domain (NBD, otherwise known as ATPase domain or the ABC protein) that binds and hydrolyses ATP. In the functional transporter, the TMD and the NBD are both dimeric and have been found in a number of configurations. Often the four domains are separate proteins. However, they have also been found to have one TMD and one ATPase domain fused into one polypeptide and the functional transporter composed of two of these so-called 'half transporters' (Dawson *et al.*, 2007).

A third structural component of ABC importers is an extracytoplasmic solute receptor (ESR, otherwise known as a solute binding protein or periplasmic binding protein). This component is found in the periplasm of Gram negative bacteria and attached to the cytoplasmic membrane by a lipid moiety in Gram positive bacteria (Gilson *et al.*, 1988). The ESR binds the substrate of the transporter with a high affinity (low μM) and presents it to the translocation machinery. The ESR is thought to confer the specificity to the transporter, the high affinity, and also unidirectionality (Doeven *et al.*, 2004, Wilkinson, 2003).

1.7.2 A short history of ABC transporters functional determination

There was a great deal of debate and controversy in the transport field when ABC transporters were first discovered, and it took almost three decades before the full picture emerged regarding the structural components and the energy utilisation of these transporters.

In the late 1970s and early 1980s transporters were originally separated on the basis of being either "shock sensitive" or "shock insensitive" permeases. This refers to the sensitivity of the transporter when the cell undergoes osmotic shock treatment (Neu & Heppel, 1965). This method compromises the integrity of the outer membrane using the metal-chelating agent, ethylenediaminetetraacetic acid (EDTA) and lysozyme while preventing cell lysis by maintaining the cells in high sucrose solution. In this way the periplasm is removed leaving sphaeroplasts. At the time, "shock insensitive" transporters were essentially all secondary

transporters as they did not require any periplasmic components to facilitate transport. The “shock sensitive” transporters were what we now refer to as ABC importers that require a periplasmic ESR for function (ignoring for a moment the other ESR-dependent transporter that will be mentioned later). These initial classifications were obviously limited by all the studied organisms being Gram negative bacteria. The classifications have since been refined with the discovery of variations, for example, anchored ESRs, ABC exporters and other ESR-dependent transporters.

The term “ABC transporter” was not coined until 1990 (Hyde *et al.*, 1990) and prior to this there was a great deal of ambiguity as to the energy coupling of substrate transport via ABC transporters. Some examples that led to this ambiguity and confusion include the *E. coli* glutamine ABC importer (now known as GlnHPQ), which was thought to require both ATP and the membrane potential in the transport mechanism (Plate, 1979). Another example of literature that produced some ambiguity was that maltose uptake in *E. coli* (by MalEFGK₂) was found not to require the PMF, but was also reported to be unaffected by depletion of intracellular ATP concentration (Ferenci *et al.*, 1977). A particularly curious example is that transport of glutamine, histidine and methionine in *E. coli* was reported to be independent of the PMF, but dependent on acetyl phosphate (Hong *et al.*, 1979). These experiments were performed in whole cells, which are complex environments where it is particularly difficult to determine accurate transport energetics (Poolman, 1994). It was also found that some of the inhibitors used in these experiments produced artifacts that would produce misleading data (Joshi *et al.*, 1989). It was only with the use of a simplified vesicle system that the energy coupling of ABC transporters was resolved (Ames *et al.*, 1989, Dean *et al.*, 1989).

1.7.3 *E. coli* maltose transport by MalEFGK₂ as a model ABC system

The maltose ABC transporter in *E. coli* is arguably the most comprehensively studied ABC transporter in Nature. It is a member of the carbohydrate uptake transporter (CUT) family as defined by Saier (Saier, 2000) and has been shown to transport di- and oligo-saccharides. Homologues of the *E. coli* maltose ABC transporter are present in a number of bacteria dwelling in various environments. Bacteria obtain maltose from the amylase-mediated degradation of starch in the

environment. In the case of soil and marine bacteria this is derived from plant matter and in the case of gut-dwelling organisms like *E. coli*, it is obtained from the human diet. The maltose ABC transporter can also transport maltodextrins, which are linear polymer of D-glucose molecules joined at the O-glycosidic bond, up to maltoheptaose (7 glucose molecules) (Wandersman *et al.*, 1979).

The maltose uptake system in *E. coli* is composed of an ESR, namely MalE (or MBP), a heterodimer of TMDs, MalF and MalG and a homodimer of the ATPase domain, MalK.

1.7.3.1 The ESR domain – MalE

MalE is a soluble, ~40 kDa protein that is found in the periplasm of Gram negative bacteria (Kellermann & Szmelcman, 1974) or anchored to the cytoplasmic membrane as a lipoprotein in Gram positive bacteria (Gilson *et al.*, 1988). MalE, like most periplasmic proteins, is synthesized in the cytoplasm and then translocated to the periplasm by the Sec translocon that recognizes an N-terminal signal peptide. The signal peptide is cleaved off during the translocation process producing the mature protein.

MalE has been purified and characterised *in vitro* and found to bind maltose and maltodextrin with a high affinity (K_D of 1 μ M (Kellermann & Szmelcman, 1974)). Production of MalE is induced by maltose to high degrees and is present in the periplasm at high concentration (~1 mM (Dietzel *et al.*, 1978)). The total number of molecules of each component of the MalEFGK₂ ABC transporter have been calculated (Dietzel *et al.*, 1978, Koman *et al.*, 1979, Shuman & Silhavy, 1981) and it has been determined that MalE is in ~40-50x excess over the other components. However, a study has revealed that only ~20% of the MalE predicted to be present in the periplasm is required for normal transport rates (Manson *et al.*, 1985). In this study, the MalE signal peptide was modified to modulate the total amount of MalE translocated to the periplasm. The levels were varied between 4-23% of normal MalE levels and transport rates of maltose were then determined, revealing that the MalE concentration is in great excess of what is required for transport. It is unknown why such a high concentration is present, however, MalE is also known to have been recruited for a use as a chemotaxis receptor (Springer *et al.*, 1977). Maltose bound to MalE interacts with a chemotactic signal transducer, Tar (taxis to

aspartate and some repellents), which generates a chemotactic response to maltose. *E. coli* requires approximately double the concentration MalE for wildtype chemotactic response compared to wildtype uptake rates (Manson et al., 1985), which may help in explaining the overabundance of MalE.

Studies using *in vitro* vesicle systems have shown that MalE stimulates ATPase activity and is locked onto the TMDs when an ATP transition state analogue is bound to the ATPase domain (Chen et al., 2001, Davidson et al., 1992). These two observations suggest that there is communication between the ESR and the ATPase domains, presumably mediated by conformational changes in the TMDs.

The high resolution structure of MalE has been determined in both the liganded and unliganded states (Sharff et al., 1992, Spurlino et al., 1991). These structures revealed that MalE is composed of two globular domains connected by a flexible hinge region (Fig. 1.7a). The ligand binding site is in the middle of the globular domains which is freely accessible to the solvent prior to binding the ligand. Upon binding ligand, the protein undergoes a large conformational change which results in complete occlusion of the ligand from the bulk solvent (Fig. 1.7b). Residues from both globular domains contribute to the formation of the binding site and make a number of van der Waals interactions, two hydrogen bonding interactions and possibly more interactions that can not be seen at the 2.3 Å resolution obtained.

The overall features of MalE as determined by the high resolution crystal structures are also found in other ABC transporter binding protein, including the leucine/isoleucine/valine binding protein (Olah et al., 1993, Trakhanov et al., 2005), the D-galactose/D-glucose binding protein (Vyas et al., 1991), the L-arabinose binding protein, and the vitamin B₁₂ specific protein (Borths et al., 2002). What is remarkable about this observation is that these proteins share little sequence identity and yet have the same overall structural features and mode of action. The major difference found in the structures of ESRs is in the folding topology and how the two globular domains are connected. These are divided into two groups; type I and type II ESRs (Fig 1.7c).

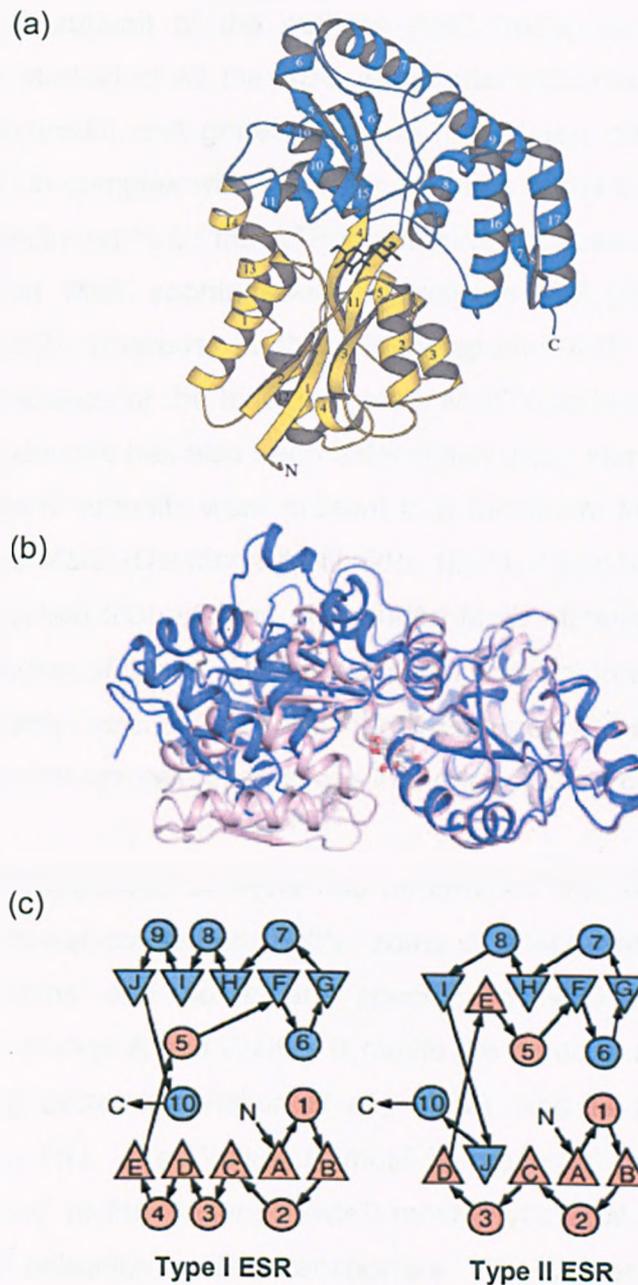


Figure 1.7 High resolution crystal structure of MalE and topological differences between type I and type II ESRs. (a) High resolution structure of MalE with bound substrate from hyperthermophile *T. litoralis* (Diez *et al.*, 2001). (b) Crystal structures of the conformational changes of MalE from *E. coli* in the presence (purple) and absence (blue) of bound substrate. Figure taken from Davidson (2008). (c) Difference in the core structure of type I and type II ESRs caused by domain dislocation. Taken from Fukami-Kobayashi (1999).

1.7.3.2 ATPase domain (NBD) - MalK

MalK is the NBD subunit of the maltose ABC transporters and is the most comprehensively studied of all the ABC transporter subunits. A large number of biochemical, biophysical and genetic studies have been performed on MalK in isolation and also in complex with the other subunits of the transporter. MalK was experimentally confirmed to be the ATPase subunit because when expressed and purified in isolation, MalK spontaneously hydrolyses ATP (Morbach *et al.*, 1993, Walter *et al.*, 1992), whereas, in the full transporter ATP hydrolysis is tightly coupled to the presence of the binding protein MalE (Davidson *et al.*, 1992). The stoichiometry of subunits has also been determined using immunoprecipitation and found that two MalK subunits were present in a functional transporter along with one MalF and one MalG (Davidson & Nikaido, 1991). It has been demonstrated in a reconstituted system that addition of liganded MalE stimulates ATPase activity, indicating that binding of liganded MalE transmits a signal through MalF and MalG to MalK presumably through conformational changes. This signalling process depends on the protein:protein interactions in the transporter complex.

Mutational and phylogenetic analysis has determined that MalK, and indeed all NBDs, contain several conserved motifs, some of which are characteristic of all ATP-binding proteins and some are specific to NBD subunits from ABC transporters. The Walker A and Walker B motifs are found in a number of different nucleotide-binding proteins (Walker *et al.*, 1982), and is proposed to be the nucleotide-binding fold. The Walker B motif is preceded by LSGGQ (or "ABC signature sequence" or the "linker peptide") motif (Hyde *et al.*, 1990). This motif is found in all NBD subunits in ABC transporters. The Walker A and B motifs are separated from the LSGGQ motif by a sequence containing a conserved glutamine and called the Q loop (otherwise known as the "helical domain" or the "lid"). The proposed functions of these motifs and folds have been deduced from the observed phenotypes produced upon mutation of residues within the motifs (extensively reviewed in Hunke (2000) and Schneider (2003)). This mutational analysis has shown that any variants of the Walker A or Walker B motifs prevent ATP binding and/or hydrolysis and therefore also inhibit transport. The Q loop has

been implicated in signaling between the NBD and the TMDs because mutations in this motif abolished transport, but left nucleotide binding unaffected. The postulate that the Q loop facilitates the interaction of the NBD with the TMD is further strengthened by the production of suppressor mutations in MalK that restore transport activity in response to mutation of a conserved region of the TMD (Hunke et al., 2000, Mourez et al., 1997). Mutations in the LSGGQ motif have been found to abolish transport, but leave nucleotide binding unaffected indicating that this motif is also involved in signaling between the NBD and TMD.

To determine the mechanism of NBDs in ABC transporters it was necessary to elucidate the high resolution structure of an NBD in complex with ATP, ADP and Pi and nucleotide free. There are a number of structures of isolated NBDs from different transporters (Diederichs et al., 2000, Gaudet & Wiley, 2001, Hung et al., 1998, Karpowich et al., 2001, Yuan et al., 2001). There are also a number of structures for complete ABC complexes determined (Hollenstein et al., 2007a).

The first structure of a MalK homologue was from the archaeon *Thermococcus litoralis*. This structure was determined to be non-physiological due the small dimer interface and peculiar arrangement of sub-domains (Diederichs et al., 2000). More recent structures of MalK in isolation (Chen et al., 2003, Lu et al., 2005) and in complex with the rest of the MalEFGK₂ complex (Oldham et al., 2007) from *E. coli* have been solved that are more consistent with the biochemical data for MalK.

The structure of MalK has revealed that it is composed of two domains, an ATPase domain and a regulatory domain (RD, Fig. 1.8a). MalK from *E. coli* is not a typical ABC transporter NBD for two reasons. Firstly, it contains an additional C-terminal domain present in MalK homologues, but not in other NBDs (Diederichs et al., 2000, Oldham et al., 2007) and secondly, MalK can be purified as a dimer in the presence and absence of nucleotide – a feature not shared with other NBDs (Chen et al., 2003, Diederichs et al., 2000, Kennedy & Traxler, 1999). These characteristics are thought to be linked as the RDs from each protomer are thought to stabilize the dimer in the absence of nucleotide (Chen et al., 2003). MalK from *E. coli* has been crystallised in three different conformations; a closed, ATP bound form, and two open forms in the presence of ADP and without bound nucleotide

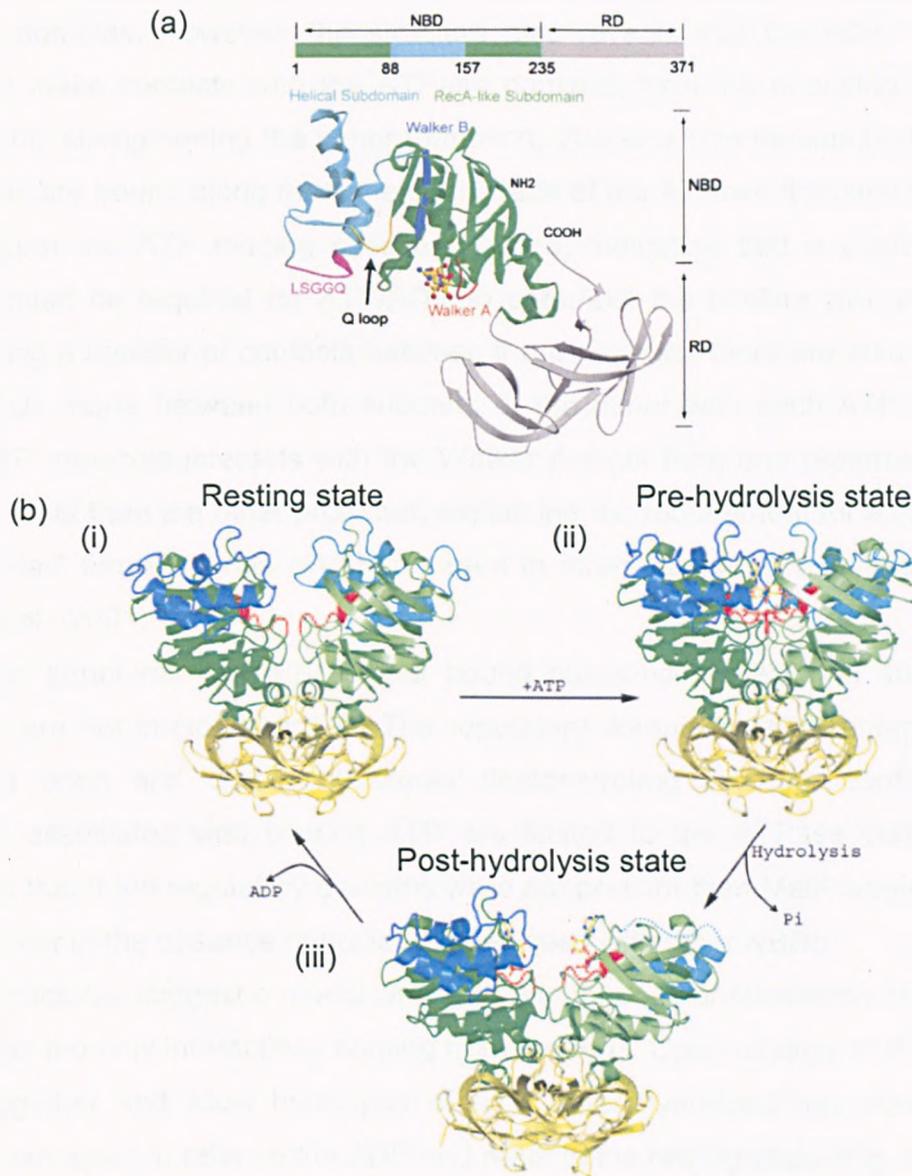


Figure 1.8 Crystal structure and conformational changes of the ATPase domain, MalK. (a) Crystal structure of a MalK monomer showing the nucleotide binding domain (NBD), the regulatory domain (RD), LSGGQ motif, Q loop and WalkerA/B motifs. Substrate is shown interacting with the Walker A motif. Figure taken from Chen *et al* (2003). (b) Conformational changes during binding and hydrolysis of ATP. (i) In the absence of nucleotide, the dimer is in the resting state and primarily interacting through the RD. (ii) In the presence of nucleotide, the NBDs come together burying ATP and allowing hydrolysis to occur. (iii) The ADP-bound state after hydrolysis is in the open state allowing release of nucleotide and resetting back to resting state. Figure modified from Lu *et al* (2005).

buried surface area. The contacts between the protomers principally come from the ATPase domains. However, the structure also reveals that the RDs from each protomer make contacts with the ATPase domains from the opposing protomer, presumably strengthening the dimer formation. The structure reveals that two ATP molecules are bound along the dimeric interface of the ATPase domains and in the closed form the ATP binding sites are buried, indicating that a conformational change must be required for ATP/ADP to enter/exit the binding site. As well as there being a number of contacts between the protomers, there are also a number of contacts made between both subunits of the dimer with each ATP molecule. Each ATP molecule interacts with the Walker A motif from one protomer and the LSGGQ motif from the other protomer, explaining the requirement for a dimer. This "head-to-tail" formation has also been seen in other NBDs, for example, MJ0796 (Yuan et al., 2001).

The open structures of MalK without bound nucleotide reveal that the ATPase domains are not in close contact. The regulatory domains are superimposable in both the open and closed structures demonstrating that the conformational changes associated with binding ATP are limited to the ATPase domain. This suggests that if the regulatory domains were not present then MalK would not be a stable dimer in the absence of nucleotide, as seen with other NBDs.

These structures suggest a model where, in the absence of nucleotide, the dimer is open, with the only interactions coming from the RDs. Upon binding ATP the NBDs come together and allow hydrolysis of ATP. After hydrolysis has occurred, the NBDs open again to release the ADP and reset to the resting state (Fig. 1.8b).

1.7.3.3 Transmembrane domains (TMD) - MalFG

The transmembrane component of the maltose ABC transporter is composed of MalF and MalG, two integral membrane proteins of unequal length. The topological models of both proteins have been experimentally determined using alkaline phosphatase fusions (Boyd *et al.*, 1993, Ehrmann *et al.*, 1990). MalG is predicted to be 32 kDa and have 6 TMHs and MalF is 57 kDa and has 8 TMHs.

Phylogenetic analysis of the amino acid sequence of MalG revealed a conserved motif in a hydrophilic, cytoplasmic loop that was also present in MalF and other ABC transporter TMDs (Dassa & Hofnung, 1985). The consensus sequence of this

motif is EAA₃GX₉IXLP (where X is any amino acid), but is referred to as the EAA motif. Due to the conserved nature of the EAA motif it was theorised to be essential for functional transport and, because of its cytoplasmic location, was thought to interact with the ATPase domain. This theory was confirmed using site-specific chemical crosslinking (Hunke et al., 2000) and by mutagenising the most highly conserved residues in the EAA motif in MalF and/or MalG (Mourez et al., 1997). This latter study revealed that mutation of the highly conserved residues leads to decreased or complete abolition of transport activity. In the mutants where transport was abolished, MalK was no longer membrane associated and was localized in the cytoplasm indicating that the EAA motif forms an interaction with MalK. This evidence was further strengthened by the isolation of suppressor mutations in MalK that regained functional activity with some of the more deleterious MalFG mutations. The suppressor mutations isolated in MalK were localized to the Q-loop motif, including residues Ala85 and Val117. These residues were also found to interact with residues in the EAA motif in the chemical crosslinking study (Hunke et al., 2000). This mutagenesis study of the EAA motif also revealed that the EAA motifs of MalF and MalG do not act symmetrically and do not contribute to the definition of the substrate specificity. The interaction of the EAA motif with MalK has also been demonstrated in a more recent study monitoring the binding of synthetic peptides from a library derived from MalFG to MalK, and *vice versa* (Bluschke et al., 2007).

MalF has been extensively mutagenised throughout helices 3-8, which has revealed that substrate differentiation between maltose and maltodextrins was defined by TM5 and 6 (Ehrle et al., 1996, Steinke et al., 2001). This indicates that there is a substrate binding site in the transmembrane domain of the maltose ABC transporter and could also be true for other ABC systems. This fact is not necessarily obvious when ligand specificity and high affinity is conferred by the ESR. Crosslinking studies have revealed that MalE interacts with MalF and MalG in periplasmic loops 1 and 2 (Daus et al., 2007).

It has been established that a wild type maltose ABC transporter has an absolute requirement for the binding protein MalE in order to transport substrate (Shuman, 1982). However, mutants have been isolated that can transport substrate in the

absence of MalE and the mutations responsible have been localised to MalF and MalG (Covitz *et al.*, 1994, Treptow & Shuman, 1985). In these mutants there is a much lower affinity for maltose (K_m of 2 mM as opposed to 1 μ M in the wildtype transporter), the ATPase activity of the MalK dimer is higher (Davidson *et al.*, 1992) and the addition of wildtype MalE actually inhibits transport by these mutants (Treptow & Shuman, 1985). The mutations were mapped to transmembrane helices 5-8 in MalF and 3 and 4 in MalG. The mutations are proposed to alter the character of the transmembrane helices such that the transporter is in the “MalE-triggered” state that has previously been shown to stimulate the ATPase activity of MalK₂. The clustering of the mutations in helices 5-8 of MalF correspond with the helices that were strongly implicated in forming the binding site/translocation channel in the mutagenesis studies (Ehrle *et al.*, 1996).

1.7.3.4 High resolution structure of complete MalEFGK₂ complex

There have recently been a number of high resolution crystal structures of intact ABC transporter complexes, including the binding protein-less Sav1866 multidrug exporter from *Staphylococcus aureus* (Dawson & Locher, 2006), BtuCD vitamin B₁₂ importer from *E. coli* (without its cognate binding protein BtuF (Locher *et al.*, 2002)), HI1470/1 metal-chelate importer from *H. influenzae* (Pinkett *et al.*, 2007) and ModB₂C₂A, the molybdate importer from *Archeoglobus fulgidus* (Hollenstein *et al.*, 2007b). The most relevant structure to this introduction is the elucidation of the intact MalEFGK₂ complex from *E. coli* (Fig. 1.9 (Oldham *et al.*, 2007)).

The structure of MalEFGK₂ was solved to a resolution of 2.8 Å. To obtain the intact structure it was necessary to introduce a single mutation into MalK (E159Q) that permitted binding of ATP but prevented its hydrolysis. An overview of the structure reveals that the MalK dimer is essentially identical to the crystal structure previously obtained (Chen *et al.*, 2003). It is a closed dimer with two ATP molecules bound. The only difference between this structure and the structure of MalK₂ crystallised on its own is that the Q loops are ordered into β -strand structures. In the structure of MalK₂ alone, this region was disordered, but in the structure of the complete transporter the Q loops interact with the TMDs and form an ordered structure. MalF and MalG form the translocation channel and reveals that the topology modelling of these proteins was accurate. There is

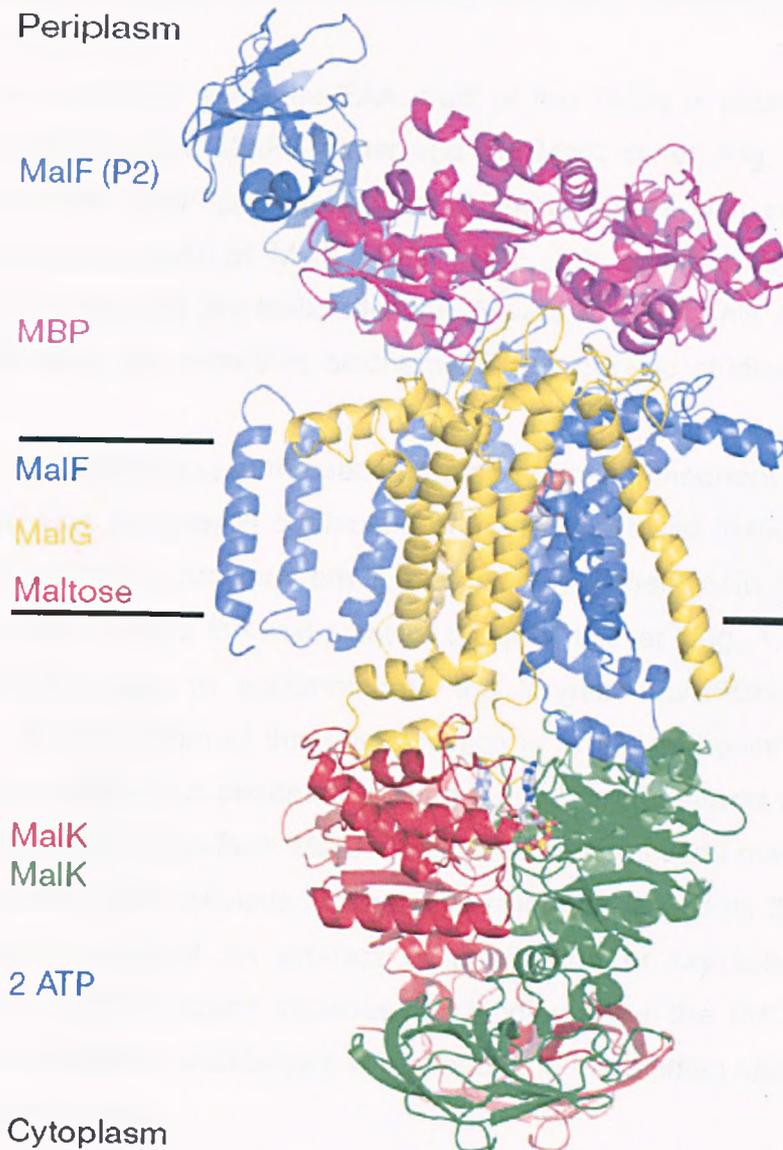


Figure 1.9 Crystal structure of the intact MalEFGK₂ complex. Overall complex of the maltose ABC transporter. Each polypeptide is colour-coded and labelled accordingly on the left hand side. The structure is of a catalytic intermediate able to bind, but not hydrolyse ATP. In this structure, the ATPase domains are closed and bound to ATP. The ESR, MalE (MBP) is open and maltose is bound in the translocation channel of the TMD. MalF (P2) is the Ig-like domain of the integral membrane component that interacts with MalE. Figure taken from Oldham *et al* (2007).

pseudosymmetry between helices 3-8 of MalF and 1-6 of MalG. The ESR, MalE is ligand-free with both globular domains forming extensive interactions with both MalF and MalG (Fig. 1.9).

As predicted from previous work, the EAA motif of the TMDs is prominent in the interaction face between the MalFG dimer and the MalK dimer (Fig. 1.10a). The EAA motifs from both MalF and MalG form 2 small helices with one of these helices docking into a cleft of MalK. This MalK cleft consists of the helical subdomain, the Q loop and the helix after the Walker A motif. This supports the results obtained from the extensive biochemical and genetic studies previously performed.

The substrate binding pathway is the second region forming important interactions in this structure. This pathway is composed of MalF, MalG and MalE and at the intersection of these three proteins there is a large solvent-filled cavity that extends from the periplasmic side to the mid-point of the lipid bilayer (Fig. 1.10b i). This cavity is sufficiently large to accommodate the longest MalEFGK₂ substrate, maltoheptaose. At the bottom of this cavity, which is at the mid-point of the lipid bilayer, a maltose molecule is observed making a number of contacts with several residues from MalF, but none from MalG. The specific interactions made between protein and substrate, and previous biochemical evidence, suggests that this is a genuine interaction and not an artefact of purification or crystallization (Fig. 1.10b ii). This is the first structural evidence of a binding site in the TMD of an ABC transporter, and the overall architecture is analogous to the binding site at the mid-point of the bilayer in LacY.

At the other end of this cavity, MalE is in its open conformation, cupping the top of the translocation channel. There is no maltose in the binding cleft of MalE, therefore binding to MalFG presumably stabilizes the open conformation allowing release of maltose in to the MalFG binding pocket. One of the most striking features from this structure is the presence a periplasmic loop from MalG in the binding cleft of MalE (Fig. 1.10b iii). This so-called "scoop loop" either dislodges maltose from MalE and/or sterically hinders its re-entry therefore facilitating unidirectional transport. Another interesting and unexpected protein:protein interaction is the formation of an Ig-like domain from second periplasmic loop of MalF that makes extensive contacts with the N-terminal globular domain of MalE.

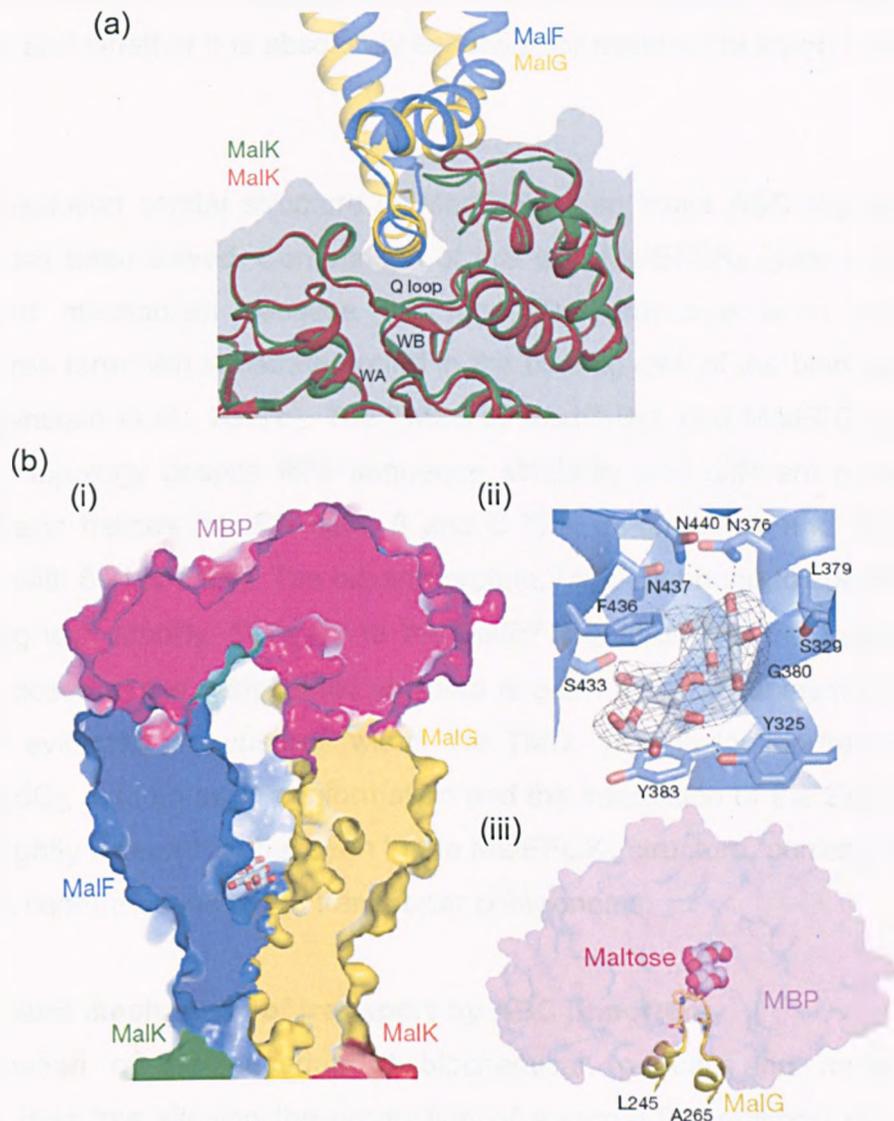


Figure 1.10 Domain interactions in the MalEFGK₂ complex. (a) The conserved EAA loop of MalF and MalG docking with a surface cleft in MalK. Position of the EAA motif from MalF (blue) and MalG (yellow) are compared by superimposing the MalK monomers (red and green). The Q loop, Walker A (WA) and Walker B (WB) are shown. (b) (i) Cross-section of a space-filling model showing the open conformation of MalE (MBP), the solvent filled cavity and the positioning of maltose at the bottom. (ii) Maltose binding site within the TMD showing interactions between maltose and amino acids. (iii) Insertion of the “scoop loop” from MalG into the binding pocket of MalE. A maltose molecule has been modelled in to demonstrate the resultant steric hindrance. Figures taken from Oldham *et al* (2007).

This Ig-like domain is not seen in any of the other structures of intact ABC transporters and whether it is absolutely essential for maltose transport is currently unknown.

The high resolution crystal structure of ModC₂B₂A, an intact ABC importer from *A. fulgidis*, has been solved. Comparison of this with MalEFGK₂ gives insight into the transport mechanism because the ModC₂B₂A structure is in the open, nucleotide-free form with substrate bound in the binding cleft of the binding protein ModA (Hollenstein et al., 2007b). The TMDs of ModC₂B₂A and MalEFGK₂ have a very similar topology despite little sequence similarity and different numbers of transmembrane helices (MalFG have 8 and 6 TMs, respectively and ModB is a homodimer with 6 TMs each). The binding protein, ModA is bound to substrate and is interacting with ModB₂. Contrary to the MalEFGK₂ structure, the translocation channel is closed at the periplasmic end and is open at the cytoplasmic end and there is no evidence of substrate within the TMD. The nucleotide-free ATPase domain, ModC₂, is in an open conformation and the interaction of the EAA motif of ModB₂ is slightly different to that seen in the MalEFGK₂ structure, corresponding to the different conformations of the transporter components.

1.7.4 Proposed mechanism of transport by ABC importers

The culmination of the decades of biochemical, genetic and more recent biophysical data has allowed the production of a model for transport by an ABC importer (Fig. 1.11). The protein complex of ATPase domain and TMDs starts in the resting state in which the translocation channel is closed to the periplasm and the ATPase domain is in the open, unliganded form. Substrate-bound ESR docks with the TMDs (this is observed in the ModC₂B₂A structure) which promotes the conformational change to the catalytic transition state (as seen in the MalEFGK₂ structure) in which the periplasmic side of the TMD is open, stabilizing the open state of the ESR and release of substrate into the translocation channel (with help from the “scoop loop”, where available). ATP hydrolysis takes place, substrate-free binding protein is released, the ATPase dimer opens and the TMDs revert to the resting state. The transport cycle is then ready to begin again.

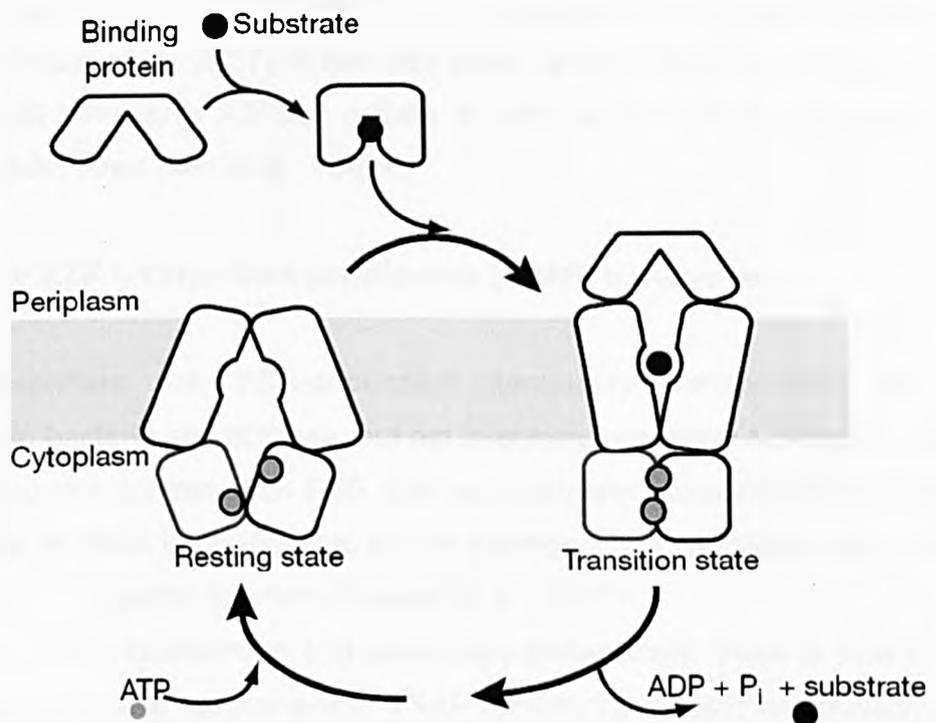


Figure 1.11 Proposed mechanism for solute transport by ABC uptake system. Substrate binds to the ESR and interacts with the TMD-NBD complex in its resting state. Interaction of liganded ESR triggers conformational changes within the TMD resulting in the opening of the periplasmic side of the TMD and release of substrate from the ESR. The NBD closes and hydrolysis of ATP occurs, leading to the opening of the NBD and subsequent opening of the cytoplasmic side of the TMD, releasing the substrate into the cytoplasm. ADP + P_i are released, resulting in the resetting to the resting state. Figure taken from Oldham *et al* (2007).

It is worth reiterating that the model suggests that it is the interaction of the liganded ESR with the TMDs and not binding of ATP that initiates the opening of the periplasmic side of the translocation channel and closing of the ATPase domain. ATP concentrations within the cytoplasm have been reported to be sufficiently high so that the ATPase domain is saturated at all times (Davidson *et al.*, 1996, Oldham *et al.*, 2007). It has also been demonstrated that introduction of liganded MalE stimulates ATPase activity *in vitro*, which further strengthens the proposed model (Davidson *et al.*, 1992).

1.8 Tripartite ATP-independent periplasmic (TRAP) transporters

TRAP transporters are ESR-dependent secondary transporters that are widespread in bacteria and archaea, but not in eukaryotes (Kelly & Thomas, 2001). TRAP transporters consist of an ESR and two unequally sized integral membrane proteins, one of which is predicted to be the translocation channel and the other is of unknown, but essential function (Forward *et al.*, 1997).

Compared to ABC transporters and secondary transporters, there is little known about the structure and mechanism of TRAP transporters. Some comparisons can be drawn between TRAP, classical secondary and ABC transporters. TRAP transporters utilize an electrochemical gradient to power transport, therefore they are secondary transporters. This is combined with the presence of a high affinity ESR, characteristic of ABC transporters (Kelly & Thomas, 2001) (Fig. 1.12).

The first TRAP transporter to be identified was the C₄-dicarboxylate (malate, succinate and fumarate)-specific DctPQM from *Rhodobacter capsulatus* (Forward *et al.*, 1997).

1.8.1 Discovery of the first TRAP transporter: DctPQM from *R. capsulatus*

1.8.1.1 C₄-dicarboxylate transport in bacteria

C₄-dicarboxylates (succinate, fumarate, malate) are transported into bacterial cells and metabolized in both aerobic and anaerobic conditions. A number of different variations of secondary transporters are known to be used to transport C₄-dicarboxylates into and out of the cell. Based on sequence similarity, structural components and energetic requirements these transporters can be split into a

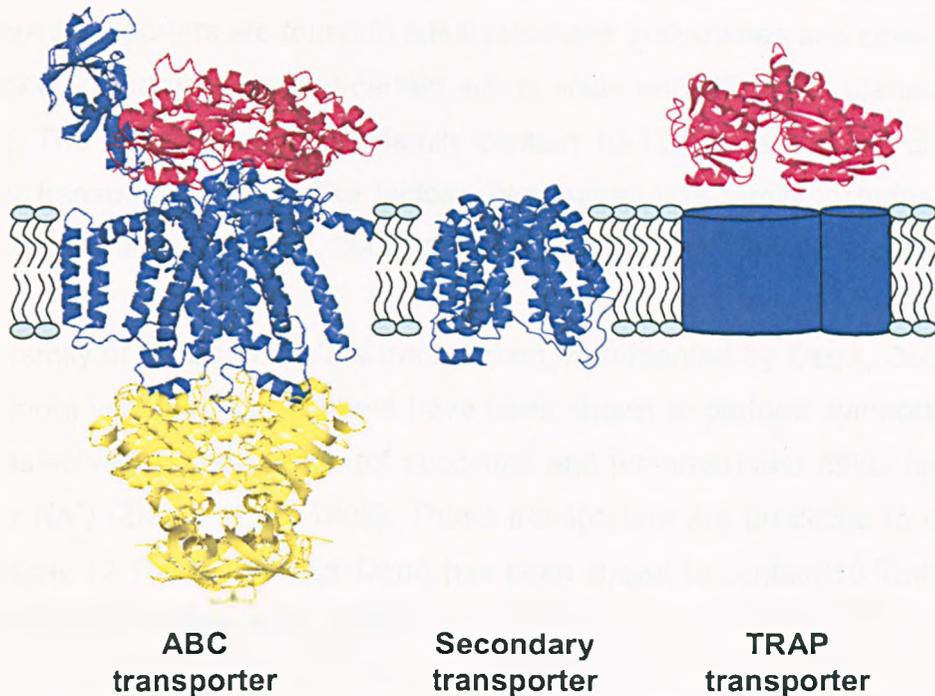


Figure 1.12 Composition of archetypal ABC, secondary and TRAP transporters. Comparison of the structural composition of three types of transporter. ABC transporters, represented by MalEFGK₂ (Oldham et al., 2007) are composed of an ESR, a transmembrane domain and a nucleotide binding domain. Secondary transporters, represented by lactose permease (Abramson et al., 2003) are composed of a transmembrane domain with no need of ancillary protein components. TRAP transporters are secondary transporters composed of a transmembrane domain and an ESR component. The ESR shown is SiaP (Muller et al., 2006). There is no structural information regarding TRAP transporter integral membrane components. Colour coding: ESR domains are red, transmembrane domains are blue and nucleotide binding domains are yellow.

number of families. The first group is the DctA family of transporters that are expressed under aerobic conditions (Janausch *et al.*, 2002) and are members of the dicarboxylate/amino acid:cation symporter (DAACS) family (Paulsen *et al.*, 1998). DctA transporters are found in eukaryotes and prokaryotes and catalyse the symport of C₄-dicarboxylates and certain amino acids with H⁺ or Na⁺ (Janausch *et al.*, 2002). The transporters in this family contain 10-12 TMHs and are classical secondary transporters, much like lactose permease. This family includes YdbH from *Bacillus subtilis* (Asai *et al.*, 2000) and DctA from *E. coli* (Davies *et al.*, 1999).

The Dcu family of C₄-dicarboxylate transporters, represented by DcuA, DcuB and DcuC, is more versatile, as members have been shown to perform symport of C₄-dicarboxylates with H⁺, exchange (of succinate and fumarate) and efflux (symport with H⁺ or Na⁺) (Zientz *et al.*, 1996). These transporters are predicted to contain approximately 12 TMHs, although DcuA has been shown to contain 10 TMHs with *dcuA-blaM* fusions (Golby *et al.*, 1998).

Members of the CitT transporter family are also predicted to possess 12 TMHs. This family has been shown to catalyse the exchange of succinate, fumarate or tartrate with citrate (the tricarboxylic acid) leading to uptake of citrate for fermentation (Pos *et al.*, 1998).

The final group of C₄-dicarboxylate transporters is the TRAP family of transporters. Although it is now known that TRAP transporter substrates are not limited to the C₄-dicarboxylates, the first TRAP transporter identified was a C₄-dicarboxylate transporter from *R. capsulatus*.

1.8.1.2 C₄-dicarboxylate transport in *Rhodobacter capsulatus*

R. capsulatus is a purple, photosynthetic bacterium that can grow in aerobic and anaerobic conditions, photo- and chemo-heterotrophically on various organic carbon sources (Madigan & Gest, 1979). The C₄-dicarboxylate-specific TRAP transporter in *R. capsulatus* was initially identified during a transposon mutagenesis study (Hamblin *et al.*, 1990). This study identified 5 separate mutations that resulted in the loss of *R. capsulatus* growth on malate, succinate

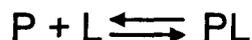
and fumarate as the sole carbon source in aerobic conditions in the absence of light. Using complementation from a library of cosmids, this study identified at least three closely associated genes that were required for uptake of the C₄-dicarboxylates into the cells and were located in the locus designated *dct* (dicarboxylate transporter, (Hamblin et al., 1990)).

The transport characteristics and substrate specificity of this transporter were studied using *in vivo* transport assays and it was confirmed to transport succinate, malate and fumarate (Shaw, 1991). This transporter, like certain ABC transporters, was found to be "shock sensitive" indicating that there was a role for an ESR during the transport cycle. The presence of an ESR was further indicated by a filter binding assay that revealed radiolabelled malate was sequestered by a protein in the periplasmic fraction prepared from *R. capsulatus*. However, unlike ABC transporters, a strong link between ΔpH , $\Delta\Psi$ and malate transport was also observed (Shaw, 1991).

The utilisation of an ESR was confirmed when it was purified and characterised *in vitro* (Shaw et al., 1991, Walmsley et al., 1992). The protein, named DctP, was found to be extremely abundant in the periplasm of wild-type *R. capsulatus*, but was absent from one of the transposon insertion mutant strains, MJH28 (Hamblin et al., 1990). DctP was purified to homogeneity and filter binding assays determined that DctP bound malate with a high affinity (K_d of $8.4 \pm 1.49 \mu\text{M}$) comparable to ABC transporter ESRs. Competition studies confirmed that succinate and fumarate were also substrates of DctP (Shaw et al., 1991). Steady-state fluorescence spectroscopy revealed that DctP bound the C₄-dicarboxylates with different affinities. L-malate bound with the highest affinity ($0.05 \mu\text{M}$), followed by fumarate and succinate (0.255 and $0.17 \mu\text{M}$, respectively) and finally D-malate ($6.3 \mu\text{M}$) (Walmsley et al., 1992). The substrate specificity of DctP *in vitro* reflected the substrate specificity of the transporter *in vivo* further reinforcing that DctP was an ESR with a similar function in transport as ABC ESRs.

The pre-steady state binding kinetics of ligand to ABC ESRs have been assessed using stopped-flow fluorescence spectroscopy for a number of different proteins with different substrate specificities, for example, galactose, maltose, histidine and

arabinose (Miller *et al.*, 1983, Miller *et al.*, 1980). In each example, the binding of protein (P) and ligand (L) were described as a rapid, 2nd order process with a simple association of ligand with a single binding site on the protein, i.e.;



The binding kinetics of C₄-dicarboxylates to DctP were found to be inconsistent with this kind of simple association. Instead, the stopped-flow analysis suggested that DctP could be in four different states during the binding mechanism; closed-unliganded (P1), open-unliganded (P2), open-liganded (P2-L) and closed-liganded (P3-L);



The data obtained suggested a slow isomerisation of the protein between P1 and P2 with the ligand only binding to P2 (Kelly & Thomas, 2001, Walmsley *et al.*, 1992). These data indicate that the ABC ESRs studied are predominantly in the open state prior to ligand binding. DctP, on the other hand is predicted to have a more stable closed-unliganded conformation. This bias in the equilibrium between the two unliganded conformations of DctP may depend on a number of variables that were not investigated, such as the ionic strength or pH of the buffer (Walmsley *et al.*, 1992). The same kinetic behaviour was observed for another TRAP ESR, RRC01191 also from *R. capsulatus* (Thomas *et al.*, 2006). Curiously, analysis of the pre-steady state kinetics of ligand binding to two other TRAP ESRs revealed kinetic behaviour similar to that seen by the ABC ESRs (Muller *et al.*, 2006, Thomas *et al.*, 2006). This demonstrates that this kinetic behaviour is not characteristic for all TRAP ESRs and may indicate a fundamental difference between ESRs within the TRAP family.

Sequencing of the 3' region downstream of *dctP* revealed two coding sequences (CDSs) of 681 and 1320 base pairs, named *dctQ* and *dctM* respectively (Forward *et al.*, 1997). Mutagenesis revealed that all three genes were required for growth on C₄-dicarboxylates (Forward *et al.*, 1997).

The gene *dctQ* encodes a protein of 24.7 kDa. Hydropathy analysis of the residues predicted 4 TMHs (62% hydrophobic residues), which was later confirmed by *dctQ*-

blaM fusions (Wyborn *et al.*, 2001) (Fig. 1.13a, left panel and b). The gene *dctM* encodes a protein with a predicted molecular weight of 46.8 kDa and 12 predicted TMHs (73% hydrophobic residues Fig. 1.13a, right panel) separated by a large hydrophilic sequence reminiscent of the 2 x 6 configuration of the 12 TMH secondary transporters such as OxlT, LacY and GlpT (Abramson *et al.*, 2003, Hirai *et al.*, 2003, Huang *et al.*, 2003). However, there is no sequence similarity between DctM and these secondary transporters; furthermore, there are none of the distinguishing features of ABC transporter, such the LSGGQ or EAA motifs (Forward *et al.*, 1997). DctQ and DctM were confirmed to be integral inner membrane proteins by heterologously co-expressing the genes from a T7-inducible vector in *E. coli*, labelling proteins with [³⁵S]-methionine and separating the inner and outer membranes with a sucrose density gradient. The proteins were visible using a fluorogram after 2 days exposure but were not visible on Coomassie-stained SDS-PAGE indicating very low abundance (Forward *et al.*, 1997).

To determine whether *dctPQM* encoded the complete transporter and to dispel the possibility that other components, like an ATP-binding cassette protein were encoded elsewhere on the genome, *dctPQM* were heterologously expressed in a strain of *Sinorhizobium meliloti* with a disruption in the sole C₄-dicarboxylate transporter, a DctA family member. Heterologous expression of *dctPQM* in this strain led to uptake of succinate, providing strong evidence that all 3 of the proteins; DctP, DctQ and DctM are necessary and sufficient for the transport of C₄-dicarboxylates (Forward *et al.*, 1997).

1.8.1.3 Transport by DctPQM is dependent on $\Delta\Psi$ and not ATP hydrolysis

To determine the energy coupling mechanism for DctPQM in *R. capsulatus*, *in vivo* transport of radiolabelled succinate was measured in relation to intracellular ATP concentration and the membrane potential. A decrease in succinate transport was observed upon dissipation of the membrane potential by titration with 3 different uncouplers; FCCP, SF6847 and 2,4-dinitrophenol. Intracellular ATP concentrations remained high during these assays indicating that the membrane potential, and not ATP hydrolysis, was responsible for succinate transport via DctPQM (Forward *et al.*, 1997). To further ratify this observation, cells were treated with orthovanadate, which is known to be a potent inhibitor of ABC transporters (Richarme *et al.*, 1993).

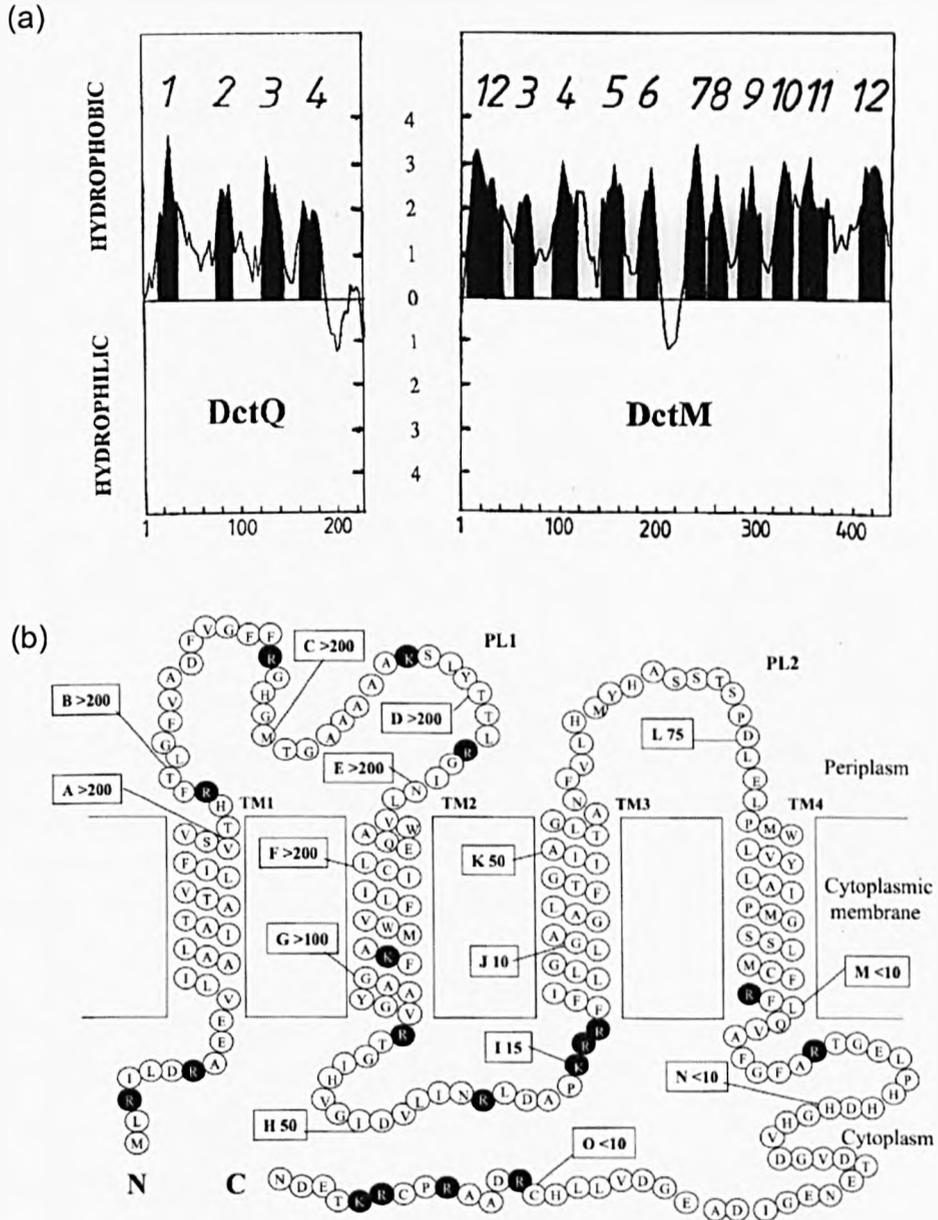


Figure 1.13 Topology modeling of DctQ and DctM from *R. sphaeroides*. (a) Hydropathy analysis of the amino acid composition of DctQ and DctM showing hydrophobic and hydrophilic residues. Transmembrane helices (TMHs) are indicated by black peaks. DctQ has 4 predicted TMHs and DctM has 12 predicted TMHs. A large hydrophilic loop is observed between helices 6 and 7 of DctM. Figure taken from (Forward et al., 1997) (b) Experimentally defined topology model of DctQ. Positively charged residues are shown in black. The locations of BlaM fusions used to define the model are labelled A-O. Figure taken from Wyborn *et al* (2001).

Transport of C₄-dicarboxylates via DctPQM was orthovanadate insensitive. The lactose permease control was also orthovanadate insensitive, whereas an alanine-specific ABC transporter was 80% inhibited. Combined, this evidence strongly suggested that the pmf, and not ATP binding and hydrolysis, was required for transport of C₄-dicarboxylates by DctPQM.

Through this work on DctPQM, predominantly performed in Prof. David Kelly's laboratory in Sheffield, a new class of transporter was identified that combined an ESR with the use of an electrochemical gradient to power transport of substrate across the membrane. DctPQM was the first TRAP transporter described and characterised in detail. As such, all subsequently discovered TRAP transporters that were homologous to these proteins were placed in the DctP, DctQ or DctM families.

1.8.2 Phylogenetic analysis of the DctP, DctQ and DctM families of proteins

Phylogenetic analyses of the components DctPQM has provided a number of insights into the prevalence, genetic organisation and divergence of TRAP transporters in bacteria and archaea (Kelly & Thomas, 2001, Rabus *et al.*, 1999).

1.8.2.1 The DctP family of ESRs

Phylogenetic analysis of ESRs can be used to group the proteins into families based on their sequence similarity, where members of a particular family will bind chemically similar ligands (Tam & Saier, 1993). This can be performed for *dctP* homologues and can be used in conjunction with the genome context to predict substrate specificities of TRAP systems. As an example of this, the YiaMNO TRAP transporter from *E. coli* is in an operon with the genes necessary for the utilization of 2,3-diketo-L-gulonate (2,3-DKG) and has been predicted to be a 2,3-DKG specific transporter (Yew & Gerlt, 2002). Although some literature implicated YiaMNO in L-xylulose transport (Plantinga *et al.*, 2004), the purified ESR, YiaO has been shown conclusively to bind 2,3-DKG and not L-xylulose using a number of complementary biophysical techniques (Thomas *et al.*, 2006). *H. influenzae* contains 3 complete TRAP transporters, one of the ESRs; HI0128 has high sequence similarity with YiaO, therefore one would predict that HI0128 also binds

2,3-DKG (Kelly & Thomas, 2001). This high degree of sequence similarity of TRAP ESRs from different organisms, including archaea, indicates that these orthologues have arisen by vertical transfer and not horizontal transfer, i.e. from speciation. This indicates that TRAP transporters are ancient, and were present before the split between bacteria and archaea occurred (Rabus *et al.*, 1999).

Multiple *dctP* homologues have been found in some organisms, and have also been found to be highly divergent indicating that they bind structurally different substrates. There are also examples of paralogues having very high sequence similarity at the amino acid level (>70%), for example PA0884, PA4616 and PA5167 from *P. aeruginosa*, suggesting that they bind chemically similar substrates and arose from a recent gene duplication event (Kelly & Thomas, 2001). Phylogenetic analysis of DctP has also revealed a TRAP ESR being used in a non-transport related context. YdbE in *B. subtilis* is an orphan TRAP ESR and is not associated with *dctQ* or *dctM* genes. *B. subtilis* does not contain any *dctQ* or *dctM* orthologues, strongly suggesting that YdbE is not part of a complete TRAP transporter (Kelly & Thomas, 2001). *ydbE* is divergently transcribed from a 2 partner sensor-regulator pair, *ydbF* and *ydbG*, and a gene encoding a C₄-dicarboxylate transporter of the DcuA-family, *ydbH*. YdbE is an experimentally determined extracytoplasmic C₄-dicarboxylate sensor that binds substrate and interacts with the 2 component sensor-regulator pair to activate transcription of *ydbH* (Asai *et al.*, 2000). This is analogous to MalE, which can work as either a transport associated ESR or as a sensor for chemotaxis (Springer *et al.*, 1977). However, YdbE does not have cognate integral membrane components so acts solely as a C₄-dicarboxylate sensor.

1.8.2.2 The DctM family of integral membrane proteins

Phylogenetic analysis of the *dctM* family has revealed that the nucleotide sequence of these genes has diverged least out of all TRAP genes, suggesting that it has a central role in transport and does not necessarily define the substrate specificity (Rabus *et al.*, 1999).

Members of the DctM family are ~350-450 amino acids long and possess 10-13 TMHs that are separated by a centrally located hydrophilic sequence of amino acids, similar to other secondary transporters that consist of 12 TMHs (Prakash *et*

al., 2003). It has been proposed that DctM is the translocation channel through which the substrate passes to cross the membrane. Analysis of the amino acid sequence has revealed that DctM shares a sequence motif with DcuC – a C₄-dicarboxylate secondary transporter that does not require any auxiliary protein components to function – indicating that DctM has a weak but significant relationship with classical secondary transporters (Prakash et al., 2003, Rabus et al., 1999).

DctM is a member of the ion transporter (IT) superfamily which consists of secondary transporters specific for anionic and cationic compounds (Prakash et al., 2003). Other members of this superfamily include the arsenite–antimonite (ArsB) efflux family, which can consist of 2 proteins ArsA, an ATPase domain and ArsB, a 12 TMH translocation channel. In the presence of ArsA the transporter is ATP-driven. However, in the absence of ArsA, ArsB uses the PMF to power transport (Dey & Rosen, 1995). This demonstrates that ArsB and DctM are translocation channels and are able to incorporate ancillary proteins to modify their transport mechanism.

Although it has not been explicitly demonstrated experimentally, these findings demonstrate that DctM is a very strong candidate to be the translocation channel for TRAP transporters.

1.8.2.3 The DctQ family of integral membrane proteins

Phylogenetic analysis reveals that DctQ is the most divergent of the TRAP components and has no sequence similarity with any other family of proteins (Kelly & Thomas, 2001, Rabus et al., 1999). DctQ is of unknown function, but has been shown to be essential for transport (Forward et al., 1997). It has been postulated that DctQ may be involved in coupling ESR recognition to the translocation channel. It has also been suggested that DctQ acts as a landing pad for the ESR, or DctQ may be involved in the folding/insertion of DctM into the membrane (Kelly & Thomas, 2001). Unfortunately, there is very little evidence for any of these theories and more experimental work will be required to ascertain what role DctQ has in the transport cycle.

1.8.3 Genetic arrangement of TRAP transporters

Phylogenetic analysis has revealed that the organization of the genes encoding TRAP transporters varies considerably (Fig. 1.14 (Kelly & Thomas, 2001)). Operons have been discovered where the three proteins are expressed from 3 separate genes, for example, DctPQM from *R. capsulatus* or YiaMNO from *E. coli* (Fig. 1.14). These two systems differ in the position of the *dctP* gene in relation to the *dctQ* and *dctM* genes. The former has *dctP* upstream and latter has *dctP* downstream of the *dctQ* and *dctM* positions.

The only example of a *dctP-dctQ* fusion is from *Rhizobium* sp. Strain NGR 234 (Fig. 1.14).

Although DctQ and DctM have been treated as separate entities until now, there are a number of examples of *dctQ-dctM* fusions, for example SiaQM from *H. influenzae* (Fig. 1.14). This is a commonly observed fusion in a number of different TRAP transporters, however, this organisation is not as common as encoding the 3 components from 3 different genes. Although fused, these fusion proteins have easily discernable DctQ and DctM domains. This fusion is found in TRAP transporters from a number of different bacteria, but curiously, all TRAP transporters in archaea have a DctQM fusion protein (Kelly & Thomas, 2001).

There is an example of a TRAP system with 2 *dctP* genes associated with *dctQ* and *dctM* in *Trepomena pallidum* (Fig. 1.14) and examples of *dctP* genes divergently and convergently transcribed from the *dctQ* and *dctM* genes from *Actinobacillus actinomycetemcomitans* and *Rhodopseudomonas palustris*, respectively (Fig. 1.14).

These observations reveal a number of insights into TRAP transporters. The example of a *dctP-dctQ* fusion adds weight to the possibility that DctQ acts as a bridge between the ESR and the translocation channel. The phylogenetic analysis reveals that there are no examples where a *dctM* gene is transcribed before a *dctQ* gene. This implies that *dctQ* must be transcribed/translated first, which suggests DctQ may be involved in the stability/insertion of DctM. To reinforce this postulate, it has also been noted that DctM can not be accumulated unless co-expressed with DctQ (Kelly & Thomas, 2001).

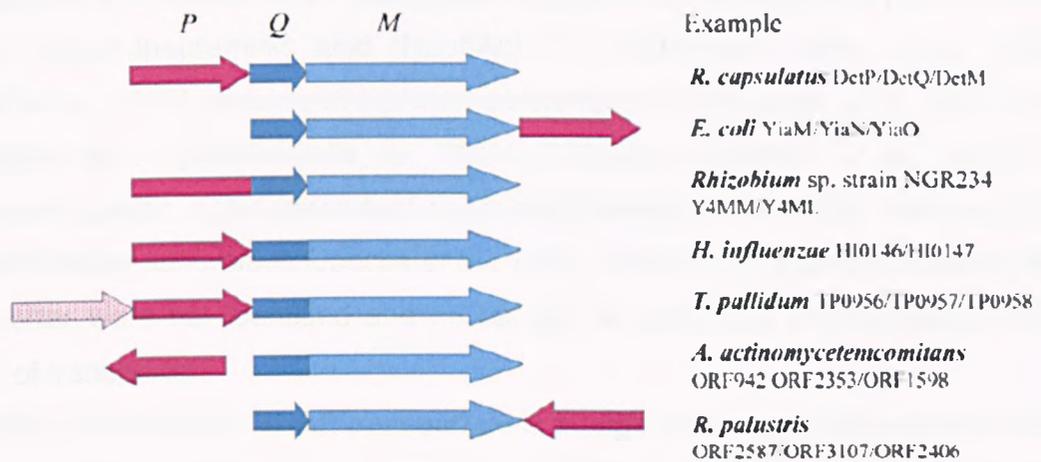


Figure 1.14 Organisation of genes encoding TRAP transporters containing DctP homologues. Genes encoding the integral membrane proteins, DctQ and DctM are coloured dark blue and light blue, respectively. In some of the cases, the *dctQ* and *dctM* genes are fused. The *dctP* homologues flanking the genes encoding DctQ and DctM are shown in red. It is unknown whether the gene coloured in pink in *T. pallidum* is a true *dctP* homologue. Examples of organisms and genes where these different organizations can be observed are shown on the right. Figure modified from Kelly & Thomas (2001).

1.8.4 Known and predicted substrates of TRAP transporters

Since their discovery, a number of TRAP transporter substrates have been identified. This has revealed that TRAP transporters are not limited to C₄-dicarboxylate transport, but actually have a wide variety of substrates. The following compounds are experimentally determined substrates of TRAP transporters (Fig. 1.15a); C₄-dicarboxylates (succinate, fumarate and malate) in *R. capsulatus* (Forward et al., 1997) and *Wolinella succinogenes* (Ullmann et al., 2000), *N*-acetylneuraminic acid (Neu5Ac) in *H. influenzae* (Allen et al., 2005, Severi et al., 2005), ectoine in *Halomonas elongata* (Grammann et al., 2002) and glutamate from *Synechocystis* sp. Strain PCC6903 (Quintero et al., 2001). A glutamate specific, ESR-dependent secondary transporter has also been reported in *Rhodobacter sphaeroides* (Jacobs et al., 1996). However, the genes encoding the transporter were not identified and so cannot be confirmed as belonging to this family of transporter.

The ESR component of TRAP transporters is thought to confer the specificity to the transporter. There have been a number of TRAP ESRs characterised and the substrate specificities determined, including (Fig. 1.15b); 2,3-diketo-L-gulonate from *E. coli* and monocarboxylate keto-acids from *R. sphaeroides* (Thomas et al., 2006). Two ESRs from *Bordetella pertussis* have been crystallised and found to be bound to pyroglutamate (Rucktooa et al., 2007).

A study of transcriptional inducers of the ESR components of TRAP transporters in *Sinorhizobium meliloti* has led to the prediction of a number of new substrates including (Fig. 1.15c); malonate, quinate, oxobutyrate, pyruvate and a number of sugars including mannose, raffinose, fucose, arabinose and lactose (Mauchline et al., 2006). These findings increase the range of possible TRAP transporter substrates considerably, but obviously need to be experimentally verified.

Finally, there are the substrates that are predicted from genome context and phylogenetic analysis, which includes (Fig. 1.15c); taurine from both *R. capsulatus* (Bruggemann et al., 2004) and *R. sphaeroides* 2.4.1 (Denger et al., 2006) and 4-chlorobenzoate from *Pseudomonas* sp. DJ-12 (Chae et al., 2000).

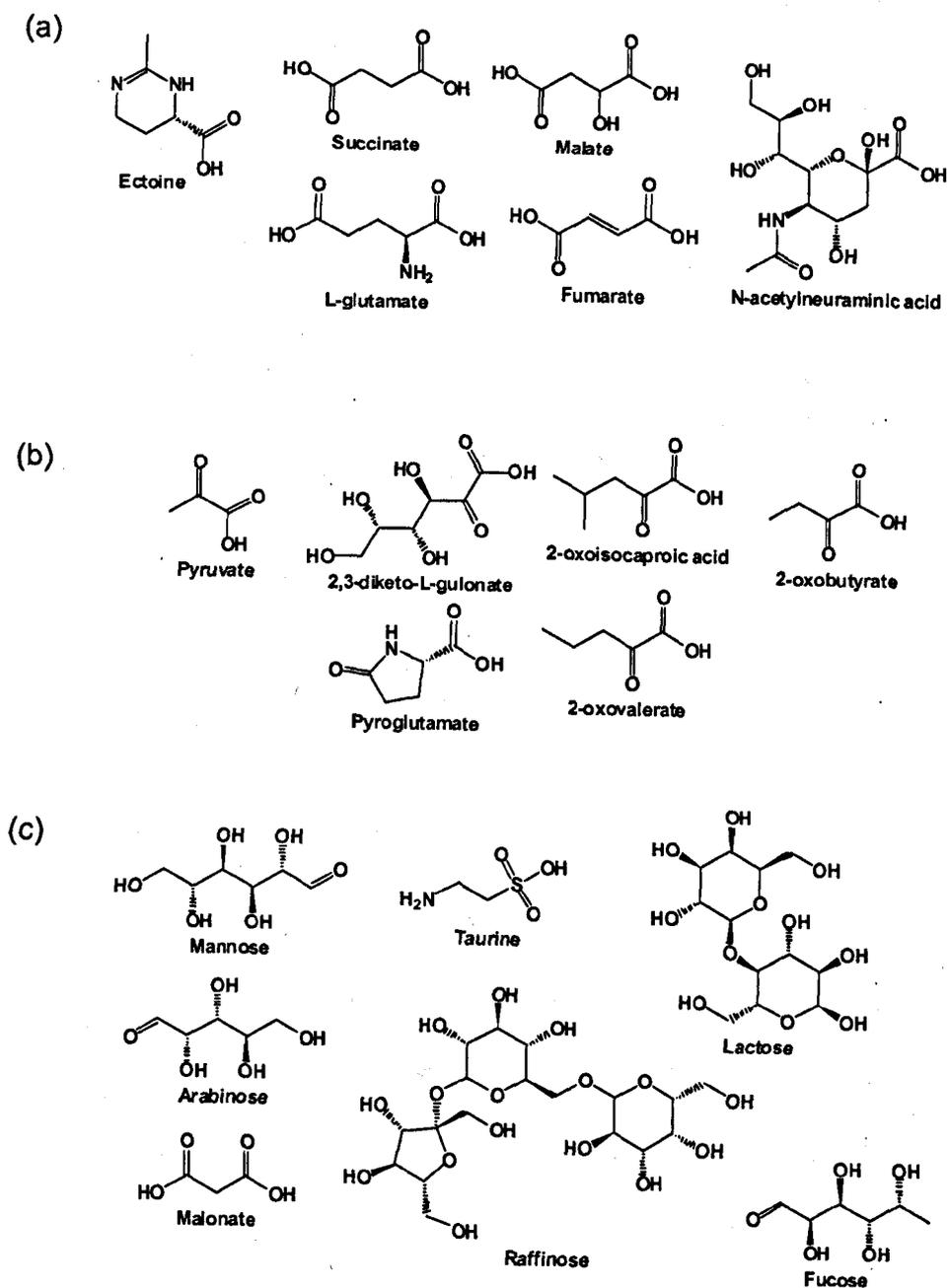


Figure 1.15 Known and predicted TRAP transporters substrates. (a) TRAP transporter substrates experimentally defined using *in vivo* studies. (b) TRAP transporter substrates defined by *in vitro* characterization of purified ESR components. (c) TRAP transporter substrates predicted from phylogenetic and genome context analysis or from transcription induction studies.

1.9 SiaPQM: an *N*-acetylneuraminic acid TRAP transporter from *H. influenzae*
SiaPQM is a TRAP transporter from *H. influenzae* that has been shown to be essential for transport of the sialic acid, *N*-acetylneuraminic acid (Neu5Ac) *in vivo* (Allen et al., 2005, Severi et al., 2005). This transporter consists of an ESR, SiaP and a single integral membrane component, SiaQM.

1.9.1 An introduction to sialic acids

Sialic acids have been studied for over 40 years and were once said to be “not only the most interesting molecules in the world, but also the most important” (Vimr, 1994), so it would be remiss to attempt to encapsulate the breadth of sialic acid research into a few paragraphs. There are however, a number of excellent reviews on this subject that will cover the areas not discussed (Severi et al., 2007, Vimr et al., 2004).

The term sialic acid encompasses over 40 different naturally occurring 9-carbon sugar acids. The parent compound from which these compounds are derived, and the most common sialic acid, is *N*-acetylneuraminic acid (Neu5Ac) (Angata & Varki, 2002). Throughout the course of this section the terms sialic acid and Neu5Ac will be used interchangeably unless specifically stated otherwise.

A feature of all higher metazoan cells is the presence of a sialic acid molecule as the outermost constituent on surface exposed glycoconjugates (glyco-proteins and -lipids). On the eukaryotic cell surface, these sialylglycoconjugates have been shown to have a number of important functions. They prevent complement activation and autoimmunity by interacting with innate immune response factors, can take part in cell-to-cell communication and adhesion, and they can act as chemical messengers, regulating receptor and membrane transport functions (Varki, 1993, Varki, 1997, Vimr et al., 2004). Bearing in mind the importance and apparent abundance of these molecules, it is not surprising to find that microorganisms that interact with higher eukaryotic organisms have developed methods of harnessing sialic acids for their own use, sometimes to the detriment of the host (Severi et al., 2007, Vimr et al., 2004). It has become apparent that a number of human commensals and pathogenic bacteria are able to utilize sialic acid as a source of carbon, nitrogen, energy and amino sugars for use in cell wall synthesis (Plumbridge & Vimr, 1999). Sialic acids are now well known to be

virulence factors in a number of pathogenic organisms that are able to sialylate their surface-exposed lipopolysaccharide (LPS) and masquerade as “self”, thus circumventing the innate immune response of the host (Bouchet *et al.*, 2003, Figueira *et al.*, 2007, Hood *et al.*, 1999, Vimr *et al.*, 2000). In some organisms there is a dual fate for sialic acid; acquired sialic acid is either catabolised and used for biosynthesis or it is used to decorate the outer surface of the organism for immune evasion tactics (Vimr *et al.*, 2000).

1.9.2 Acquisition of sialic acid

There are two major methods by which bacteria acquire sialic acid; either by *de novo* synthesis or by scavenging from the environment (Vimr *et al.*, 2004). Synthesis of sialic acid is performed by a number of organisms including *E. coli* K1, *Neisseria meningitidis* and *Campylobacter jejuni* (Vimr *et al.*, 2004) using the cell wall synthesis precursor, CMP-GlcNAc, as a starting point. The actions of two enzymes, NeuC and NeuB, produce Neu5Ac via ManNAc (Vimr *et al.*, 2004). Neu5Ac can then be catabolised or used to decorate the LPS. The second method of obtaining sialic acid is from the environment, in the case of pathogens, this is the mammalian host. Although most of the sialic acids will be bound to glycoconjugates, there is such an abundance of sialoglycoconjugates on the cell surfaces of mammalian cells that even a conservative estimate of the percentage (0.05%) that is in the free form produces an appreciable amount of free sialic acid for bacteria to scavenge (Vimr *et al.*, 2000). Certain bacteria, for example *Vibrio cholerae* also encode a neuraminidase (also known as a sialidase) that can be used to free sialic acid from the host glycoconjugate (Corfield, 1992). Neuraminidases can also be used to desialylate other bacteria to free sialic acid (Shakhnovich *et al.*, 2002), or they can scavenge sialic acid freed from host sialidase action in response to inflammation (Sohanpal *et al.*, 2004, Sohanpal *et al.*, 2007).

1.9.3 Utilisation of sialic acid in bacteria

The enzymes that are involved in sialic acid catabolism have been identified in *E. coli*, and homologues of these enzymes are found in a number of other bacteria (Vimr *et al.*, 2004). The first isolated mutants defective in Neu5Ac utilisation were designated *nan* mutants (for *N*-acetylneuraminate) and were found to have

disruptions in the genes, *nanA* and *nanT* (Vimr & Troy, 1985a). These were found to encode the transporter NanT, and an aldolase NanA. A number of biochemical studies elucidated the pathway of Neu5Ac catabolism (Fig. 1.16a); Neu5Ac is transported into the cytoplasm and broken down into ManNAc and pyruvate by the action of NanA (Vimr & Troy, 1985a, Vimr & Troy, 1985b). ManNAc is phosphorylated by the ATP-dependent enzyme Nank to produce ManNAc-6-P, which is then converted to GlcNAc-6-P by the actions of a reversible epimerase, NanE (Ringenberg *et al.*, 2003). GlcNAc-6-P is deacetylated by NagB producing GlcN-6-P which is deaminated to Fructose-6-P and can be used in glycolysis (Plumbridge & Vimr, 1999).

One of the more exciting aspects of sialic acid biology is the fact it has potentially two roles to play; it can be used for biosynthesis or can be used to decorate the bacterial cell surface where it has a number of functions including host immune response evasion and a role in adhesion/invasion to host cells (Karlsson, 1998, Vimr *et al.*, 2000). There are a number of different methods by which cell surface glycoconjugates are sialylated. The method most pertinent to this study is a process termed precursor scavenging, best described in *H. influenzae*, in which Neu5Ac is taken into the cytoplasm and activated by CMP-Neu5Ac synthetase (SiaB in *H. influenzae* or NeuA in *E. coli*) (Vimr *et al.*, 2000). Activated Neu5Ac is then added to the lipooligosaccharide (LOS) by the actions of sialyltransferases, of which *H. influenzae* has 3; Lic3A, Lic3B and SiaA (Fox *et al.*, 2006, Hood *et al.*, 2001) (Fig. 1.16a). The sialylated LOS is then presented on the cell surface using the normal LOS biosynthesis machinery (Bos *et al.*, 2007).

1.9.4 Bacterial sialic acid (Neu5Ac) transporters

Whether scavenged from the environment or released by neuraminidase action, bacteria still need to transport Neu5Ac into the cytoplasm because it has been demonstrated that the catabolic enzymes and enzymes involved in LPS sialylation are cytoplasmic (Lilley *et al.*, 1998, Vimr *et al.*, 2000). To do this, bacteria have employed a number of different classes of transporter. The first bacterial Neu5Ac transporter identified and characterised was NanT from *E. coli* (Vimr & Troy, 1985a). This is the sole uptake route for Neu5Ac in *E. coli*, a classical secondary

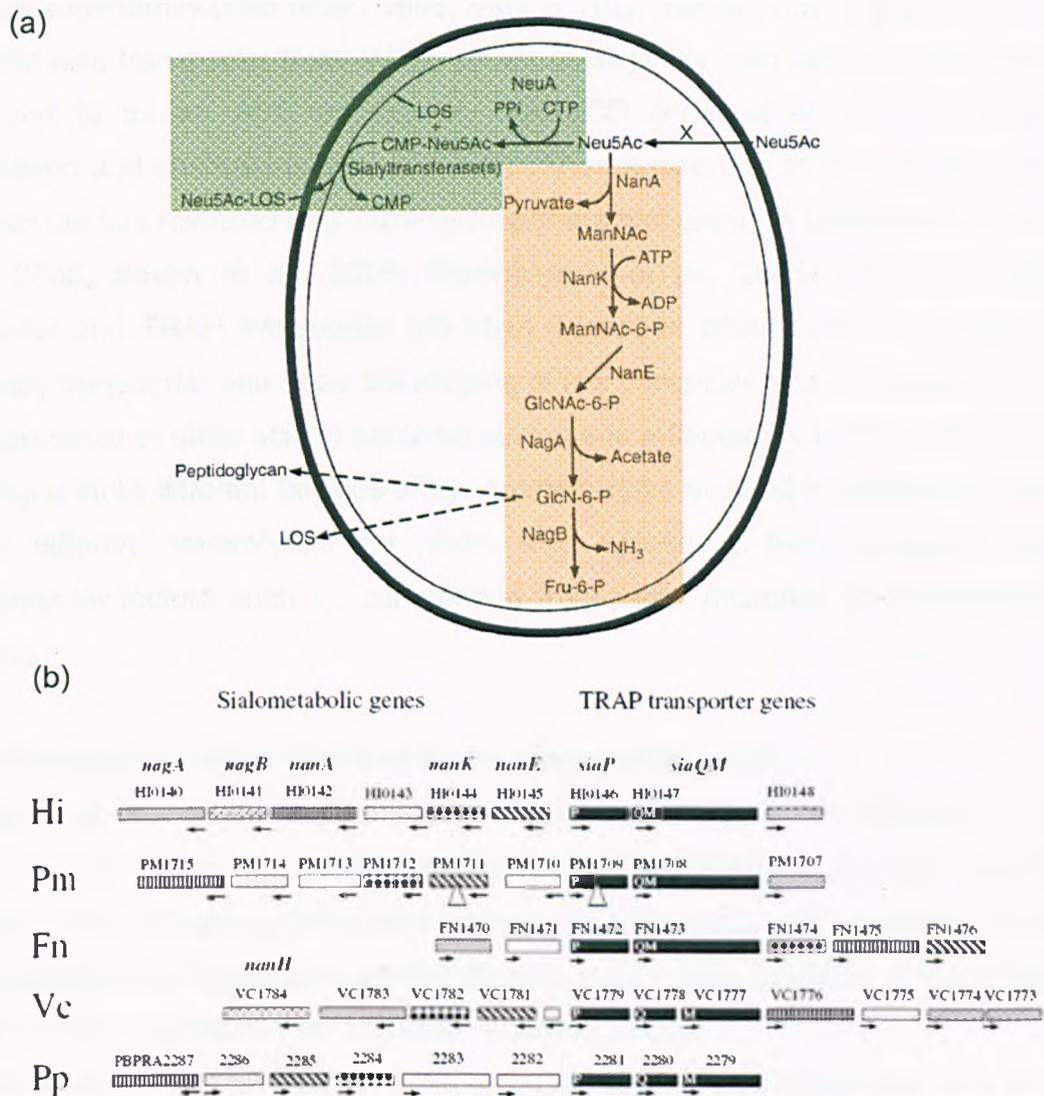


Figure 1.16 (a) Dual function of Neu5Ac in *H. influenzae* and (b) genetic arrangement of sialometabolic operon in *H. influenzae* and related species.

(a) Catabolic pathway of Neu5Ac resulting in formation of fructose-6-phosphate (Fru-6-P) shaded orange. Pathway for LOS sialylation by Neu5Ac shaded green. Figure modified from Vimr *et al* (2000). X is the Neu5Ac specific transporter, which is SiaPQM in *H. influenzae*. (b) Schematic of sialometabolic operons from 5 organisms containing genes for TRAP transporters (solid black, *siaP* and *siaQM*) and sialometabolic genes. Organisms include *H. influenzae* (Hi), *P. multocida* (Pm), *Fusobacterium nucleatum* (Fn), *Vibrio cholerae* (Vc), *Photobacterium profundum* (Pp). Figure taken from Severi *et al* (2005).

transporter and member of the sialate:H⁺ symporter subfamily (2.1.12) of the major facilitator superfamily (Pao et al., 1998, Vimr & Troy, 1985a, Vimr & Troy, 1985b). The sialic acid transporter from *Haemophilus ducreyi* has also been characterised and found to be an ABC transporter, SatABCD (Post et al., 2005). Finally, identification and characterisation of the Neu5Ac transporters in *P. multocida* and *H. influenzae* has revealed that TRAP transporters also perform this function (Allen et al., 2005, Severi et al., 2005, Steenbergen et al., 2005). Both the ABC transporter and TRAP transporter will have increased affinity over the classical secondary transporter and allow scavenging of low concentrations of Neu5Ac.

The importance of sialic acid in bacterial systems is reflected by the fact that there are at least three different families of transporters have evolved to transport it. The use of different transporters for sialic acid uptake is likely influenced by environmental factors such as competition from other microbes and scarcity of substrate.

1.9.5 *Haemophilus influenzae* and N-acetylneuraminic acid

H. influenzae is a common Gram negative nasopharyngeal commensal that is found in up to 80% of the population (Bouchet et al., 2003). It is also a human pathogen with encapsulated strains causing septicaemia and meningitis and capsule-deficient (nontypeable) strains causing otitis media (infection of the middle ear) and acute respiratory tract infections (Moxon, 2000).

H. influenzae contains all the genes necessary to enable it to utilise sialic acid as a carbon and nitrogen source (Vimr et al., 2000) and it has also been observed that a high proportion of *H. influenzae* strains, both capsule-deficient and capsular, contain sialic acid in their LPS (Hood et al., 1999). Disruptions of genes in the LPS sialylation pathway have demonstrated that sialylation of the LPS in *H. influenzae* increases the resistance to the complement-mediated killing effect of normal human serum (Hood et al., 1999). Sialylation of the LPS has been shown to be an important virulence factor for *H. influenzae* mediated otitis media in chinchilla (Bouchet et al., 2003). In *H. influenzae*, the catabolic pathway and the sialylation of the LPS both require sialic acid to be in the cytoplasm. It is unknown how the utilization of sialic acid is regulated between these two separate pathways; however, it has been shown that disruption of the first gene in the catabolic

pathway, *nanA*, leads to hypersialylation of the LPS and increased resistance to human serum (Vimr et al., 2000).

H. influenzae does not have the enzymes necessary to synthesize sialic acid *de novo* and has been shown to incorporate sialic acid into its LPS by scavenging Neu5Ac from its host (Bouchet et al., 2003).

The TRAP transporter SiaPQM was first implicated in sialic acid transport in *H. influenzae* due to its close proximity to sialometabolic genes in the genome (Kolker et al., 2004, Severi et al., 2005, Steenbergen et al., 2005). The linkage of TRAP transporter and sialometabolic genes has also been observed in other organisms including *Pasteurella multocida*, *Vibrio cholerae*, *Photobacterium profundum* and *Fusobacterium nucleatum* (Fig. 1.16b). In *H. influenzae*, the *siaP* gene that encodes the ESR and the *siaQM* gene, which encodes a fused integral membrane component are transcribed with *HI0128* (a homologue of *nanM* (Severi et al., 2008)) and divergently transcribed from the sialometabolic genes *nanE*, *nanK*, *nanA*, *nagB* and *nagA*. This close association of transporter and metabolic genes strongly suggested that the substrate of the transporter was Neu5Ac.

1.9.6 Characterisation of SiaPQM from *H. influenzae*

1.9.6.1 Characterisation of SiaPQM *in vivo*

SiaPQM was confirmed to be Neu5Ac specific transporter in two separate, but complementary studies on the unencapsulated laboratory *H. influenzae* strain RW118 (Rd) and the nontypeable strain, 2019 (Allen et al., 2005, Severi et al., 2005). Disruptions in either of the genes *siaP* and *siaQM* in *H. influenzae* prevented transport of radiolabelled Neu5Ac into the cell, measured using *in vivo* transport assays (Allen et al., 2005, Severi et al., 2005). These same studies also showed that HI1104, a homologue of NanT from *E. coli* is not involved in transport of Neu5Ac under these conditions in either strain of *H. influenzae*.

A sialylated glycoform of LPS is easily differentiated from LPS lacking sialic acid using SDS-PAGE due to the differences in molecular weight imparted by sialylation (Hood et al., 2001, Hood et al., 1999). When the LPS profile of *siaP* and *siaQM* mutants were compared to the LPS profile from wildtype *H. influenzae* it was found to be significantly different, consistent with lack of sialylation of the LPS in the

mutant strains. The LPS profile from wildtype *H. influenzae* cells was found to be sensitive to neuraminidase treatment which led to removal of Neu5Ac from LPS and a decrease in the molecular weight. The LPS profile of the *siaP* and *siaQM* mutants was insensitive to neuraminidase treatment further confirming the absence of sialylated LPS in the mutant strains (Allen et al., 2005, Severi et al., 2005).

Sialylated LPS is known to confer resistance to human serum and consistent with SiaPQM being essential for Neu5Ac uptake and therefore sialylation of LPS, there was decreased survivability of the *siaP* and *siaQM* mutants in normal human serum (Allen et al., 2005, Severi et al., 2005). These data show that SiaPQM is essential for the uptake of Neu5Ac indicating that it is the sole uptake route for Neu5Ac in *H. influenzae* under these conditions.

1.9.6.2 In vitro characterisation of SiaP

It has been demonstrated for ABC transporters that it is the ESR component that confers the specificity to the transporter (Doeven et al., 2004, Wilkinson, 2003), therefore, the *in vitro* characteristics of SiaP were defined.

SiaP was purified to homogeneity in the absence of Neu5Ac and the ligand binding was characterised using electrospray ionization mass spectroscopy (ESI-MS), tyrosine fluorescence spectroscopy and stopped-flow fluorescence spectroscopy (Allen et al., 2005, Muller et al., 2006, Severi et al., 2005). The mass of SiaP in the presence and absence of Neu5Ac was determined using ESI-MS. Unliganded SiaP was revealed to have a mass of 34165 Da. Addition of Neu5Ac led to an increase of 309 Da to 34473 Da indicating that SiaP was bound to a single molecule of Neu5Ac (Severi et al., 2005). The binding of Neu5Ac to SiaP was confirmed with tyrosine fluorescence spectroscopy, which revealed that SiaP bound Neu5Ac with a high affinity of 120 nM (Severi et al., 2005). These two techniques were also used to determine that SiaP did not bind to other known TRAP transporter substrates such as C₄-dicarboxylates, pyruvate or α -ketobutyrate (Severi et al., 2005). The high affinity binding of Neu5Ac to SiaP and the 1:1 stoichiometry was confirmed using isothermal titration calorimetry (ITC) with a calculated K_d of 28 nM (Johnston et al., 2008).

The pre-steady state kinetics of Neu5Ac binding to SiaP were established using stopped-flow fluorescence spectroscopy (Muller et al., 2006). The data obtained

was consistent with a simple bimolecular interaction similar to that seen in ABC ESRs (Miller et al., 1983, Miller et al., 1980) and to another TRAP ESR, YiaO from *E. coli* (Thomas et al., 2006). The kinetic behaviour of SiaP and YiaO are not consistent with the kinetics measured for the TRAP ESRs, DctP and RRC01191 both from *R. capsulatus* suggesting that there may be more than one kinetic mechanism used by TRAP ESRs.

1.9.6.3 High resolution crystal structure of SiaP

Two independent studies have produced high resolution crystal structures of SiaP in the presence and absence of bound ligand (Johnston et al., 2008, Muller et al., 2006). The first structure to be solved was of SiaP as it is found in *H. influenzae* RW118 (Rd). This structure was solved with a resolution of 1.7 Å in both the absence of ligand and in the presence of a Neu5Ac analogue, Neu5Ac2en (Fig. 1.17a (Muller et al., 2006)) This analogue has been shown to have a binding affinity of 20 µM with SiaP, 167-fold lower than that of the true substrate, Neu5Ac (Severi et al., 2005). The second SiaP structure to be solved was the protein from *H. influenzae* strain 2019 to a 1.4 Å resolution in the presence of Neu5Ac (Johnston et al., 2008). The two structures are predominantly the same except that the binding site interactions deviate slightly due to the difference in the bound substrates. Neu5Ac2en differs from Neu5Ac by having a double bond between carbons (C) 2 and 3 of the sugar ring resulting in a partial flattening of the normal chair conformation and removal of the hydroxyl group on C2 (Fig. 1.17c).

SiaP is composed of 2 globular α/β domains connected by 3 polypeptides segments forming a “hinge” region, one of which is a long α -helix that is unique to this ESR structure. This structure is very similar to that seen for ABC ESRs. The glycine-betaine ABC ESR (PDB 19RL) and a number of other ABC transporter ESRs were the most similar structures in a DALI search of the SiaP structure despite there being no sequence similarity (Muller et al., 2006). The structure of SiaP reveals that it is also a type II ESR as seen for ABC ESRs (Fukami-Kobayashi *et al.*, 1999).

The structure of the ligand bound versions of SiaP reveal that it is closed with a single molecule of ligand in the binding site, confirming the previous ITC and ESI-

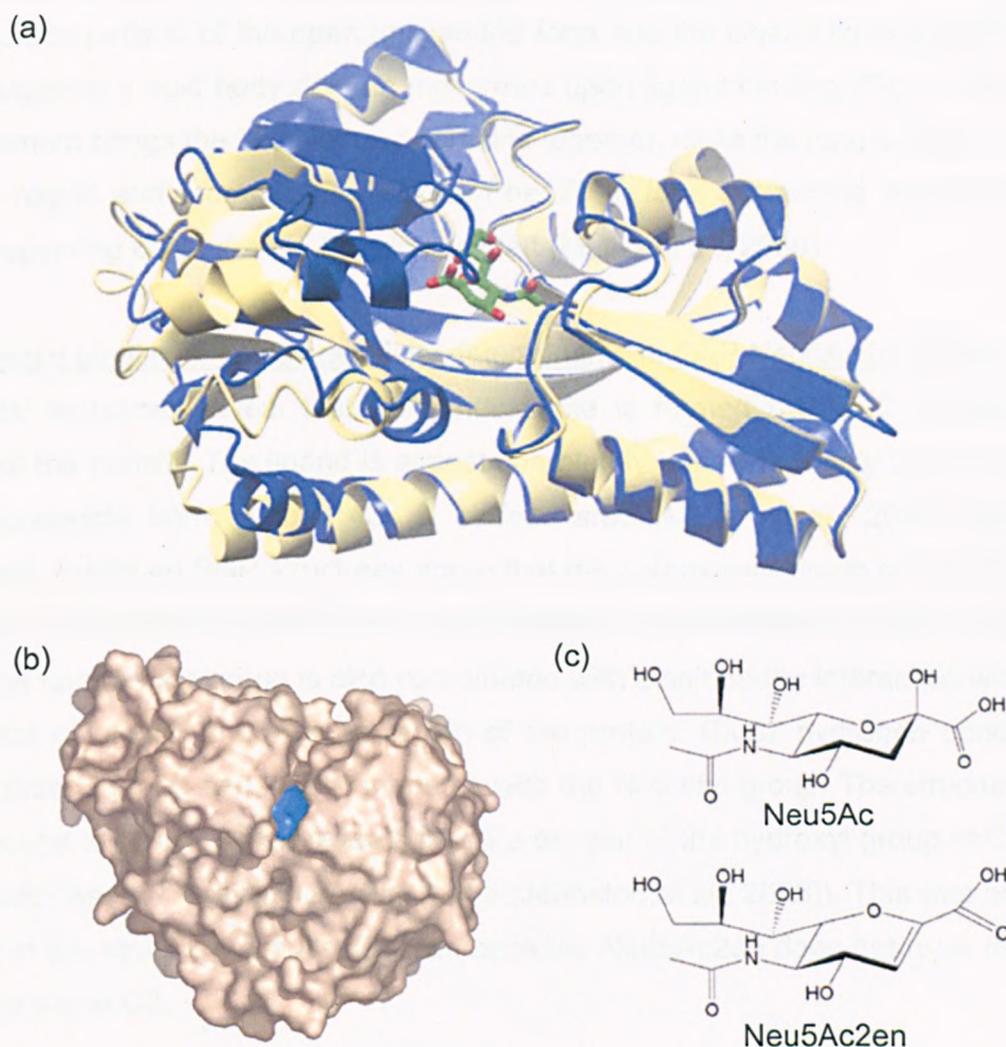


Figure 1.17 Crystal structure of SiaP in ligand-free and ligand-bound forms. (a) Superimposition of ligand-free SiaP (yellow) and ligand-bound SiaP (blue) showing conformational changes induced upon ligand binding. Neu5Ac2en is present in the binding site (green). (b) Space filling model of ligand-bound SiaP showing almost complete envelopment of Neu5Ac2en (blue) by SiaP (light brown). (c) Structural differences between Neu5Ac and Neu5Ac2en. Dehydration of carbon 2 in Neu5Ac leads to double bond formation and loss of hydroxyl group to form Neu5Ac2en. Figures modified from Muller *et al* (2006).

MS results showing a 1:1 ratio of ligand to protein (Johnston et al., 2008, Severi et al., 2005). Comparison of the open unliganded form and the closed liganded form of SiaP suggests a rigid body domain movement upon ligand binding (Fig. 1.17a). This movement brings the two globular domains together, kinks the long α -helix in the hinge region and reorients the residue Phe170 to form a stacking interaction with the sugar ring of Neu5Ac in the binding cleft (Muller et al., 2006).

1.9.6.4 Ligand binding site and specificity determinants of SiaP:Neu5Ac interaction

The crystal structures reveal that the binding site is formed by the 2 globular domains of the protein. The ligand is almost completely buried with only 32 Å² still solvent accessible from its total 435 Å surface area (Muller et al., 2006) (Fig. 1.17b). Both published SiaP structures agree that the carboxylate group of Neu5Ac forms a salt bridge with Arg147 and polar interactions with Asn187 of SiaP (Fig. 1.18a). The carboxylate group is also coordinated with a salt bridge interaction with Arg127 that is located in the hinge region of the protein. Glu67 hydrogen bonds with the glycerol group and Asn10 interacts with the N-acetyl group. The structure of SiaP bound to Neu5Ac also reveals that the oxygen of the hydroxyl group at C2 interacts with Asn187 and Arg127 (Fig. 1.18b (Johnston et al., 2008)). This was not observed in the structure with Neu5Ac2en because Neu5Ac2en does not have the hydroxyl group at C2.

The importance of these interactions was ascertained by two different approaches; by altering the nature of the ligand by using Neu5Ac analogues (Fig. 1.19), or altering the nature of the protein by introducing mutations into the binding site (Johnston et al., 2008, Muller et al., 2006). In the first approach, ESI-MS and tyrosine fluorescence spectroscopy were used to measure binding of ligand to SiaP (Muller et al., 2006). In the second approach an *in vivo* phenotype of LPS sialylation (measured by whole cell ELISA assays and matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (MALDI-TOF MS)) was used to determine the effects of point mutations in the binding site of SiaP (Johnston et al., 2008).

It is clear from the structures that residue Arg147 forms a salt bridge with the carboxylate group in Neu5Ac. The binding affinity for sialyl amide (Fig. 1.19f), an

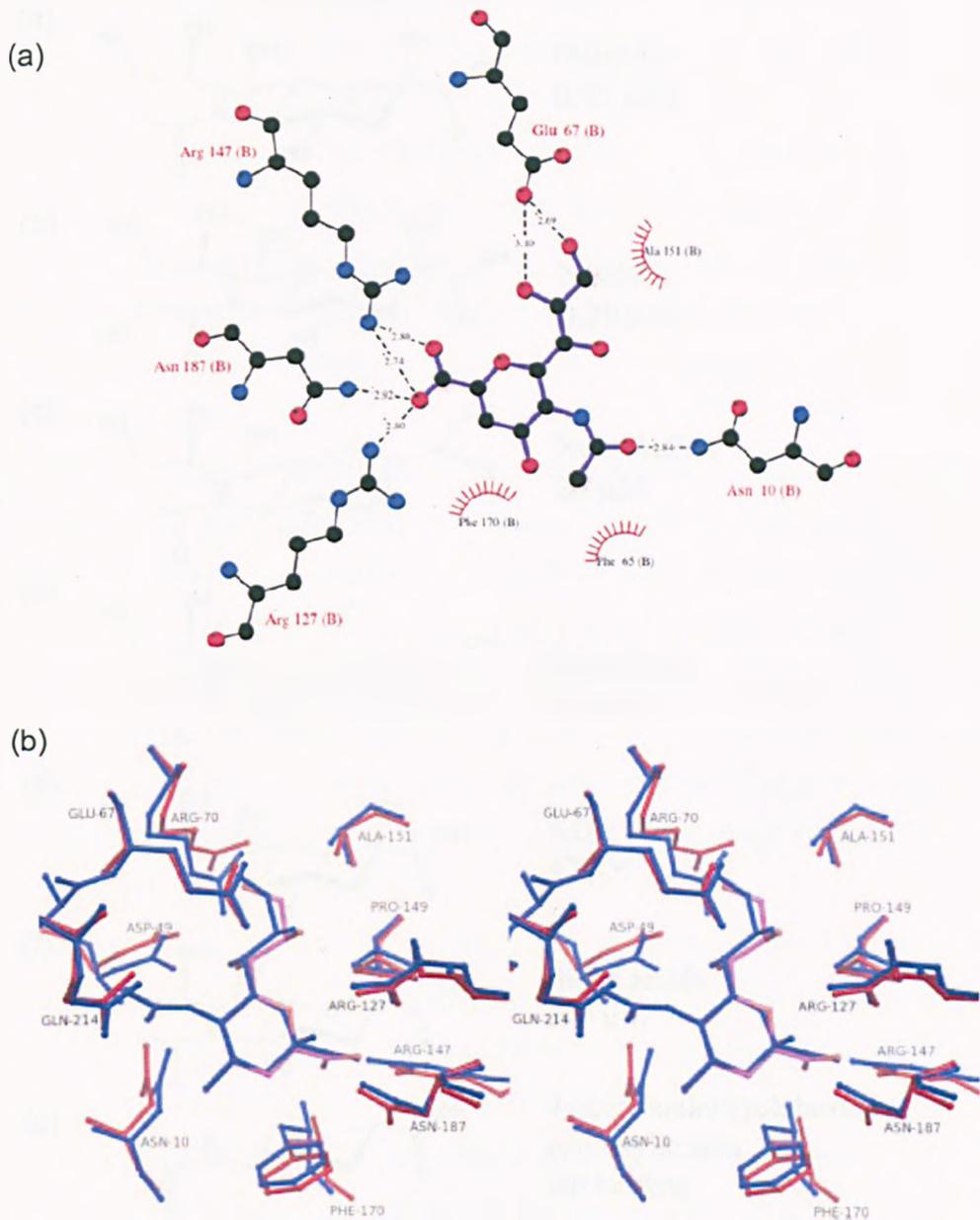
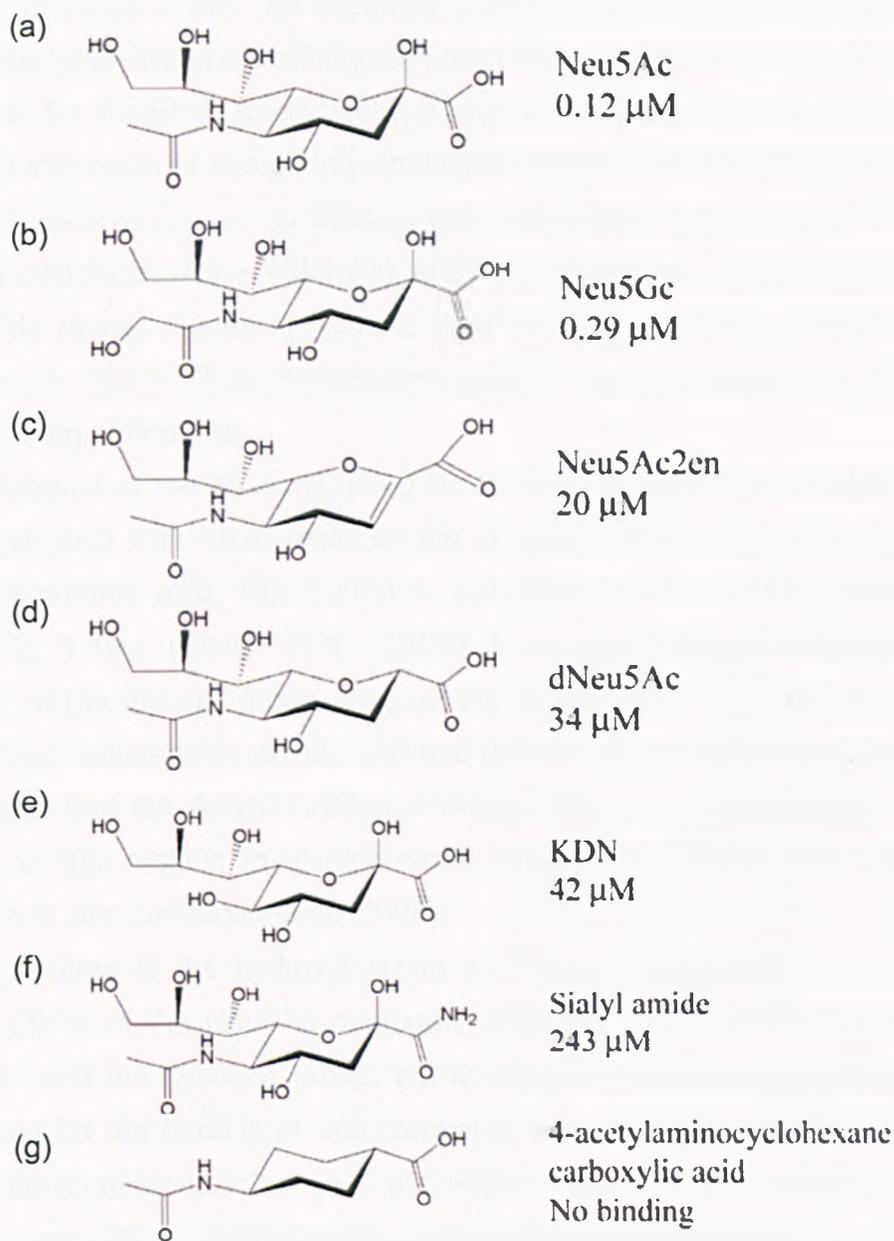


Figure 1.18 Binding site interactions between Neu5Ac2en and SiaP. (a) Interactions between binding site residues of SiaP and Neu5Ac2en. Figure taken from Muller, *et al* (2006). **(b)** Stereo diagram of SiaP binding site in the presence of Neu5Ac (blue) or Neu5Ac2en (red) showing almost identical binding site architecture. Figure taken from Johnston *et al* (2008).



Figures 1.19 Neu5Ac orthologues used in the biophysical analysis of SiaP binding determinants. Structures shown; *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), Neu5Ac2en, dNeu5Ac, 2-keto-3-deoxy-D-glycero-D-galactonononic acid (KDN), sialyl amide and 4-acetylamino-cyclohexane carboxylic acid. The binding affinity of each compound for SiaP is indicated on the right. Figure modified from Muller *et al* (2006).

analogue of Neu5Ac that has replaced the carboxylate with an amide group, is found to be 2000-fold lower affinity than the affinity for Neu5Ac (243 μM compared to 0.12 μM for Neu5Ac). Reciprocally, if Arg147 is mutated to an alanine or lysine then LPS sialylation is completely abolished indicative of no transporter function. The importance of Arg147 for binding was also shown by the inability of SiaP to bind *N*-acetylglucosamine (GlcNAc) and *N*-acetylmannosamine (ManNAc), two compounds closely related to Neu5Ac that do not possess the carboxylate group (Severi et al., 2005). This indicates that Arg147 is an important factor in the high affinity binding of Neu5Ac.

The importance of the *N*-acetyl group for Neu5Ac binding to SiaP was assessed using analogues that either replaced the *N*-acetyl group for a glycolyl group (*N*-glycolylneuraminic acid, Fig. 1.19b) or completely removed the *N*-acetyl group (KDN) (Fig. 1.19 e (Severi et al., 2005)). Biophysical analysis revealed that the presence of the glycolyl group reduced the affinity 2-fold and removal of the *N*-acetyl group reduced the affinity 350-fold (Severi et al., 2005). The structures of SiaP reveal that the Asn10 residue interacts with the *N*-acetyl group. However, mutation of this residue to alanine did not reduce the observable level of LPS sialylation *in vivo* (Johnston et al., 2008).

The contribution of the hydroxyl group at C2 was assessed by measuring the binding affinity of the Neu5Ac analogue, dNeu5Ac (Fig. 1.19d). This compound does not have the hydroxyl group, but is otherwise identical to Neu5Ac. The K_d determined for the binding of this compound was 34 μM , which is similar to the binding affinity of Neu5Ac2en (~ 20 μM , Fig. 1.19c (Severi et al., 2005)) indicating that the lower affinity of Neu5Ac2en is due to the absence of the C2 hydroxyl group and not the partial flattening of the sugar ring. The C2 hydroxyl group was shown to interact with Arg127 and Asn187 in the SiaP structure with Neu5Ac bound (Johnston et al., 2008). If Asn187 is mutated to a glutamine then LPS sialylation still occurs at wildtype levels. If Arg127 is mutated, then LPS sialylation is completely abolished. However, Arg127 also interacts with the carboxylate group so is likely to be a more important residue than Asn187. The binding of Neu5Ac to the SiaP R127A mutant was also assessed using ITC and no binding was detectable, reinforcing the importance of this residue in the binding mechanism (Johnston et al., 2008).

To determine whether the carboxylate and amino groups of Neu5Ac alone were sufficient for binding to SiaP, the binding affinity of the compound, 4-acetylaminocyclohexane carboxylic acid was assessed using tyrosine fluorescence spectroscopy (Muller et al., 2006). This molecule is a 6 carbon ring with the amino and carboxylate groups in analogous positions to Neu5Ac, but with all other Neu5Ac functional groups missing (Fig. 1.19g). There was no measurable binding indicating that the carboxylate group and the amino group are essential but not sufficient for high affinity binding.

1.10 Other ESRs working with TRAP integral membrane components

1.10.1 TAXI-type TRAP transporters

During phylogenetic analysis of the TRAP transporters, a family of ESRs with no sequence similarity to the DctP family of ESR has been found to be associated with the TRAP transporter integral membrane components (Fig. 1.20a (Kelly & Thomas, 2001, Rabus et al., 1999)). The first example of these proteins was a cell surface component of *Brucella abortus* and many, if not all genes encoding these proteins are annotated as 'immunogenic proteins' (Mayfield et al., 1988). These ESRs are associated with the membrane components of TRAP transporter, and with no DctP homologue also associated, these immunogenic proteins are the obvious candidate for the ESR role. They also have all the characteristics of ESRs; hydrophilic proteins, 300-400 amino acids long and possessing an N-terminal signal sequence (Kelly & Thomas, 2001). With no functional information available on these proteins, they were termed TRAP-associated extracytoplasmic immunogenic proteins (TAXI). Phylogenetic analysis has revealed that they form a discrete family of ESRs and analysis of the genome context of TAXI genes has shown them to be associated only with *dctQM* fusion genes (Fig. 1.20a (Kelly & Thomas, 2001)). Another curious observation is that TAXI-TRAP transporters are the only type of TRAP transporter archaea possess, however, TAXI-TRAP and DctP-TRAP transporters are widespread in bacteria possibly indicating that the TAXI-TRAP ESRs are the older of the two (Kelly & Thomas, 2001).

The only published functional data on TAXI-type ESRs is from a high-throughput structural genomics project in which the high resolution structure of a TAXI ESR

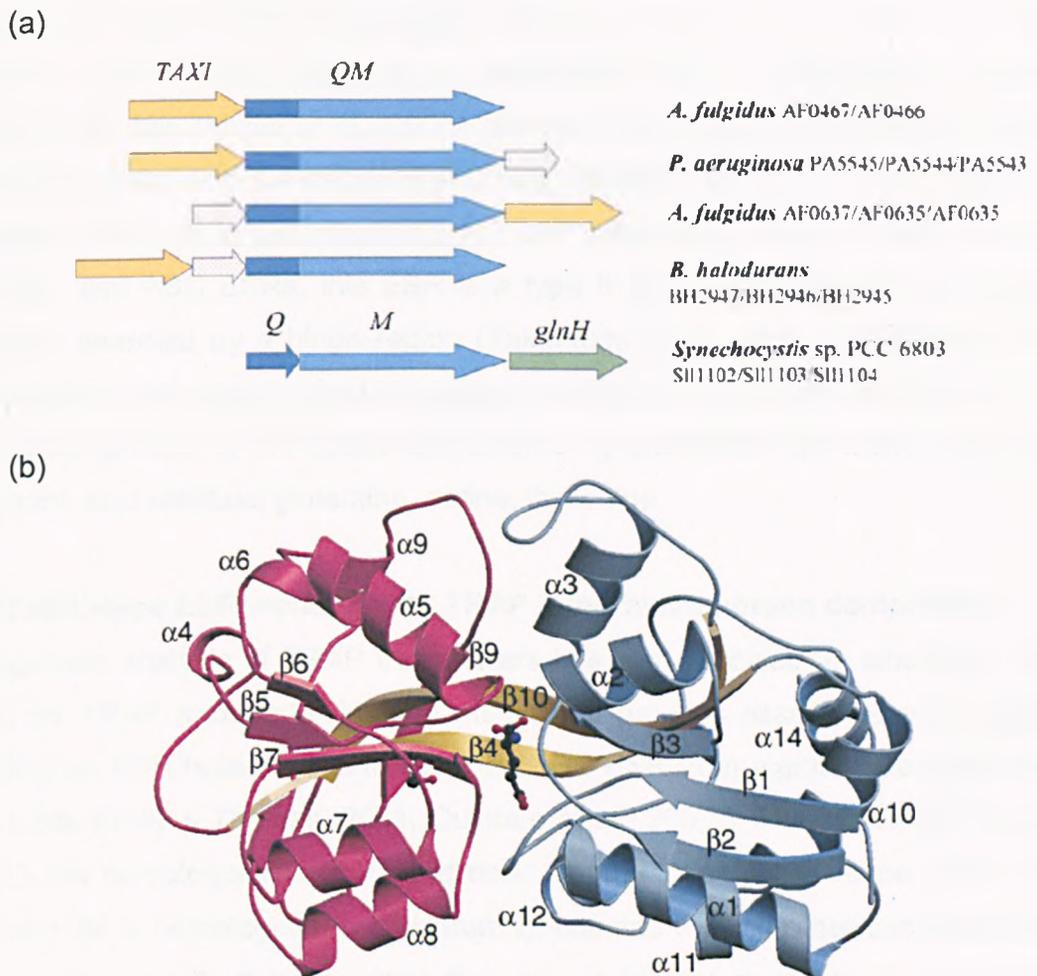


Figure 1.20 Other ESRs interacting with TRAP integral membrane components. (a) Organisation of genes encoding TRAP transporters associated with TAXI-type ESRs. DctQ and DctM homologues shaded in dark and light blue respectively. Genes encoding TAXI-type ESRs shaded in yellow. TRAP-linked genes (*tlg*, white with spots) encode proteins of unknown function. The ABC-type ESR, GlnH, shaded in green. Examples of the genetic organizations shown are indicated. Figure taken from Kelly & Thomas (2001). (b) Crystal structure of TAXI-type ESR bound to L-glutamate or L-glutamine. Figure taken from Takahashi *et al* (2004).

was solved, however, the authors did not identify this protein as such (Fig. 1.20b (Takahashi *et al.*, 2004). The crystal structure solved is of TTHA1157 from *Thermus thermophilus* HB8 and is associated with a DctQM-type integral membrane protein (Muller *et al.*, 2006). Structure was solved in the presence and absence of ligand and the bound ligand was identified as either L-glutamate or L-glutamine, which fits in with other known TRAP substrates as it is an organic anion. Like SiaP and ABC ESRs, this ESR is a type II ESR composed of two globular domains separated by a hinge region (Takahashi *et al.*, 2004). Interestingly, the architecture of the binding site is completely different to that seen in SiaP with the carboxylate groups of the ligand coordinated by interaction via water molecules and amino acid residues glutamine, serine, threonine.

1.10.2 ABC-type ESR working with TRAP integral membrane components

Phylogenetic analysis of TRAP transporters has also revealed an example of the genes for TRAP transporter integral membrane proteins associated with a gene encoding an ESR homologous to GlnH, an ABC ESR from a glutamate transporter (Fig. 1.20a (Kelly & Thomas, 2001, Quintero *et al.*, 2001)). The genes *sll1102* and *sll1103* are homologous to *dctQ* and *dctM* from *R. capsulatus*, respectively. The gene *sll1104* is homologous to *glnH* from *E. coli* and with no other candidate ESR genes in the vicinity it is assumed they are co-regulated and form a functional transporter (Kelly & Thomas, 2001). When the genes for the integral membrane components (*sll1102* and *sll1103*) or the putative ESR (*sll1104*) are disrupted separately, Na⁺-dependent transport of glutamate decreases to 70-80% of wildtype. This indicates that all three genes encode proteins involved glutamate transport. The background (the remaining 70-80%) glutamate transport is caused by a Na⁺-dependent secondary transporter homologous to GltS from *E. coli* (Quintero *et al.*, 2001).

1.11 Other notable TRAP transporters

1.11.1 A dimeric DctP-type ESR

The gene *takP* encodes a DctP-type (20% identity with SiaP) ESR from *R. sphaeroides* (Gonin *et al.*, 2007). TakP and an orthologue from *R. capsulatus*

have both been characterised using tryptophan fluorescence spectroscopy and found to be specific for the α -keto acids, pyruvate, oxobutyrate and oxoalate (Gonin et al., 2007, Thomas et al., 2006). The structure of TakP was solved and was found to be a homodimer with intricate helix swapping between the protomers ruling out crystal contacts as the source of dimerisation (Fig. 1.21a and b (Gonin et al., 2007)). Dimer formation was confirmed in solution using analytical size exclusion chromatography (SEC) and protein cross-linking (Gonin et al., 2007). The protomers are in a back-to-back configuration with a number of hydrogen bonds and salt bridges stabilizing the complex. The 55 Å long kinked helices that are swapped from one protomer to another are conserved in 100 homologues of TakP, including RRC01191 from *R. capsulatus*, but not DctP from *R. capsulatus* (Mulligan and Thomas, unpublished data). The structure was solved in complex with sodium pyruvate indicating that sodium plays an important role in binding (Fig. 1.21c).

It is currently unknown whether this dimeric configuration is a common feature of TRAP ESRs homologous to TakP, if there are any advantages to a dimeric ESR, or indeed how a dimeric ESR would interact with the membrane components.

1.11.2 A glutamate specific ESR-dependent secondary transporter from *R. sphaeroides*

A strain of *R. sphaeroides* 2.4.1 was isolated that was unable to grow on L-glutamate as the sole carbon and nitrogen source, however, upon addition of Na⁺ growth was restored (Jacobs et al., 1996). *In vivo* transport assays revealed high affinity transport of glutamate that was Na⁺-stimulated, however, osmotic shock treatment abolished transport indicating the presence of an ESR. Experiments also revealed that this transporter was inhibited by ionophores, valinomycin and nigericin, but insensitive to orthovanadate. The ESR was purified from the periplasmic fraction and was shown to bind glutamate with a high affinity (1.2 μM) and the binding of glutamate was not stimulated by the addition of Na⁺ ions (Jacobs et al., 1996).

Vesicles from *R. sphaeroides* were isolated and transport was characterised *in vitro*. Uptake of radiolabelled glutamate was only observed in the presence of high concentration of Na⁺ ions and the purified ESR and uptake activity was reduced in the presence of ionophores and uncouplers. Radiolabelled glutamate taken into the

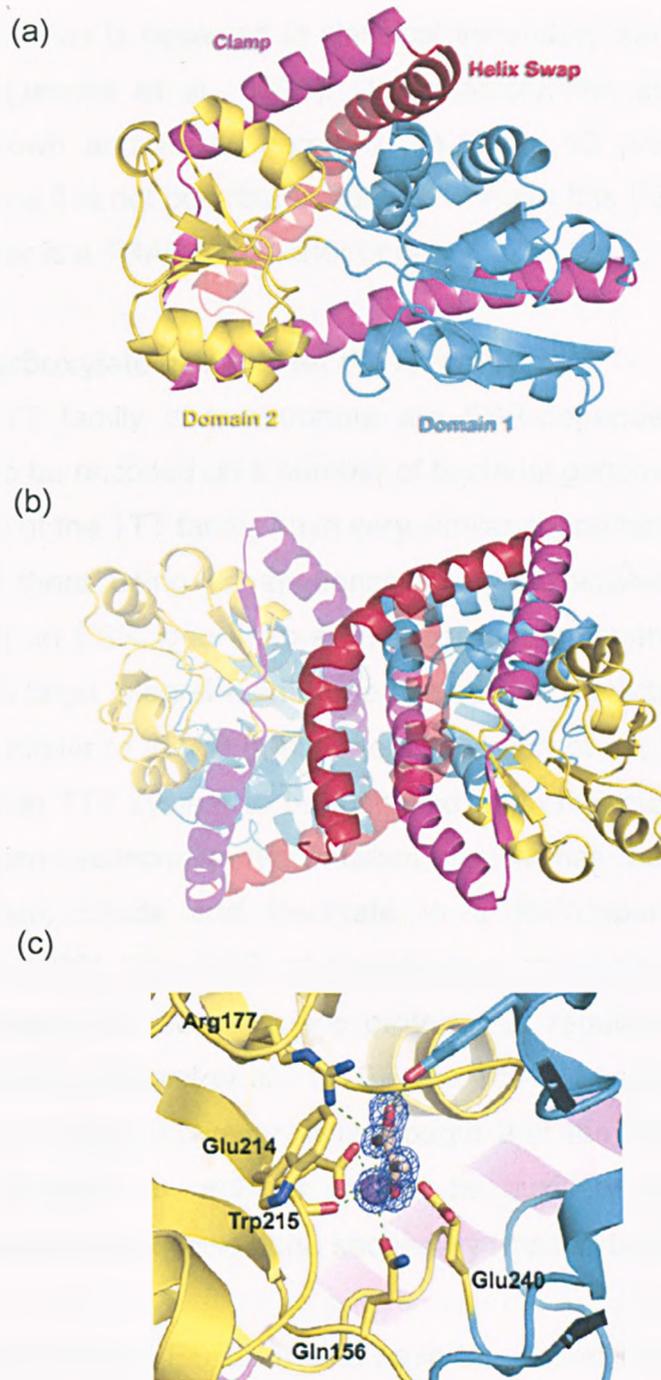


Figure 1.21 Crystal structure of dimeric TRAP ESR in complex with sodium pyruvate. (a) TakP monomer colour-coded to differentiate the 2 globular domains (yellow and blue) and the dimer forming helix (red). (b) Dimeric form of TakP. The same colour scheme is used as in (a). (c) TakP binding site showing presence of pyruvate and sodium (purple sphere). Figures taken from Gonin *et al* (2007).

vesicles was not released (exchanged) upon addition of 500-fold excess non-radiolabelled substrate as is observed in classical secondary transporters, like lactose permease (Jacobs et al., 1996). Unfortunately, the identity of this transporter is unknown and *R. sphaeroides* 2.4.1 has 13 predicted TRAP transporters, therefore it is not possible to confirm whether this ESR-dependent secondary transporter is a TRAP transporter or not.

1.12 Tripartite tricarboxylate transporter (TTT) family

Members of the TTT family of transporters are ESR-dependent secondary transporters found to be encoded on a number of bacterial genomes (Winnen et al., 2003). Members of the TTT family have very similar characteristics to TRAP transporter despite there being no sequence similarity between them. TTT systems consists of an ESR, one integral membrane component predicted to have 4 TMHs and a large integral membrane component predicted to have 12 THMs, remarkably similar to TRAP transporters (Forward et al., 1997, Winnen et al., 2003). The first TTT system to be identified and characterised was the TctABC system from *Salmonella typhimurium*, which has been shown to transport fluorocitrate, citrate and isocitrate in a Na⁺-dependent manner (Widenhorn et al., 1988) The ESR of this system has been purified and characterised revealing it to be a dimeric protein that requires Na⁺ ions for optimal substrate binding (Sweet et al., 1979). The TctABC-type systems have only been found in bacteria, however, it is thought that the large membrane component, TctA is found in archaea without its auxiliary proteins, which suggests it is functioning as a standalone secondary transporter (Winnen et al., 2003).

79 genes encoding predicted ESR proteins have been identified in *Bordetella pertussis*. The majority of these so-called *Bordetella* uptake genes (*Bug*) are orphan ESRs and not encoded with any translocation machinery (Antoine et al., 2003). These *Bug* genes have also been identified in a number of other bacteria, but the only other bacteria to have as many *Bug* genes as *Bordetella pertussis* are *Bordetella parapertussis*, *Bordetella bronchiseptica* and *Ralstonia metallidurans*, suggesting that this enrichment is lineage-specific (Antoine et al., 2003). TctC has been found to be orthologous to these *Bug* genes (Antoine et al., 2003). High resolution structures of two *Bug* proteins, BugD and BugE, from *B. pertussis* have been elucidated (Huvent et al., 2006b, Huvent et al., 2006a).

Both ESRs were crystallised with bound ligand. BugD binds aspartate and BugE binds glutamate (or glutamine), however, the overall binding site architecture is very similar with both proteins utilising two H₂O molecules to coordinate the ligand.

1.13 Expression and purification of bacterial integral membrane proteins

It is predicted that 5-10% of bacterial genomes and 3% of the human genome are dedicated to encoding transport proteins (Paulsen *et al.*, 2000, Paulsen *et al.*, 1998, Venter *et al.*, 2001). However, there is a dearth of structural and functional information on integral membrane proteins when compared to soluble proteins. To study transport proteins biochemically and biophysically, large amounts of purified protein are required, which is the major bottleneck for this work.

There are a number of difficulties involved in the production of integral membrane protein that makes them challenging to work with. Generally, integral membrane transporters make up a very low proportion of the membrane proteome (<0.1% in most cases). This means harvesting transporters at their natural abundance would be incredibly laborious and often unfruitful, thus necessitating the development of overexpression systems for transport proteins. However, overexpressing integral membrane proteins brings its own problems. Integral membrane proteins are inherently insoluble and are often found to be unstable in solution. They also need to be correctly inserted into the membrane to function properly, which is a process reliant on a host of other proteins. It has been observed in a number of instances that integral membrane protein expression leads to growth perturbation and/or cell death (Miroux & Walker, 1996). A proteomic study of the consequences of integral membrane overexpression revealed that growth perturbation was caused, in part, by the saturation of the cytoplasmic membrane protein translocation machinery (Wagner *et al.*, 2007). Saturation of the translocation machinery leads to build up of cytoplasmic aggregates that include the overexpressed proteins themselves, chaperones, proteases and precursors of periplasmic and outer membrane proteins. There was also a significant decrease in the levels of respiratory chain complexes in the membrane and activation of less efficient forms of ATP production, which would decrease cell processes significantly (Wagner *et al.*, 2007). A further problem with the production of integral

membrane proteins is that once the protein has been successfully inserted into the membrane, it must then be extracted using detergents for purification. These processes can often lead to inactivation and aggregation of the transporters. To overcome these problems a number of key variables can be altered, two of the most important are; the expression strain and the expression vector used. However, the overriding message from the literature is that expression of integral membrane proteins is very subjective. It is therefore often more fruitful to use a number of approaches simultaneously in order to attain successful overexpression.

1.13.1 Choice of expression strain

For expression of bacterial integral membrane transporters, *E. coli* is the expression strain of choice due to its low cost, high growth rate, well characterised genetics and plethora of available cloning vectors. Probably the most common *E. coli* expression host is BL21 (DE3) (Novagen). This strain is for use with genes expressed from a T7 RNA polymerase based system such as the pET series of vectors from Novagen. The (DE3) designation denotes the presence of a chromosomal copy of the T7 RNA polymerase gene under the control of the isopropyl- β -D-thiogalactopyranoside (IPTG)-inducible *lacUV5* promoter (Studier *et al.*, 1990). BL21 (DE3) is deficient in the proteases *lon* and *ompT*, to prevent interference with overexpressed protein accumulation. C41 (DE3) and C43 (DE3) are strains of *E. coli* derived from BL21 (DE3) (Miroux & Walker, 1996). These strains were derived by selecting for BL21 (DE3) mutants with increased resistance of expression-induced toxicity. The exact mutations in C41 (DE3) and C43 (DE3) that allow it to overcome the toxic effects of integral membrane protein expression are not known. However, it is believed that the mutations help to prevent the uncoupling of transcription and translation which is believed to contribute to these toxic effects. MC1061 is another strain of *E. coli* that is used for protein expression. MC1061 is used primarily with the arabinose-inducible pBAD vectors because it has the region of its genome that encodes the catabolic genes for arabinose deleted (Casadaban & Cohen, 1980). This allows persistence of arabinose in the cell and continued protein induction.

Aside from *E. coli*, there are a number of expression hosts that have been used to express many proteins with different properties. However, the other

expression host most pertinent to study is *Lactococcus lactis*. *L. lactis* is a Gram positive, aerotolerant, lactic acid bacterium that is reported to possess a number of characteristics that make it ideal for membrane protein expression (Kunji *et al.*, 2003). *L. lactis* is able to grow rapidly and requires no aeration for growth. It has tightly controlled promoter systems making expression of toxic proteins possible. *L. lactis* possesses a small genome and has only one membrane making membrane protein purification easier. There are now a number of studies available on the expression of eukaryotic membrane proteins, which are particularly difficult to express in bacteria (Kunji *et al.*, 2005, Kunji *et al.*, 2003, Monne *et al.*, 2005). The most commonly used *L. lactis* expression strain is NZ9000 (Kunji *et al.*, 2003). This strain is a derivative of MG1363 and has the genes encoding NisR and NisK integrated in the chromosome for use with the nisin A inducible expression vectors (see below).

1.13.2 Choice of expression vector

The choice of expression vector is another major consideration when expressing integral membrane proteins. There are a number of desirable traits an expression vector should have for the expression of membrane proteins. Expression should be tightly regulated, which prevents leaky expression that can lead to proteolysis and cell death. The promoter system should not be too strong to prevent the formation of inclusion bodies. Ideally, the promoter should be tightly controlled so that expression can be modulated.

The series of pET vectors offers a variety of options (>30 variants) and has been used to express a number of membrane proteins, however, they are not ideal candidates. Expression from pET vectors is controlled by the T7 promoter, which is not recognised by the *E. coli* RNA polymerase, so is in theory tightly regulated. However, the T7 promoter is very strong, which has been shown to lead to cell death (Miroux & Walker, 1996). The arabinose inducible pBAD vectors are a better choice as expression is highly repressed in the absence of inducer and induction can be varied linearly between 0.00002% and 0.2% arabinose (v/v) (Guzman *et al.*, 1995). The pTTQ18 vector is under the control of the moderately strong *tac* promoter (Stark, 1987). Expression is not tightly repressed as leaky expression has been observed when cells enter stationary phase (Ward, 2000). However, this vector has been extremely efficacious in the expression of a number of integral membrane proteins (Liang *et al.*, 2005,

Morrison *et al.*, 2003, Saidijam *et al.*, 2005, Saidijam *et al.*, 2003). The expression vectors most commonly used in *L. lactis* are pNZ8048 variants. Expression from pNZ8048 is under the control of *nisA* promoter and is induced by addition of nisin A to the medium. Alternatively, it has been demonstrated that expression can be induced by addition of the culture supernatant of NZ9700, which is a strain of nisin A-producing *L. lactis* (Kuipers *et al.*, 1993). The nisin expression system is composed of a sensor kinase, NisK and the response regulator, NisR, which are both integrated into the chromosome of strain NZ9000 (Kuipers *et al.*, 1993). Expression from pNZ8048 is tightly regulated controllable between 100- 10000-fold dilutions of NZ9700 culture supernatants making it a good alternative for integral membrane protein expression (Kunji *et al.*, 2003).

Aims of this investigation

The aim of this study was to investigate the structure and mechanism of TRAP transporters using biochemical techniques.

The first priority was to develop an overexpression system for the production of the integral membrane proteins from the *Escherichia coli* TRAP transporter, YiaMNO and the *Haemophilus influenzae* TRAP transporter, SiaPQM. It was hoped that at least one of these transporters would be reconstituted into liposomes and its transport mechanism characterised using *in vitro* transport assays. It was also hoped that these integral membrane proteins would be characterised *in vitro* using other biochemical and biophysical techniques in order to answer some fundamental questions regarding these transporters.

Chapter 2

Materials and methods

2.1 Suppliers

All chemicals, reagents and media were purchased from New England Biolabs (NEB), Sigma-aldrich, GE Healthcare, Oxoid, Promega, Pierce, Merck, Avestin, Applied Biosystems, Bio-rad, Bioline, Fermentas, Invitrogen, Qiagen and Avanti Polar Lipids.

2.2 Table of primers

See table 2.1 for details of the oligonucleotides used during PCR for the amplification of DNA fragments.

2.3 Strains and plasmids

See table 2.2 and 2.3 for details of the plasmids and bacterial strains used.

2.4 Media and antibiotics

2.4.1 M9 minimal medium

M9 minimal salts contained 6 g Na_2HPO_4 , 3 g K_2HPO_4 , 0.5 g NaCl and 1 g NH_4Cl per litre. For growth of *E. coli*, these salts were supplemented with 0.2% glucose, 1 mM MgSO_4 and antibiotics, if required. The growth of *E. coli* strain MC1061 in M9 required the addition of the following; 40 mg/ml L-leucine, 40 mg/ml L-isoleucine, 50 mg/ml 4-methyl-5-(β -hydroxyethyl) thiazole (THZ), 2 mg/ml thiamine, 50 mg/ml nicotinic acid 100 mM FeSO_4 .

2.4.2 Enhanced M9 minimal medium

Enhanced M9 minimal medium was composed of the same components as M9 minimal medium except all the weights, percentages and concentrations were doubled. In addition, the following supplements were added; 25 $\mu\text{g/ml}$ $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 1 $\mu\text{g/ml}$ riboflavin, 1 $\mu\text{g/ml}$ nicotinamide and 1 $\mu\text{g/ml}$ thiamine.

2.4.3 Luria-Bertani (LB) medium

LB was composed of 10 g tryptone, 5 g yeast extract and 10 g NaCl per litre.

2.4.4 M17 medium

M17 medium was made by dissolving 37.25 g of M17 (Oxoid) into 1 L dH_2O and autoclaving. In addition, 0.5% glucose was added for growth.

Name of Primer	Sequence of primer (5'→ 3')	Restriction site
YiaMF	CCGGAATTTCGCATATGAAAAAATACTCGA ACGAATACTGG	EcoRI
YiaMR	AAA <u>ACTGCAGCAGCTCCTT</u> GCGGTGGCGAC GTTAG	PstI
YiaNF	CCGGAATTTCGCATATGGCTGTGCTGATTTT TCTGGGCTG	EcoRI
YiaNhisR	AAA <u>ACTGCAGCATT</u> AATCCATTTCAAAGGG AGGATG	PstI
SiaQMpTTQHisF	CCGGAATTTCGCATATGAAATATATTAATAA GCTTGAGG	EcoRI
SiaQMpTTQHisR	AAA <u>ACTGCAGCTGGT</u> TATCAATAGATTTGGC ACAAATGTG	PstI
SiaP_Nfor_minSP	ATGGTGAGAATTTATATTTTCAAGGTGCTG ATTATGACTTAAAATTCGG	LIC
SiaP_Cfor	ATGGGTGGTGGATTTGCTATGAAATTGACA AAACTTTTCCTTGCC	LIC
SiaP_Nrev	TGGGAGGGTGGGATTTTCATTATGGATTAA TTGCTTCAATTTGTTTTAA	LIC
SiaP_Crev	TTGGAAGTATAAATTTTCTGGATTAATTGC TTCAATTTGTTTTAA	LIC
SiaQM_Nfor	ATGGTGAGAATTTATATTTTCAAGGTAAGT ATATTAATAAGCTTGAGGAA	LIC
SiaQM_Cfor	ATGGGTGGTGGATTTGCTAAATATATTAAT AAGCTTGAGGAATGG	LIC
SiaQM_Nrev	TGGGAGGGTGGGATTTTCATTATGGTGGTA TCAATAGATTTGGCACAAATGT	LIC
SiaQM_Crev	TTGGAAGTATAAATTTTCTGGTATCAATAG ATTTGGCACAAATGT	LIC

Table 2.1 Details of oligonucleotide primers used for amplification of DNA fragments.

Plasmid name	Description	
pTTQ18	<i>E. coli</i> expression vector	Amp ^r
pBADnLIC2005	pBAD vector (Invitrogen) modified with LIC cassette	Amp ^r
pBADcLIC2005	pBAD vector (Invitrogen) modified with LIC cassette	Amp ^r
pNZ8048	<i>L. lactis</i> expression vector	Chl ^r
pERL	<i>L. lactis</i> vector containing the pNZ8048 origin of replication	Eryth ^r
pRY11	pET20b containing gene VC1779	Amp ^r
pES7	pWSK30 vector containing <i>E. coli nanT</i>	Amp ^r
pCM14	<i>yiaM</i> in pTTQ18 with C-terminal RGSH ₆	Amp ^r
pCM19	<i>YiaN</i> in pTTQ18 with C-terminal RGSH ₆	Amp ^r
pCM20	<i>yiaMN</i> in pTTQ18 with C-terminal RGSH ₆ <i>yiaN</i> (G368D) mutant	Amp ^r
pCM22	<i>yiaMN</i> in pTTQ18 with C-terminal RGSH ₆	Amp ^r
pCM23	<i>siaQM</i> in pTTQ18 with C-terminal RGSH ₆ <i>siaQM</i> (D322Y) mutant	Amp ^r
pCM23.1	<i>siaQM</i> in pTTQ18 with C-terminal RGSH ₆	Amp ^r
pBADnSiaQM	<i>siaQM</i> in pBAD with N-terminal H ₁₀	Amp ^r
pBADcSiaQM	<i>siaQM</i> in pBAD with C-terminal H ₁₀	Amp ^r
pBADnSiaP	<i>siaQM</i> in pBAD with N-terminal H ₁₀	Amp ^r
pBADcSiaP	<i>siaQM</i> in pBAD with C-terminal H ₁₀	Amp ^r
pNZnSiaQM	<i>siaQM</i> in pNZ9000 with N-terminal H ₁₀	Chl ^r
pNZcSiaQM	<i>siaQM</i> in pNZ9000 with C-terminal H ₁₀	Chl ^r

Table 2.2 Details of plasmids created or used during this study.

Bacterial strain	Bacterial species	Genotype	Source
BL21 (DE3)	<i>E. coli</i>	F- <i>ompT hsdSB</i> (rB-, mB-) <i>gal dcm</i> (DE3)	Novagen
MC1061	<i>E. coli</i>	<i>araD139 Δ(ara-leu)7696 ΔlacX74 galU galK hsdR2</i> (r _K - m _K +) <i>mcrB1 rpsL</i> (F-)	(Casadaban & Cohen, 1980)
DH5α	<i>E. coli</i>	K-12 F' ϕ 80 <i>dlacZΔM15 recA1 endA1 gyrA26 thi-1 hsdR17 supE44 relA1 deoR Δ(lacZYA-argF)U169]</i>	Invitrogen
NZ9000	<i>L. lactis</i>		(Kuipers et al., 1993)
NZ9700	<i>L. lactis</i>		(Kuipers et al., 1993)
BW25113Δ <i>nanT</i>	<i>E. coli</i>	<i>Δ(araD-araB)567, ΔlacZ4787(::rrnB-3), lambda⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514, ΔnanT</i>	(Baba et al., 2006)

Table 2.3 Details of bacterial strains used during this study.

2.4.5 Solid media

Solid media were made by mixing liquid media with 1.5% (w/v) agar.

2.4.6 Antibiotics

Antibiotic stock solutions were made and filter sterilized. The antibiotics were added to liquid or solid media such that the final concentrations were; 100 µg/ml ampicillin, 50 µg/ml kanamycin and 30 µg/ml chloramphenicol.

2.5 Buffers and solutions

2.5.1 SDS-PAGE

10x running buffer was composed of 30 g/L Tris, 140g/L glycine and 10 g/L SDS.

Coomassie Brilliant Blue (CBB) dye consisted of 45% methanol, 10% acetic acid and 0.25% Coomassie brilliant blue dye.

2.5.2 Western blotting

TBST (10x) was composed of 24.2 g/L Tris, 80 g/L NaCl and 10 ml/L Tween 20 and adjusted to pH 7.6 with concentrated HCl.

Transfer buffer (1x) was composed of 14.41 g/L glycine, 100 ml/L methanol and 25 mM Tris, pH 8.3.

PBST (1x) was composed of 1.088 g/L KH_2PO_4 , 4.544 g/L Na_2HPO_4 , 13.444 g/L NaCl and 1% Tween 20.

TBSTT (1x) was composed of 20 mM Tris pH 7.5, 500 mM NaCl, 0.05% (v/v) Tween 20 and 0.2% Triton X-100.

2.5.3 Agarose gel electrophoresis

TBE (10x) was composed of 108 g/L Tris, 55 g/L orthoboric acid and 9.3 g/L sodium EDTA (ethylenediaminetetraacetic acid). This solution was diluted 10-fold to produce a working solution for agarose gel electrophoresis of DNA samples.

Agarose gels were constructed by dissolving 1% (w/v) agarose into 1x TBE buffer.

Sample buffer consisted of 10% (v/v) glycerol, 0.025% (w/v) bromophenol blue in 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA.

2.6 General cloning techniques

2.6.1 Agarose gel electrophoresis

1 % (w/v) agarose gels were made by pouring molten agarose directly into a horizontal electrophoresis tank, adding combs to produce appropriately sized wells and allowed to set. Solid agarose gels were submerged in TBE buffer (1x).

Samples were prepared by mixing 8 µl DNA sample with 5x stock of sample buffer and loading into a well. 50µl of 10 mg/ml ethidium bromide solution was added to the TBE buffer at the anode end of the electrophoresis tank and electrophoresis was performed for ~45 minutes at a constant 80 V. DNA was visualized under UV light using a transilluminator (Syngene).

2.6.2 Polymerase chain reaction (PCR)

PCR was used for the amplification of genes for cloning. High fidelity DNA polymerases were used throughout, for example, KOD Hotstart (Novagen). A standard PCR reaction consisted of KOD PCR buffer (1x), 0.2 mM each dNTPs, 1 mM MgSO₄, 0.5 µM 3'- and 5'-primers, 1 unit KOD polymerase and either genomic or plasmid DNA template. A thermocycler (Techne) was used and the standard program was initial denaturation at 94°C for 5 minutes followed 30 cycles of 94°C for 30 seconds, 60°C for 60 seconds and 72°C for 6 minutes.

2.6.3 Preparation of DNA fragments for cloning

2.6.3.1 Preparation of plasmid DNA

Plasmid DNA was prepared from bacterial strains using a Miniprep kit (Qiagen) that employs the alkaline lysis method for plasmid recovery. The manufacturers recommended protocol was used.

2.6.3.2 Gel extraction

This method was used when the target DNA fragment was contaminated with other DNA. The commercially available QIAquick gel extraction kit (Qiagen) was used. The manufacturers recommended protocol was used.

2.6.3.3 PCR clean-up

This method was used to remove unwanted, non-DNA contaminants, such as PCR reaction components, from a pure DNA fragment. The commercially available QIAquick PCR clean-up kit (Qiagen) was used. The manufacturers recommended protocol was used.

2.6.4 Ligation-dependent cloning

This cloning method consisted of treating vector DNA ("vector") and the DNA fragment to be inserted ("insert") with restriction enzymes and then ligating the insert and vector together. A standard 35 μ l restriction enzyme treatment consisted of 15 μ l DNA (either insert or vector), 1 μ l of restriction enzyme(s) of choice, 3.5 μ l compatible 10x buffer and made up to 35 μ l with dH₂O. The reaction is incubated for between 1 and 18 hours at 37°C. The DNA is then purified, either by gel extraction or PCR clean-up.

A standard 15 μ l ligation reaction consisted of an insert DNA to vector DNA molar ratio of at least 2:1, 1 μ l *E. coli* ligase (NEB) or T4 DNA ligase (Promega), Ligase buffer and made up to 15 μ l with dH₂O. The reaction was incubated at room temperature overnight.

2.6.5 Ligation-independent cloning (LIC)

This method was used in the cloning of SiaP and SiaQM into pBAD vectors; pBADnLIC2005 or pBADcLIC2005 (Geertsma & Poolman, 2007). These vectors are modified versions of the commercially available pBAD vectors with the inclusion of a LIC cassette at the site of cloning (modified by Dr. Eric Geertsma).

Primers were designed such that a dedicated LIC tail was added to the PCR product, which was designed to be complementary to the LIC cassette in pBADnLIC2005 and pBADcLIC2005 (see Methods section 2.1 for Table of Primers). The gene of interest was amplified using PCR and purified using gel extraction. The vector was prepared by digestion with *Swa*I (1 μ l/2.5 μ g DNA) for 3 hours at 25°C. The linearised vector was purified by gel extraction.

Both the insert and vector DNA were treated with T4 DNA polymerase (Promega). The 3'→5' exonuclease activity of T4 DNA polymerase produces long, complementary overhangs on the vector and the insert. To do this, 200 ng of linearised vector was made up to 10 μ l with dH₂O and treated with T4 DNA

polymerase in the presence of dCTP (1.5 μ l 25 mM dCTP, 3 μ l 5x buffer and 0.5 μ l T4 DNA polymerase). The reaction was incubated at room temperature for 30 minutes, after which T4 DNA polymerase was deactivated by incubation at 75°C for 20 minutes. A volume of insert DNA corresponding to a 1:1 molar ratio with vector was treated with T4 DNA polymerase but this time in the presence of dGTP. Both the "LIC-ready" insert and vector can be stored stably at 4°C for several months.

The LIC-ready components were mixed in a 1:3 ratio of vector to insert and 3 μ l of this mixture was transformed into competent cells using standard methods.

2.6.6 Vector backbone exchange (VBEx) into *L. lactis* pNZ8048 vector (Geertsma & Poolman, 2007)

The cloning of SiaQM into the *L. lactis* vector pNZ8048 was achieved using vector backbone exchange (VBEx). This method was developed to avoid DNA manipulations in *L. lactis*. An intermediate vector was developed by fusing an *E. coli* vector with the region of the pNZ8048 vector (*L. lactis* vector) relevant for cloning. All DNA manipulations, i.e. insertion of gene into the vector, could then take place in *E. coli*. When the desired construct was achieved the pNZ8048 region of the vector could be excised and ligated into the *L. lactis* vector, pERL, which can then be transformed into *L. lactis*.

2.6.7 Transformations

2.6.7.1 Making chemically competent *E. coli*

To chemically induce competence in *E. coli*, a 10 ml LB culture of the desired strain was incubated overnight. 1 ml of this culture was diluted 10-fold and allowed to grow to an optical density (OD₆₅₀) of 0.4-0.5 at which point the cells were kept on ice. The cells were separated into 1 ml aliquots in sterile 1.5 ml tubes. The cells were harvested by centrifugation at 5000 rpm for 10 minutes and then resuspended in 1 ml ice-cold 100 mM CaCl₂. The cells were incubated on ice for 1 hour and then harvested by centrifugation at 5000 rpm for 10 minutes. The aliquots of cells were resuspended in 200 μ l ice-cold 100 mM CaCl₂ each and incubated on ice for 20 minutes. The chemically competent cells could be used immediately or, with the inclusion of 20% glycerol in the final

resuspension, could be snap frozen with liquid nitrogen or dry-ice/ethanol and stored at -80°C for future transformations.

2.6.7.2 Transformation of chemically competent cells: freeze-thaw method

This transformation method was used solely for the movement of plasmid DNA into a desired strain and not for transforming competent cells with a ligation reaction. A 5 ml culture of the desired strain was grown overnight in LB and then refreshed in the morning by diluting 0.5 ml of the overnight culture with 4.5 ml of LB. This culture was then grown aerobically for a further hour at 37°C. The cells were harvested by centrifugation at 5000 rpm for 10 minutes and the supernatant was discarded. The cell pellet was resuspended in 600 µl ice-cold 50 mM CaCl₂ and divided into 6 x 100 µl aliquots. These aliquots were kept on ice until they were required. 1-2 µl of concentrated plasmid DNA was added to an aliquot of cells, mixed by vortexing and snap frozen using liquid nitrogen or a dry-ice/ethanol bath. Once the mixture was frozen (~1 minute) it was quickly thawed by incubation at 37°C for no longer than 2 minutes. 1 ml of LB was added to the mixture and then incubated for 1 hour at 37°C, after which, 100 µl of the cell suspension was spread on a selective LB agar plate. The plates were incubated overnight at 37°C.

2.6.7.3 Transformation of chemically competent cells: heat shock method

This method of transformation can be used for the transformation of the desired *E. coli* strain with both plasmid DNA or ligation products. This method required a stock of ready-made chemically competent cell (for details, see Methods section 2.6.7.1). An aliquot of chemically competent cells was allowed to thaw on ice. 1-2 µl of plasmid DNA or a volume of ligation reaction was added to the cells. The mixture was mixed by vortexing and incubated on ice for a further 30 minutes. The mixture was incubated for 1.5 minutes at 42°C, then had 1 ml of LB added and was incubated at 37°C for 1 hour. After 1 hour, 100 µl of the cell suspension was spread on a selective LB agar plate, the remaining cell suspension was harvested by centrifugation and resuspended in 100 µl LB and spread on a second selective LB agar plate. All plates were incubated at 37°C overnight.

2.7 Growth of bacteria

2.7.1 Initial trial for expression from pTTQ18-derived vectors in *E. coli*

2 x 250 ml flasks containing 50 ml LB (supplemented with 20 mM glycerol and 100 µg/ml of ampicillin) were inoculated to an optical density (OD₆₅₀) of 0.1 from a 5 ml overnight culture of the particular *E. coli* expression strain. The cultures were grown to mid-log phase (OD₆₅₀ of 0.5) at 37°C with aeration, at which point one flask was induced with 0.5 mM IPTG and the other was left uninduced. The cultures were grown for a further 3 hours before being harvested by centrifugation.

2.7.2 Initial trial for expression from pBAD-derived vectors in *E. coli*

2 x 6 ml LB media (supplemented with 0.5% glycerol and 100 µg/ml ampicillin) were inoculated with 60 µl (1% inoculum) of a 2 ml overnight culture of the particular *E. coli* expression strain. The culture was grown aerobically to mid-log phase (OD₆₅₀ of 0.5) at 37°C, at which point expression was induced by the addition of 0.005% L-arabinose. The culture is grown for a further 2 hours before being harvested by centrifugation at 5000 rpm for 15 minutes.

2.7.3 Initial trial for expression from pNZ8048-derived vectors in *L. lactis*

10 ml of (M17 supplemented with 1% glucose and 50 µg/ml chloramphenicol) was inoculated with 200 µl (2% inoculum) of a 2 ml overnight culture of the particular *L. lactis* expression strain. The culture was grown anaerobically to mid-log phase (OD₆₀₀ of 0.5) at 30°C, at which point expression was induced by the addition of 1/5000 dilution of culture supernatant from the nisin A producing strain of *L. lactis* NZ9700. The culture was grown for a further 2 hours before being harvested by centrifugation at 5000 rpm for 15 minutes.

2.7.4 Standard timecourse experiment for expression from pTTQ18 derivatives

This protocol was used for the optimization of growth temperature, growth medium and harvest time for accumulation of integral membrane proteins expressed from the pTTQ18 system.

650 ml LB (supplemented with 20 mM glycerol and 100 µg/ml of either carbenicillin or ampicillin) or M9 minimal media (supplemented with 0.2%

casamino acids and 100 µg/ml of ampicillin) were inoculated to OD₆₅₀ of 0.1 with overnight cultures of the particular *E. coli* strain. Cultures were grown, with aeration, at 37°C or 25°C. All cultures were grown to mid-log phase and induced with 0.5 mM IPTG. Samples of the cultures were taken every hour after induction for 4/5 hours. A sample was also taken from the culture after overnight growth.

2.8 Preparation of bacterial extracts

2.8.1 Preparation of total membrane extracts using the water lysis method (Witholt & Boekhout, 1978)

This method was primarily used for the analysis of membrane fraction content during expression optimization of integral membrane proteins.

Bacterial cell culture is harvested by centrifugation at 5000 rpm for 15 minutes at 4°C (SLC6000 rotor in Sorvall Evolution centrifuge). The cell pellet was resuspended in 10 ml 200 mM Tris, pH 8, and incubated at room temperature for 20 minutes. After 20 minutes, 4.85 ml buffer containing 1 M sucrose, 200 mM Tris, pH 8, 1 mM EDTA was added and incubated for 1.5 minutes. 65 µl 10 mg/ml lysozyme solution (prepared freshly in dH₂O) was added after 1.5 minutes and incubated for a further 30 seconds, at which point 9.6 ml dH₂O was added. The suspension was incubated at room temperature for 20 minutes.

The resultant sphaeroplasts were collected by centrifugation at 18000 rpm for 20 minutes at 4°C (SS34 rotor in Sorvall Evolution centrifuge). The supernatant (periplasmic fraction) was discarded. The sphaeroplast pellet was resuspended in 15 ml dH₂O and incubated at room temperature for 30 minutes. The membrane fraction was collected by centrifugation at 18000 rpm for 20 minutes at 4°C (SS34 rotor in Sorvall Evolution centrifuge). The supernatant (cytoplasmic fraction) was discarded. The membrane fraction was washed twice with 30 ml of buffer containing 100 mM sodium phosphate, pH 7.2, and 1 mM β-mercaptoethanol and then resuspended in a small volume of this buffer (typically ~50-200 µl). Membrane fractions were stored at -20°C.

2.8.2 Preparation of total membrane extracts using the a FastPrep machine

This method was used to prepare total membrane fractions from small bacterial

cultures during expression test and expression optimization.

A small volume bacterial culture (typically 5ml) of bacterial culture was harvested by centrifugation and resuspended in 0.6 ml 50 mM potassium phosphate, pH 7, 10% glycerol, 10 mM MgSO₄ and 0.1 mg/ml DNase. 300 mg of glass beads were added to the cell suspension and the cells are ruptured using a Fastprep machine (Q-biogene) at force 6 setting for 20 seconds. Samples were cooled on ice for 5 minutes and 5 mM EDTA was added. Samples are centrifuged at 20000xg for 15 minutes at 4°C (Sigma 2-18K refrigerated centrifuge) and the supernatant is centrifuged again at 100000 rpm for 30 minutes at 4°C (TLA100.1 rotor in a Beckman Coulter TL100 ultracentrifuge). The resultant pellet (the mixed membrane fraction) is resuspended in 1x SDS PAGE sample buffer and loaded onto an SDS-PAGE gel for analysis.

2.8.3 Preparation of total membrane extracts from *E. coli* and *L. lactis*

This method was used primarily for the preparation of membrane extracts from large volume cultures.

Harvested *E. coli* cells were ruptured by at least 2 passes through a French pressure cell (Thermo) at 10000 psi in the presence of 100 µg/ml deoxyribonuclease I (Sigma) and 1 mM MgCl₂. Lysate was incubated on ice for 5 minutes and 5 mM EDTA, pH 7, was added.

Harvested *L. lactis* cells were ruptured by 30 minute incubation with 10 mg/ml lysozyme followed by at least 4 passes through a French pressure cell (Thermo) at 10000 psi in the presence of 100 µg/ml deoxyribonuclease I (Sigma) and 1 mM MgCl₂.

Cell lysate from both cultures were incubated on ice for 5 minutes and 5 mM EDTA, pH 7, was added. Unbroken cells and cell debris were removed by centrifugation at 13000 rpm for 20 minutes at 4°C (SS34 rotor in Sorvall Evolution centrifuge). The supernatant was centrifuged at 45000 rpm for 1 hour at 4°C (Ty70Ti rotor in a Beckman L7-65 ultracentrifuge) and the membrane pellets were resuspended in 50 mM KPi, pH 7. Mixed membrane vesicles were snap-frozen in liquid nitrogen and stored at -80°C.

2.8.4 Separation of *E. coli* inner and outer membranes using sucrose density gradient

A sucrose density gradient was prepared with percentages of sucrose (in 20 mM Tris pH 7.5) ranging from 30% to 55%, in 5% increments, in a 70 ml Ti45 ultracentrifuge tube (Beckman Coulter) with the highest density sucrose solution at the bottom and the lowest density (30%) at the top.

E. coli mixed membrane vesicles were prepared (see Methods section 2.8.3) and were applied to the top of the sucrose density gradient. The gradient was centrifuged at 38000 rpm for 16 hours at 4°C (Ty45Ti in a Beckman L60 ultracentrifuge). To maintain the integrity of the gradient slow acceleration was used and no brakes were applied for deceleration.

After the 16 hour centrifugation step, the gold-coloured inner membrane (at 35-40%) was separated from the white outer membrane (45-50%) using a large needle. The inner and outer membrane fractions were diluted by at least 3-fold and placed into separate, clean Ty45Ti ultracentrifuge tubes. The membrane fractions were harvested by ultracentrifugation at 45000 rpm for 2 hours at 4°C (Ty45Ti in a Beckman L60 ultracentrifuge). The membrane fractions were finally resuspended in 20 mM Tris pH7.5 and snap frozen in liquid nitrogen or dry-ice/ethanol and stored at -80°C.

2.8.5 Periplasmic preparation using lysozyme treatment (Neu & Heppel, 1965)

The cells were harvested by centrifugation at 5000 rpm for 15 minutes (SLC6000 or SS34 rotor in Sorvall Evolution centrifuge). The subsequent pellet was resuspended in 15 ml ice-cold 50 mM KPi pH7.8 per litre of cells in the original culture. The resuspended cells were centrifuged at 12000 rpm for 5 minutes at 4°C (SS34 rotor in Sorvall Evolution centrifuge). After centrifugation, the supernatant was discarded and the cell pellet was resuspended in SET buffer (0.5 M sucrose, 5 mM EDTA, 50 mM KPi pH 7.8 and 600 µg/ml lysozyme) in a ratio of 10 ml per litre of cells in the original culture. This suspension was incubated for 1 hour at 30°C. The subsequent sphaeroplasts were harvested by centrifugation at 12000 rpm for 30 minutes (SS34 rotor in Sorvall Evolution centrifuge); the supernatant from this centrifugation was the periplasmic fraction.

2.9 Protein concentration determination

2.9.1 Bradford assay

A set of protein standard was produced by serially diluting a stock solution of bovine serum albumin (BSA) from 400 µg/ml to 6.25 µg/ml. The protein of unknown concentration was serially diluted by at least 6 times. 50 µl of each dilution of BSA and unknown sample was added to 1 ml of Bradford reagent (Bio-rad) in a cuvette. The absorbance was measured at 595 nm wavelength. A graph of protein concentration versus absorbance (A_{595}) was plotted and the concentrations of the unknown samples calculated by comparison to the standard curve.

2.9.2 Bicinchoninic acid (BCA) protein assay

A commercial BCA assay (Pierce) was used. The assay was performed according to the manufacturer's recommendations. Briefly, a BSA standard curve was produced ranging from 900 to 28.125 µg/ml. Protein samples of unknown concentration were serially diluted 5x, 10x, 20x, 40x, 80x and 160x. Working reagent was made by mixing 50 parts Reagent A (Bicinchoninic acid) with 1 part Reagent B (4% cupric sulphate). 20 µl of each standard and sample was mixed with 180 µl working reagent in a 96-well plate and incubated at 37°C for 30 minutes. A plate reader was used to read the absorbance values at 562 nm, a standard curve was plotted of absorbance (A_{562}) versus protein concentration and the concentrations of the unknown samples was determined.

2.9.3 Schaffner-Weissmann protein assay (Schaffner & Weissmann, 1973)

A number of BSA standards were made ranging from 0-20 µg protein made up to a volume of 270 µl. The sample of unknown concentration was diluted 10-fold and 4 µl was added to 266 µl dH₂O. Each sample was treated with 30µl of 1M Tris, pH 7.5, 10% (w/v) SDS and vortexed. The protein in each sample was precipitated by addition of 60 µl 60% (w/v) trichloroacetic acid (TCA). Using a vacuum manifold the sample was applied to filter paper (Type HA, 0.45 µm, Millipore). The filter paper was then washed with 2 x 2 ml 6% TCA.

The filter paper with bound protein were stained for 3 minutes in a Petri dish containing 30 ml 0.1% Naphthalene Black (Sigma-aldrich) in methanol, acetic acid and H₂O (45:10:45 v/v). The filters were rinsed in dH₂O and destained in

30 ml methanol, acetic acid and dH₂O (90:2:8 v/v). The individual protein spots were excised and transferred to a 1.5 ml eppendorf tube. The stain was eluted with 1 ml 25 mM NaOH and 50 μ M EDTA in 50% aqueous ethanol. The absorbance was measured at 630 nm, the standard curve of absorbance versus concentration was plotted and the concentration of the unknown sample was determined.

2.9.4 Using A_{280} /extinction coefficient

The molecular weight and the extinction coefficient of purified proteins were calculated using the peptide property calculator tool on the Innovagen website (www.innovagen.se). Absorbance (A_{280}) of protein solution was measured using either a Fluoromax-3 fluorimeter or a Nanodrop ND-1000 Spectrophotometer system. Protein concentration was determined using the Beer-Lambert law.

2.10 SDS-PAGE

2.10.1 Molecular weight markers

During this project a number of different molecular weight markers have been used. Rainbow marker (GE Healthcare, #RPN756) was used for Western blots performed in Leeds, "SDS-7 Dalton Mark VII-L for SDS gel electrophoresis" (Sigma-Aldrich) was used for Coomassie stained SDS PAGE gels in Leeds and Prestain protein marker, Broad range (New England Biolabs) was used for both Western blotting and Coomassie stained SDS PAGE gels in York. Prestained ladder (Fermentas) was used for Western blots in Groningen and a low molecular weight ladder was used for Coomassie stained SDS PAGE gels.

2.10.2 Sample preparation for SDS-PAGE

The desired amount of protein (μ g) was mixed with dH₂O and 4 x sample buffer consisting of 200 mM Tris, pH 6.8, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue, 40% glycerol and 50 μ l/ml β -mercaptoethanol.

Soluble protein samples (e.g. periplasmic binding proteins) were composed as above, boiled at 95°C for 5 minutes and centrifuged for 1 minute at 13000 rpm before being loaded onto the gel. Samples containing integral membrane proteins were composed as above, but were not boiled or centrifuged before loading.

2.10.3 Mini-SDS PAGE gels

Most applications of SDS-PAGE were performed with the Mini-PROTEAN 3 Cell system (Bio-rad) using 12% agarose resolving gels and 4% agarose stacking gels. The manufacturer's guidelines were used for the assembly of the SDS-PAGE apparatus. A 10 ml solution of 12% resolving gel was made by mixing 3.4 ml dH₂O, 4 ml Protogel (30% acrylamide, 0.8% bisacrylamide), 2.5 ml 1.5 M Tris, pH 8.8, 100 µl 10% SDS, 100 µl ammonium persulphate (APS) and finally 10 µl N, N, N', N-tetramethylethylenediamine (TEMED). A 10 ml solution of 4% stacking gel was made by mixing 6.1 ml dH₂O, 1.3 ml Protogel (30% acrylamide, 0.8% bisacrylamide), 2.5 ml 0.5 M Tris, pH 6.8, 100 µl 10% SDS, 100 µl APS and finally 10 µl TEMED. The SDS-PAGE tank buffer consisted of 3 g/L Tris, 14 g/L glycine and 1 g/L SDS. SDS-PAGE was performed at a constant 200 V for ~45 minutes.

2.10.4 Large SDS-PAGE gels

The SDS-PAGE gels performed in Professor Peter Henderson's laboratory in Leeds were large (~17 x 11 cm gels) 15% gels, however, a similar protocol to that used for the Mini-PROTEAN 3 Cell system was applied with the following modifications; the 15% resolving gel consisted of 4.22 ml 40% acrylamide, 0.49 ml bisacrylamide, 2.81 ml 1.5 M Tris, pH 8.7, 100 µl 10% SDS, 3.48 ml dH₂O, 37 µl 10% APS and 12 µl TEMED. The 3% stacking gel consisted of 0.77 ml 40% acrylamide, 0.39 ml bisacrylamide, 0.75 ml 0.5 M Tris, pH 6.8, 50 µl 10% SDS, 3.2 ml dH₂O, 30 µl 10% APS and 9 µl TEMED. Electrophoresis was performed at a constant 23 mA for approximately 2 hours.

2.10.5 Staining/destaining SDS-PAGE gels

Coomassie Brilliant Blue (CBB) dye was used to detect proteins on SDS-PAGE gels. CBB consists of 2.5 g Coomassie Brilliant Blue, 450 ml methanol, 450 ml dH₂O and 100 ml acetic acid. Gels were removed from the SDS-PAGE apparatus and washed briefly with H₂O to remove excess SDS. The gel was submerged in approximately 30-50 ml of CBB dye and incubated for 5-10 minutes with agitation. The excess CBB was then poured off and the gel was briefly rinsed with H₂O to remove the residual dye. Approximately 100 ml of destain containing 10% acetic acid and 10% ethanol was added and heated at

full power in a 750 W microwave for 30 seconds. The gel is left to destain with agitation overnight.

2.11 Western Blotting

Different Western blotting protocols were used in the three laboratories where the work was performed. Almost all of the Western blotting for YiaM expression and purification was performed in Leeds using semi-dry transfer. Western blotting of YiaN and SiaQM expressed from the pTTQ18 system were performed in York with wet transfer system. Western blotting of SiaQM and SiaP produced from pBAD and pNZ8048 was performed in Groningen with a semi-dry transfer system. The following protocols detail the Western blotting procedures used in each laboratory.

2.11.1 Wet transfer of protein

Wet transfer from SDS-PAGE gel to nitrocellulose membrane (GE Healthcare) was performed using a Mini Trans-Blot electrophoretic transfer cell (Bio-rad). Assembly of the apparatus and protein transfer was performed according to the manufacturer's guidelines. The wet transfer was performed at 350 mA for 1 hour in the presence of a cooling unit and constant stirring.

2.11.2 Semi-dry transfer of protein

Semi-dry transfer was performed using a Trans-Blot SD Semi-dry electrophoretic Transfer Cell (Bio-rad). Assembly of the apparatus and protein transfer was performed according to the manufacturer's guidelines. In the Leeds laboratory, the transfer buffer was 0.5x SDS-PAGE running buffer and PVDF membrane (Pall) was used. In the Groningen and York laboratories and the transfer buffer consisted of 5.82 g/L Tris, 2.93 g/L glycine and 200ml/L methanol and nitrocellulose membrane (GE Healthcare) was used. The transfer was performed at either 52 mA for 2 hours (in Leeds), 80 mA for 30 minutes (in Groningen) or 15 V for 30 minutes (in York).

2.11.3 Processing the membrane after protein transfer (York)

The post-transfer membrane is incubated with 10 ml Ponceau S stain (Sigma Aldrich) for 15 minutes at room temperature. After analysis of the Ponceau S

stained membrane, the stain was washed off with H₂O and incubated in 20 ml blocking solution (20 ml TBST and 5% milk powder) for 30-60 minutes at room temperature. The blocking solution was removed and the membrane was washed with 20 ml TBST for 15 minutes followed by a further 2 x 5 minute washes. The membrane was incubated with 1^o hybridization buffer (5 ml TBST, 1% milk powder and a 1/2000 dilution of anti-tetrahis (mouse) antibody (Qiagen)) overnight at 4°C. After incubation with the 1^o hybridization buffer, the membrane was washed with TBST for 15 minutes followed by 3 washes at 5 minute intervals at room temperature. The membrane was then incubated with 2^o hybridization buffer (20 ml TBST, 1% milk powder and 1/5000 dilution of rabbit anti-mouse IgG (H+L) HRP conjugate (Zymax)) for 1 hour at room temperature. The membrane was washed with TBST for 15 minutes followed by 6 washes at 5 minute intervals. The visualization of the histidine-tagged proteins took place in the dark room and was performed using either the Amersham ECL kit or the Pierce Supersignal ECL kit. Both were performed using the manufacturer's guidelines.

2.11.4 Processing the membrane after protein transfer (Groningen)

The membrane is blocked PBST and 0.2% I-Block (Applied Biosystems) overnight at 4°C. The rest of the protocol is performed by a home-made, automated Western blotting machine named "Blot-o-matic". The protocol that is performed is as follows; the membrane is blocked for at least 60 minutes, but can be left as long as is necessary. The membrane is then incubated with a solution containing PBST, 0.1% I-block and 100 mM EDTA for 60 minutes. It is then washed 3 times with PBST containing 0.1% I-Block and then incubated with 1^o hybridization solution. This solution contains 10 ml PBST containing 0.1% I-block and a 1/3333 dilution of 1^o antibody. The membrane is then washed 3 times with PBST containing 0.1% I-block and incubated with 2^o antibody solution which consists of 10 ml PBST containing 0.1% I-block and a 1/5000 dilution of 2^o antibody. The membrane is then washed 3 more times with PBST containing 0.1% I-block.

The membrane is washed twice with assay buffer. This consists of 5.62 g diethanolamine, 1 mM MgCl₂ and 0.5 ml concentrated HCl in 500 ml ddH₂O. The membrane is incubated for 5 minutes with Nitroblot solution which is 2 ml of assay buffer and 200 µl Nitroblock (Applied Biosystems). The membranes are

then washed twice with assay buffer with 5 minute intervals and the CSPD solution is added and incubated for 5 minutes. The CSPD solution contains 2 ml assay buffer and 10 μ l CSPD (Applied Biosystems). The CSPD solution is poured off and the membrane is rinsed with assay buffer. The his-tagged proteins are visualized using a Fujifilm Intelligent Dark Box II.

2.11.5 Processing the membrane after protein transfer (Leeds)

After transfer the membrane is washed with TBSTT for 10 minutes and then incubated with blocking solution overnight at 4°C. The blocking solution consists of 20 ml TBSTT and 3% BSA. The blocking solution is removed and the membrane is washed twice with 10 ml TBSTT after which the membrane is incubated with 1/5000 dilution (in TBSTT) of INDIA Hisprobe-HRP (Pierce). The use of this particular 1° antibody foregoes the need for a 2° antibody, so after an hour the antibody solution is discarded and the membrane is washed 4 times with 10 ml TBSTT with 10 minute intervals. A Pierce supersignal ECL kit is used to visualize the his-tagged protein.

2.12 Tryptophan fluorescence spectroscopy

Tryptophan fluorescence spectroscopy was performed with the Fluoromax-3 spectrophotometer (Jobin Yvon Horiba). Titrations were performed using time-based acquisition with an excitation wavelength of 295 nm, an emission wavelength of 328 nm and a slit width of 5 nm. All reactions were performed with the ESR in 3 ml 50 mM Tris-HCl, pH 8, at 25°C. Increasing concentrations of ligand was added to the ESR solution and the fluorescence change was measured. The cumulative fluorescence change was plotted against the cumulative concentration of ligand in SigmaPlot 10. The data was fitted to a single rectangular hyperbola and non-linear regression was used to solve the equation and produce values for the dissociation constant (K_d).

2.13 Integral membrane protein detergent solubility screening

Inner membrane or mixed membrane samples were solubilised with several different detergents to determine which detergent the integral membrane proteins were most stable in. A standard 200 μ l solubilisation reaction was composed of 2 μ g/ μ l membrane vesicles, 200 mM NaCl, 10% glycerol, 20 mM HEPES, pH 7, 1% detergent. The reaction was mixed for between 1 hour and

overnight at 4°C. The reactions were centrifuged at 100000xg (TLA100 rotor, Beckman Coulter) for 1 hour to separate the soluble and insoluble protein. The 200 µl soluble supernatant is removed, the insoluble pellet is resuspended in 200 µl and the protein content of each fraction is analysed using SDS-PAGE and Western blotting.

2.14 Protein purification techniques

2.14.1 Extraction of integral membrane proteins for purification by DDM solubilisation

Prior to chromatography-based purification procedures on integral membrane proteins, the protein of interest must be extracted from the hydrophobic environment of the membrane by solubilisation in detergent. *n*-dodecyl- β -D-maltoside (DDM) is a common detergent used in the solubilisation of integral membrane proteins for purification.

Membrane vesicles containing overexpressed integral membrane protein were resuspended to a concentration of 3.33 mg/ml and solubilised in solubilisation buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.5% *n*-dodecyl- β ,D-maltoside (DDM), 10 mM imidazole). The solubilisation reaction was mixed by inversion and incubated on ice for 30 minutes. Insoluble material was removed from the solution by centrifugation at 53000 rpm (using TLA100.3 rotor, Beckman Coulter) for 20 minutes at 4°C. The supernatant from this centrifugation contains solubilised integral membrane protein and can now be applied to chromatographic apparatus.

2.14.2 Ni²⁺-NTA affinity chromatography using batch washing and disposable columns

The soluble fraction containing the integral membrane protein of interest is incubated with equilibrated Ni²⁺-NTA superflow resin (Qiagen, 1.5 ml/10 mg protein) for 2 hours at 4°C on a rotating platform. After 2 hours the resin suspension is centrifuged at 1000xg for 1 minute (Sigma refrigerated benchtop centrifuge) and the supernatant is removed – this contains protein that did not bind to the resin. The resin is resuspended in 10 ml wash buffer (20 mM Tris, pH 8, 10% glycerol, 300 mM NaCl, 20 mM imidazole and 0.05% DDM) and centrifuged as before. This is repeated 10 times. The resin is resuspended in 5

ml wash buffer and applied to a disposable 2 ml polystyrene column (Bio-rad). Bound protein is eluted with elution buffer (20 mM Tris, pH 8, 10% glycerol, 200 mM imidazole and 0.05% DDM) and collected in fractions of 300 μ l followed by 5 x 750 μ l fractions.

2.14.3 Column-based Ni²⁺ affinity chromatography

A volume between 1-8 ml of Ni²⁺-NTA superflow resin (Qiagen) is packed into an 8 ml C 10/10 column (GE Healthcare) and connected to a Biologic FPLC (Bio-rad) or a Peristaltic Pump P-1 (GE Healthcare). The resin is equilibrated with 5 column volumes of equilibration buffer (20 mM Sodium Pi, pH 7.2, 10% glycerol, 200 mM NaCl, 0.05% DDM (w/v), 10 mM imidazole) and the soluble fraction after DDM solubilisation is applied at a flow rate of 0.25-0.5 ml/min. The column flowthrough is collected for analysis. The resin is washed with 10 column volumes of wash buffer (20 mM Sodium Pi, pH 7.2, 10% glycerol, 200 mM NaCl, 0.05% DDM (w/v), 20 mM imidazole). Bound protein is eluted with elution buffer (20 mM Sodium Pi, pH 7.2, 10% glycerol, 200 mM NaCl, 0.05% DDM (w/v), 250 mM imidazole).

2.14.4 Ni²⁺ affinity chromatography using disposable gravity-flow columns

The soluble fraction containing the integral membrane protein of interest is incubated with equilibrated Ni²⁺-NTA superflow resin (Qiagen, 0.1 ml/10 mg protein) for 1 hour in a disposable polystyrene column (Pierce). The resin is allowed to sediment and the flow-through is collected. The resin is washed with 20 column volumes of DDM wash buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.04% DDM (w/v), 40 mM imidazole). Bound protein is eluted with Elution buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.04% DDM (w/v), 500 mM imidazole). 5 mM EDTA was added to the elution fractions.

2.14.5 Ni²⁺ affinity chromatography using HisTrap HP columns

The soluble fraction containing the protein of interest is dialysed into 50 mM Tris, pH 8, 300 mM NaCl overnight and then applied to the 1 ml HisTrap column (GE Healthcare) using a Biologic FPLC (Bio-rad). The column is washed with wash buffer (50 mM Tris, pH 8, 300 mM NaCl, 20 mM imidazole) and protein is

eluted using elution buffer (50 mM Tris, pH 8, 300 mM NaCl, 200 mM imidazole).

2.14.6 Anion exchange chromatography

Diethylaminoethanol (DEAE) resin (Sigma-Aldrich) is packed into a 8 ml column (GE Healthcare) and connected to a Biologic FPLC (Bio-rad). The resin is washed 10 CV dH₂O and equilibrated with low salt buffer (20 mM Tris, pH8). The supernatant is applied to the resin, which is then washed with 3-5 column volumes of low salt buffer. The flow through is collected. A linear gradient is applied over 20 column volumes from 0 to 500 mM NaCl with a flow rate of ~0.5 ml/minute. Fractions are collected throughout and the resulting A₂₈₀ spectrum is used to identify the location of the eluted protein.

2.14.7 Size exclusion chromatography (SEC)

SEC was performed using a Superdex 20 10/300 GL column (GE Healthcare) in combination with an AKTA HPLC (GE Healthcare). The column was equilibrated with filtered SEC buffer (50 mM potassium phosphate buffer, pH 7.8, 200 mM NaCl, 0.04% (w/v) DDM). Protein sample was centrifuged at 20000xg for 20 minutes at 4°C to remove any precipitants. Protein sample was loaded onto the column and 1.5 column volumes of SEC buffer were applied at a rate of 1 ml/min and 0.5 ml fractions were collected.

2.15 Reconstitution of integral membrane protein into liposomes

2.15.1 Purification of *E. coli* lipid extract and production of liposomes

This method details the extraction of total phospholipids from total *E. coli* lipids (Avanti Polar lipids). The extracted phospholipids are then mixed with Egg phosphatidylcholine (Egg-PC, Avanti Polar lipids) which compose the liposomes that will be used to reconstitute integral membrane proteins.

1 g of total *E. coli* lipids in ~30 ml chloroform is completely dried in a round-bottomed flask by rotary evaporation and then resolubilised in 5 ml of chloroform.

While stirring the mixture vigorously at 4°C in a stoppered conical flask, the phospholipid is added dropwise to 150 ml ice-cold acetone and 1 mM DTT – a

white precipitant should form. The conical flask is flushed with N₂, the stopper is replaced and the mixture is allowed to stir overnight at 4°C in the dark.

The precipitant suspension is centrifuged at 6000 rpm for 10 minutes at 4°C (SS34 rotor in polypropylene copolymer (PPCO), Teflon (FEP) or glass centrifuge tubes). After centrifugation, the supernatant is discarded and the pellets are completely dried by a stream of N₂. The pellets are then solubilised in 150 ml diethyl ether, transferred to a foil-wrapped conical flask and stirred for 10-20 minutes at room temperature. This solution is then 6000 rpm for 10 minutes at 4°C (SS34 rotor) and the supernatant is added to a pre-weighed round-bottomed flask and dried using rotary evaporation at 30°C. The pure, dried phospholipids are then dissolved in chloroform such that the final concentration of phospholipids is 100 mg/ml.

The 100 mg/ml solution of purified *E. coli* lipids is mixed with 20 mg/ml Egg-PC in a ratio of 3:1 (w/w) and dried using a rotary evaporation at 30°C. Once the phospholipid/Egg-PC mixture is dried, it is rinsed with 96% ethanol and dried again. The dried phospholipids/Egg-PC mixture is then resuspended in 50 mM potassium phosphate, pH 7, to a final concentration of 20 mg/ml. The suspension undergoes 4-6 cycles of sonication with a Misonix sonicator 3000 (15 seconds on/45 seconds off with 4 µm intensity) in a N₂ environment. The liposomes are separated into 1 ml aliquots, go through at least 3 freeze-thaw cycles and are finally snap frozen and stored at -80°C

2.15.2 Reconstitution of SiaQM into liposomes by the rapid dilution method

Integral membrane protein was solubilised in DDM and purified by Ni²⁺ affinity chromatography as previously described in section 2.14.4 with the following modifications; after ultracentrifugation, the solubilised protein sample is incubated with the Ni-NTA resin and then washed with 10 column volumes of DDM wash buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.04% DDM (w/v), 40 mM imidazole) followed by 10 column volumes of decyl-β,D-maltoside (DM) wash buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.1% DM (w/v), 40 mM imidazole). The bound protein is eluted with DM-containing elution buffer (50 mM Potassium Pi, pH 7.8, 20% glycerol, 200 mM NaCl, 0.1% DM (w/v), 500 mM imidazole). 5 mM EDTA was added to the elution fractions.

For the reconstitution of this DM-solubilised purified protein, liposomes were prepared as a 3:1 ratio of purified *E. coli* lipids and L- α -phosphatidylcholine from egg yolk (see Methods section 2.15.1). Purified membrane protein was added to liposomes in a 1:40 ratio (w/w) and incubated on ice for 10 minutes. The mixture was rapidly diluted by mixing with 80 ml 50 mM potassium phosphate, pH 7. Proteoliposomes were collected by centrifugation at 45000 rpm for 2 hours at 4°C. Proteoliposomes were repeatedly freeze-thawed before used in *in vitro* transport assays.

2.15.3 Reconstitution by the destabilized liposome method (Knol *et al.*, 1996)

Integral membrane protein was solubilised in DDM and purified by Ni²⁺ affinity chromatography as described in Methods section 2.14.4. Liposomes were thawed to room temperature and extruded 11 times through 400 nm filters (Avestin). The liposomes were diluted to 4 mg/ml with 50 mM potassium phosphate, pH 7, and titrated with 10% (w/v) Triton X-100. Using a UV-visible spectrometer (Shimadzu UV-1601) the absorbance (A_{540}) of the liposome solution was followed while Triton X-100 was slowly added. At the point of liposome solubilisation, purified protein is added at a 100:1 lipid to protein ratio. The mixture is incubated at room temperature for 20 minutes, at which point 40 mg Biobeads are added (SM2 – Bio-rad) per 1 ml of mixture. This is incubated at room temperature for 30 minutes. Another 40 mg/ml Biobeads are added and incubated for 1 hour at 4°C followed by a further 40 mg/ml addition of Biobeads. The mixture is incubated overnight at 4°C. The Biobeads are removed from the mixture by filtration and the proteoliposomes are harvested by centrifugation at 80000 rpm for 15 minutes at 4°C (Beckman TLA 100.1 rotor). The proteoliposomes are resuspended in 50 mM potassium phosphate, pH 7, snap frozen and stored at -80°C.

2.16 *In vitro* transport assays

For all transport assays, *N*-acetylneuraminic acid was used, which was radiolabelled on carbon 6 and had a specific activity of either 0.025 or 0.05 mCi/ml.

2.16.1 Preparation of proteoliposomes for *in vitro* [¹⁴C]-Neu5Ac transport assays

On the day of the uptake assay, the proteoliposomes (at 400 µg/ml integral membrane protein concentration) were thawed on ice and collected by centrifugation at 50000 rpm for 20 minutes at 4°C. The proteoliposomes were resuspended in Inside buffer (100 mM KAc pH 7, 20 mM KPi and 2 mM MgSO₄, unless stated otherwise) and extruded 11 times through a 400 nm filter (Avestin). The extruded proteoliposomes were collected by centrifugation at 50000 rpm for 20 minutes at 15°C. The proteoliposomes were then resuspended in Inside buffer to a final concentration of 4 µg/µl SiaQM.

2.16.2 Standard *in vitro* [¹⁴C]-Neu5Ac transport assay

The standard *in vitro* uptake assay was performed by incubating 60 µl of appropriate external buffer (see sections 2.16.3.1-2.16.3.5) for 1 min at 30°C and adding 5 µM [¹⁴C]-Neu5Ac. Electrochemical gradients are formed upon addition of 12 µl 4 µg/µl proteoliposomes. 100 µl samples are taken at 20, 60, 100, 140 and 180 seconds and are then incubated for 10 seconds with 100 µl cold wash buffer (these buffers have the same constituents as the external buffers for each assay, except no SiaP is present and 1 mM Neu5Ac is included instead of [¹⁴C]-Neu5Ac) and bound to filter paper (BA85, Schleicher and Schuell). Filters are washed with 2 ml 50 mM potassium phosphate buffer, pH 7. The counts per minute (cpm) for each sample are evaluated using a liquid scintillation analyzer (Packard Tri-carb 2900TR). *In vitro* transport assays can also be performed with half the volumes of all the solutions involved with minimal difference in transport activity and standard deviations.

2.16.3 External buffers for formation of electrochemical gradients

2.16.3.1 Na⁺ gradient alone ($\Delta\mu_{\text{Na}^+}$)

Na⁺ gradient alone ($\Delta\mu_{\text{Na}^+}$) was formed by dilution of proteoliposomes containing inside buffer (see above) into buffer containing 100 mM Sodium acetate, 20 mM Sodium PIPES, pH 7, 2 mM MgSO₄ and 5 µM SiaP.

2.16.3.2 Na⁺ gradient plus membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$)

Na⁺ gradient plus membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$) was formed by dilution of

proteoliposomes containing inside buffer into buffer containing 100 mM Sodium acetate, 20 mM Sodium PIPES, pH 7, 2 mM MgSO₄, 2 μM Valinomycin and 5 μM SiaP.

2.16.3.3 Na⁺ gradient plus pH gradient ($\Delta\mu\text{Na}^+ + \Delta\text{pH}$)

Na⁺ gradient plus pH gradient ($\Delta\mu\text{Na}^+ + \Delta\text{pH}$) was formed by dilution of proteoliposomes containing inside buffer into buffer containing 120 mM Sodium PIPES, pH 7, 2 mM MgSO₄ and 5 μM SiaP.

2.16.3.4 Membrane potential alone ($\Delta\Psi$)

A membrane potential ($\Delta\Psi$) was formed by dilution of proteoliposome containing inside buffer into buffer containing 100 mM *N*-methyl glucamine acetate, pH 7, 20 mM *N*-methyl glucamine phosphate, pH 7, 2 mM MgSO₄ and 5 μM SiaP.

2.16.3.5 pH gradient alone (ΔpH)

A pH gradient (ΔpH) was formed by dilution of proteoliposomes containing inside buffer into buffer containing 120 mM potassium phosphate, pH7, 2 mM MgSO₄ and 5 μM SiaP.

2.16.4 Solute counterflow assay

SiaQM-containing proteoliposomes are treated as described in section 2.16.1, except that the proteoliposomes are loaded with counterflow internal buffer (100 mM KAc, pH7, 20 mM KPi, pH7, 10 mM NaPIPES and 1 mM Neu5Ac). The solute counterflow assay was performed by incubating 600 μl of counterflow external buffer (100 mM KAc, pH7, 20 mM KPi, pH7, 10 mM NaPIPES, and 5 μM SiaP (where applicable)) for 1 min at 30°C and adding 5 μM [¹⁴C]-Neu5Ac. Electrochemical gradients are formed upon addition of 12 μl 4 μg/μl proteolipomes. 100 μl samples are taken at 30, 60, 120, 240 and 480 seconds and are then incubated for 10 seconds with 100 μl cold wash buffer (these buffers have the same constituents as the external buffers for each assay, except no SiaP is present and 1 mM Neu5Ac is included instead of [¹⁴C]-Neu5Ac) and bound to filter paper (BA85, Schleicher and Schuell). Filters are washed with 2 ml 50 mM potassium phosphate buffer, pH 7. The counts per

minute (cpma) for each sample are evaluated using a liquid scintillation analyzer (Packard Tri-carb 2900TR).

2.17 *In vivo* transport assays

2.17.1 Preparation of cells for *in vivo* [¹⁴C]-Neu5Ac transport assay

2x20 ml of LB (supplemented with 100 µg/ml ampicillin, where necessary) was inoculated with 0.2 ml (1% inoculum) from cultures of strains BW25113 Δ *nanT* pES7 and BW25113 Δ *nanT* grown overnight. The 20 ml cultures were grown aerobically to mid-log phase (OD₆₆₀) at 37°C, induced with 1 mM IPTG and allowed to grow for a further 2 hours. Cells were harvested by centrifugation at 8000 rpm for 8 minutes at 4°C (using SS34 rotor in Sorvall evolution centrifuge) and washed twice with 20 ml of 50 mM KPi, pH 7, 2 mM MgSO₄. The cells were finally resuspended in a small volume such that the OD₆₆₀ would equal 120. Cells were stored on ice until required.

2.17.2 *In vivo* [¹⁴C]-Neu5Ac transport assay: application of PMF ($\Delta\mu$ H⁺ and $\Delta\mu$ Li⁺)

600 µl of external buffer (50 mM KPi, pH 7, 2 mM MgSO₄ and 10 mM D-Li-lactate) was incubated for 1 min at 30°C and 7.6 µl cell suspension (OD₆₆₀=120) was added to attain a final OD₆₆₀ of 1.5. This was incubated at 30°C for 2 minutes under continuous airflow (water-saturated), at which point 8 µM Neu5Ac (a 50:50 mixture of unlabelled Neu5Ac and [¹⁴C]-Neu5Ac) was added. 100 µl samples are taken at 0.5, 1, 2, 3 and 4 minutes and bound to filter paper (BA85, Schleicher and Schuell) Filters are washed with 2 ml 50 mM potassium phosphate buffer, pH 7. The counts per minute (cpma) for each sample are evaluated using a liquid scintillation analyzer (Packard Tri-carb 2900TR).

2.17.3 *In vivo* [¹⁴C]-Neu5Ac transport assay: application of SMF ($\Delta\mu$ Na⁺)

600 µl of external buffer (50 mM KPi, pH 7, 2 mM MgSO₄ and 20 mM NaCl) was incubated for 1 min at 30°C and 8 µM Neu5Ac (a 50:50 mixture of unlabelled Neu5Ac and [¹⁴C]-Neu5Ac) was added. This was incubated for 1 minute at 30°C at which point 7.6 µl cell suspension (OD₆₆₀=120) was added to attain a final OD₆₆₀ of 1.5. 100 µl samples are taken at 0.5, 1, 2, 3 and 4

minutes and bound to filter paper (BA85, Schleicher and Schuell) Filters are washed with 2 ml 50 mM potassium phosphate buffer, pH 7. The counts per minute (cpm) for each sample are evaluated using a liquid scintillation analyzer (Packard Tri-carb 2900TR).

2.17.4 *In vivo* [¹⁴C]-Neu5Ac transport assay: application of $\Delta\mu\text{Li}^+$

Exactly the same protocol as 2.17.3, except 20 mM LiCl was in the external buffer, and not 20 mM NaCl.

Chapter 3

Genome-wide analysis of TRAP transporter gene frequency and organisation

3.1 The TRAPdb relational database

As more prokaryotic genome sequences have been published, it has become evident that TRAP transporters are widespread in both bacteria and archaea (Kelly & Thomas, 2001). With hundreds of TRAP transporter gene sequences now known there is a need for a unified resource of TRAP transporter data. TRAPdb (www.trapdb.org) was created to be this resource and to store data on all known TRAP transporters and make this data accessible in a user-friendly, intuitive fashion. The data collected for TRAPdb would not only be used to populate TRAPdb with data on individual systems, but could also be used to determine trends in frequency and distribution of TRAP transporters in a wide variety of prokaryotes.

TRAPdb was created using MySQL and the web interface was created using Macromedia Coldfusion software (by Peter Bryant, a bioinformatics MRes student) and was populated initially with 103 TRAP genes comprising 49 separate systems from 23 different organisms, taken from the TRAP transporter review of 2001 (Kelly & Thomas, 2001). Information can be retrieved from TRAPdb using either a text search function or by selecting search criteria from a number of pull-down menus (Fig. 3.1a). These criteria include searching for transporters with the same substrate (predicted or experimentally determined), all TRAP transporters with a particular gene organisation (i.e. system type), all TRAP transporters within a particular organism and all genes of a particular family, i.e. DctP, TAXI, DctQ, DctM, etc (Fig. 3.1a).

Once the target TRAP system has been selected, the information regarding that system is displayed, information such as; gene name, locus tag, accession number, sequence length and molecular weight. The amino acid sequence of one or more of the genes within a particular system can also be displayed (Fig. 3.1b). If searching by organism name then the number of TRAP transporters in that particular organism is displayed. There are also a number of external links available; such as the taxonomy of the organism and the Swissprot/trEMBL webpage for the particular protein.

TRAPdb was further populated with amino acid sequences and TRAP protein data such as locus tag, length and molecular weight collected from NCBI using

(a)

TrapDB

The tripartite ATP-independent periplasmic (TRAP) transporters database

[Introduction](#) [Credits](#) [Contact](#)

Gene Name

Search TRAPdb

Search by organism:

Get Systems

Search by substrate:

Get Systems

Search by system type:

Get Systems

Search by gene family:

Get genes

Supported by:



Website Designed by Peter Bryant 2004, updated 2007 by Richard Horler and curated by Christopher Mulhgan and Gavin Thomas.
University of York, UK

(b)

TrapDB

[Home](#)

TRAP system

Organism	System Type	Substrate Specificity	System Id
Haemophilus influenzae Rd KW20	DctP and DctQM	Sialic acid	11

Components of system

Gene Name	SiaP
Gene Family	DctP
Locus Tag	HI0146
TRAPdb ID	TDB_019
Accession Number	NP_438315
SwissProt/TrEMBL	P44542
Genome context	161558-162547
Sequence Length	329
Molecular Weight	36513.0
Retrieve sequence	<input type="checkbox"/>

Gene Name	SiaQM
Gene Family	DctQM
Locus Tag	HI0147
TRAPdb ID	TDB_020
Accession Number	NP_438316
SwissProt/TrEMBL	PP44543
Genome context	162558-164459
Sequence Length	643
Molecular Weight	69378.0
Retrieve sequence	<input type="checkbox"/>

Figure 3.1 Screen captures of TRAPdb web interface. (a) TRAPdb front page showing input box for free text searching and pull-down menus. (b) TRAPdb record page showing SiaP and SiaQM from *Haemophilus influenzae*.

BLAST searches to find homologues of the ESR subunits that were already in TRAPdb. Details of TRAP systems that were missed in these searches were obtained from Interpro (www.ebi.ac.uk/interpro). A complete TRAP transporter system was defined as the three TRAP component genes (*dctP*, *dctQ* and *dctM*, whether fused or not) present in close proximity in the genome.

As of February 2008, TRAPdb held information on 2412 TRAP component genes that made up at least 770 different complete TRAP transporter systems spanning 243 prokaryotic genomes, providing a significant resource for analysis.

The TRAP transporter review in 2001 detailed a number of varieties of genetic organisation of TRAP transporters (Kelly & Thomas, 2001). This extensive data-mining was an ideal opportunity to update the list of known TRAP transporter organisation. To do this, a web-based application, The Seed (www.theseed.org) was used to analyse the genome context for the TRAP transporters collected. Analysis of the flanking genes in conjunction with a multiple sequence alignment of all DctP-type ESR amino acid sequences collected, using the program ClustalX, has also allowed some substrate predictions to be made.

3.2 Specific enrichment of TRAP transporters in marine bacteria

The extensive data mining that was performed allowed the frequency and distribution of TRAP transporters to be assessed for the first time. Analysis of the number of DctP-type and TAXI-type TRAP transporters in certain organisms in conjunction with other data, for example, the phylum of the organism, the genome length and the environment, could be used to investigate whether there are any TRAP transporter distribution trends.

Using the data collected for the compilation of TRAPdb, the total number of complete systems were calculated for each organism. A complete TRAP system was defined by having the gene for at least one ESR component in the same operon as the genes for the two integral membrane components. These systems were then divided into either a DctP-type system or a TAXI-type system.

TRAP densities were calculated for 193 organisms by dividing the number of complete TRAP transporters (DctP-type and TAXI-type) encoded by an organism by the size of the genome (Mb). The number of TRAP transporters in an organism was also normalised by the total number of transporters in the organism, where the data was available (data provided by Quingu Ren and Ian Paulsen from TransportDB). This was expressed in a percentage and was performed for 70 organisms.

The total number of complete TRAP transporters was defined for 193 organisms. The phylum and environment of the organism were also collected and the list of organisms was ordered by TRAP density, with the top 17 being presented (Table 3.1). Table 1 also contains other organisms mentioned in the text plus *E. coli* and *Streptomyces coelicolor* A3 (2). *Streptomyces coelicolor* A3 (2) is included as it has a large genome, but is devoid of TRAP transporters indicating the size of genome is not directly proportional to the number of TRAP transporters. *E. coli* was included as a baseline because it is a common species with an average genome size, but only 1 TRAP transporter. There were 4 organisms with a TRAP density greater than 5, and for *Silicibacter pomeroyi* DSS-3 and *Jannaschia* sp. CCS1 data were available from TransportDB, revealing that TRAP transporters make up 10 and 9.13% of the total transporters in these organisms. For the 70 organisms where data were available on TransportDB, the average percentage contribution of TRAP transporters to the total number of transporters was calculated. The average was found to be 2.05, indicating that *Silicibacter pomeroyi* DSS-3 and *Jannaschia* sp. CCS1 have 5 fold higher than average TRAP density. The four organisms with the highest TRAP densities and a number of other organisms with above average TRAP densities were defined as marine dwelling organisms, i.e were all found in high NaCl concentrations. These organisms were not limited to one phylum, therefore this increased use of TRAP transporters was not lineage-dependent expansion. Indeed, there are examples where closely related organisms have vastly different numbers of TRAP transporters, which is proposed to be due to different environments and a requirement for NaCl in the growth media for one of the organisms. For example, *Bacillus subtilis* is soil-dwelling and has no

Organism name	Phylum	Environment	TRAP systems (DctP/TAXI)	Genome length (Mbp)	TRAP density	Total no. of transporters	% of total transporters
<i>Aurantimonas</i> sp. SI85-9A1	α -proteobacteria	Aquatic	27 (22/5)	4.33	6.24	-	-
<i>Silicibacter pomeroyi</i> DSS-3	α -proteobacteria	Aquatic	28 (24/4)	4.60	6.09	280	10
<i>Chromohalobacter salexigens</i> DSM 3043	γ -proteobacteria	Aquatic	21 (17/4)	3.70	5.68	-	-
<i>Jannaschia</i> sp. CCS1	α -proteobacteria	Aquatic	23 (19/4)	4.40	5.22	252	9.13
<i>Bordetella bronchiseptica</i> RB50 NCTC-1325	β -proteobacteria	Host	25 (23/2)	5.34	4.68	332	7.53
<i>Bordetella parapertussis</i> 12822 NCTC-1325	β -proteobacteria	Host	21 (19/2)	4.77	4.40	290	7.24
<i>Silicibacter</i> sp. TM1040	α -proteobacteria	Aquatic	15 (14/1)	4.15	3.61	-	-
<i>Pasteurella multocida</i> Pm70	γ -proteobacteria	Host	8 (6/2)	2.26	3.54	167	4.79
<i>Actinobacillus succinogenes</i> 130Z	γ -proteobacteria	Host	7 (6/1)	2.05	3.42	-	-
<i>Paracoccus denitrificans</i> PD1222	α -proteobacteria	Multiple	17 (16/1)	5.16	3.28	-	-
<i>Polaromonas</i> sp. JS666	β -proteobacteria	Multiple	18 (17/1)	5.90	3.05	-	-
<i>Mesorhizobium</i> sp. BNC1	α -proteobacteria	Multiple	15 (11/4)	4.94	3.04	-	-
<i>Oceanicola granulosus</i> HTCC2516	α -proteobacteria	Aquatic	12 (11/1)	4.04	2.97	-	-
<i>Bordetella pertussis</i> Tohama I NCTC-13251	β -proteobacteria	Host	12 (10/2)	4.09	2.94	231	5.19
<i>Oceanospirillum</i> sp. MED92	γ -proteobacteria	Aquatic	11 (8/3)	3.87	2.84	-	-
<i>Thermus thermophilus</i> HB27	Deinococcus-Thermus	Aquatic	6 (4/2)	2.13	2.82	108	5.56
<i>Rhodobacter sphaeroides</i> 2.4.1	α -proteobacteria	Multiple	12 (10/2)	4.45	2.70	224	5.36
<i>Archaeoglobus fulgidus</i> DSM 4304	Euryarchaeota	Aquatic	3 (3/0)	2.18	1.38	104	2.88
<i>Escherichia coli</i> K12	γ -proteobacteria	Host	1 (1/0)	4.64	0.22	383	0.26
<i>Bacillus subtilis</i>	Firmicutes	Terrestrial	0 (0/0)	4.21	0	297	0
<i>Streptomyces coelicolor</i> A3(2)	Actinobacteria	Multiple	0 (0/0)	9.34	0	442	0

Table 3.1 Organisms containing high frequencies of TRAP transporters ordered by their TRAP density. Seventeen organisms with a TRAP density over 2.5 are presented, with additional statistics for four genomes mentioned in the text.

TRAP transporters, whereas its relative, the alkaliphilic *Bacillus halodurans* has 8, which is proposed to be due to the requirement for high NaCl concentrations for growth (Takami *et al.*, 2000). This difference was also observed for *Rhodobacter capsulatus* and *Rhodobacter sphaeroides*, which have 4 and 12 complete TRAP transporters, respectively. It has been shown that Na⁺ is essential for growth of *R. sphaeroides* (Sistrom, 1960). The data strongly suggest that marine-dwelling organisms generally have high numbers of TRAP transporters. A possible reason for this is that TRAP transporters function better in high NaCl environments.

As well as this lineage-independent correlation, this analysis has also revealed an example of lineage-dependent expansion of TRAP transporters. The organisms, *Bordetella bronchiseptica* RB50 NCTC-1325, *Bordetella parapertussis* 12822 NCTC-1325 and *Bordetella pertussis* Tomaha I NCTC-13251 all have above average TRAP densities. These pathogens are also massively enriched for ESRs from the TTT transporter family (Antoine *et al.*, 2003). *B. pertussis* Tomaha I NCTC-13251 has 91 ORFs for TTT ESRs, the majority of which are orphan ESRs and not associated with transport machinery. The exact reason why there is massive amplification of ESRs is unknown, however, it has been suggested that there is a correlation in this branch of the β -proteobacteria between soil-dwelling bacteria and high numbers of TTT ESRs (Antoine *et al.*, 2003).

Analysis of the frequency of TAXI-type systems and DctP-type systems revealed that in contrast to the enrichment of DctP-type TRAP transporters, the TAXI-type TRAP transporters are much more limited in number. *Aurantimonas* sp. SI85-9A1 was found to have the highest number of TAXI-type TRAP transporters with 5 systems, whereas most other prokaryotes encode a maximum of 2 TAXI-type systems. *Archeoglobus fulgidus* was found to be the archaeon with the most TRAP transporters with a total of 3, all of which are TAXI-type systems. Indeed, TAXI-type systems are the only TRAP transporters to be found in archaea.

3.3 Genetic arrangement of *dctQ* and *dctM* genes

Prior to this extensive data mining exercise, all TRAP transporter operons encoding membrane components had the small membrane component (*dctQ*) encoded upstream of the large subunit (*dctM*). In the TRAP systems with a

single ORF encoding the fused small and large integral membrane components (*dctQM*), the *dctQ* component domain was always encoded first. There are currently 627 different DctP-type systems with these gene organisations in TRAPdb. This led to the hypothesis that the *dctQ* subunit had to be expressed first as the protein DctQ was involved in the stability or insertion of DctM into the membrane. As a consequence of the data mining there are now 4 examples where the *dctM* gene is encoded upstream of the *dctQ* gene (Fig. 3.2, *Roseovarius* sp. 217). This is only 0.64% of all the DctP-type TRAP systems in TRAPdb, but these new observations suggest that the ordering of *dctQ* upstream of *dctM* is not as essential as previously thought.

3.4 TRAP transporter systems with multiple ESR components

ABC transporter operons have been discovered with multiple ESR components, for example, GlnPQ, the four ESR containing glutamine/glutamate transporter from *Lactococcus lactis* has four ESR components associated with it (Schuurman-Wolters & Poolman, 2005). The data held within TRAPdb was used to determine whether there are examples of TRAP transporters with multiple ESR components. In total, 14 DctP-type TRAP systems with multiple ESR components were discovered.

A DctP-type TRAP system in *Mannheimia succinoproducens* was found to encode a *dctQM* gene with two downstream *dctP* genes, MS0049 and MS0050 (Fig. 3.2, *M. succiniciproducens* MBEL55E). Analysis of the genome context of this system revealed that the operon contained the same genes as the operon containing *yiaMNO*, the 2,3-diketo-L-gulonate TRAP transporter of *E. coli*. Phylogenetic analysis of the DctP-type ESR components contained within TRAPdb revealed that MS0049, MS0050 and YiaO clustered together in the same clade in the phylogenetic tree (data not shown). This information and the genome context data strongly suggest that this transporter is specific for 2,3-diketo-L-gulonate, however, the reason for having two separate ESRs is unknown.

Another DctP-type TRAP system with two *dctP* genes, this time associated with a *dctQM* gene was found in *Polaromonas* sp. JS666. In this system, the ORFs for the ESRs are flanking the *dctQM* gene, one upstream and one downstream (Fig. 3.2, *Polaromonas* sp. JS666 Bpro_2838-2841). Curiously, phylogenetic

Figure 3.2 Genetic organization of selected TRAP transporters contained within TRAPDb. The layout of TRAP genes is centered on the *dctP* (black arrow [P]), *dctQ* (white arrow [Q]) and *dctM* (grey arrow [M]) genes of the *R. capsulatus* *dctPQM* operon. All DctP homologues are indicated by solid black lines, while TAXI ESRs are shaded in checkerboard [TAXI]. The malonyl-CoA decarboxylase domain fused to DctM is shaded with vertical stripes, while the malonyl-CoA synthase is shaded in dark grey with white spots. The small gene encoding a UspA homologue is shaded with diagonal checks. TRAP-linked genes (*tlg* – see (Kelly & Thomas, 2001)) are shaded with small black spots with a white background. The *glnH* gene, an ABC ESR, is shaded with grey horizontal lines. The DctRS two-component system and dicarboxylate transporter in *B. subtilis* are shaded with diagonal black stripes and horizontal black stripes, respectively. The fused helix-turn-helix domain is shaded with black horizontal waves, while the kinase fusion to a TAXI ESR is shaded with grey and small white spots. Finally, the potential ABC transporter in an operon with a DctP ESR is shaded with a grey checkerboard pattern. Other abbreviations for species names used are *E. coli* - *Escherichia coli* K12; *H. influenzae* - *Haemophilus influenzae* RD KW20; *M. succiniciproducens* - *Mannheimia succiniciproducens* MBEL55E; *D. hafniense* - *Desulfitobacterium hafniense* Y51; *A. actinomycetemcomitans* - *Actinobacillus actinomycetemcomitans*; *S. meliloti* 1021 - *Sinorhizobium meliloti* 1021; *A. fulgidus* - *Archaeoglobus fulgidus* DSM 4304; *P. aeruginosa* PAO1 - *Pseudomonas aeruginosa* PAO1; *B. halodurans* - *Bacillus halodurans* C-125; *A. succinogenes* 130Z - *Actinobacillus succinogenes* 130Z; *B. subtilis* - *Bacillus subtilis*; *M. loti* - *Mesorhizobium loti* MAFF303099; *R. baltica* SH1 - *Rhodopirellula baltica* SH 1 and *M. magneticum* AMB-1 - *Magnetospirillum magneticum* AMB-1.

analysis of these ESRs revealed that they clustered in the same clade as YiaO, MS0049 and MS0050, therefore could possibly also bind 2, 3-diketo-L-gulonate. The third example of a TRAP system with multiple ESR genes was again discovered in *Polaromonas* sp. JS666. This system contained 3 *dctP* genes separated from the *dctQ* and *dctM* genes by a gene encoding a protein of unknown function (Fig. 3.2, *Polaromonas* sp. JS666 Bpro_1753-1748). The ESRs in this system have very high sequence identity with each other, which means they probably bind the same unknown substrate and are the product of relatively recent gene duplication.

Analysis of the data held within TRAPdb revealed several TRAP systems with multiple ESR components. The multiple ESRs appear to be the product of recent gene duplication. The exact reason for multiple ESRs is unknown. However, it may be that the ESRs bind a similar, but not identical compound; therefore this duplication is a method by which to increase the repertoire of substrates.

3.5 Recruitment of TRAP ESR domains for non-transport functions

ABC transporter ESR domains are known to have been recruited for functions other than transport, for example the LacI/GalR family of repressors are proteins that have a DNA-binding helix-turn-helix domain N-terminally fused to an ABC ESR domain (Fukami-Kobayashi *et al.*, 2003). The data within TRAPdb was analysed to determine whether there are any examples of TRAP ESRs being recruited for purposes other than transport.

A DctP-type TRAP transporter system was discovered in *Desulfitobacterium hafniense* Y51 that contained two ESR genes upstream of *dctQ* and *dctM* genes. Phylogenetic analyses of these ESR genes revealed that they were both predicted to bind C₄-dicarboxylates, however, the gene DSY4256 clustered with ESRs from characterised TRAP transporters and DSY2457 clustered with YdbE (DctB), a characterised protein from *B. subtilis* shown to regulate C₄-dicarboxylate transport (see Introduction section 1.8.2.1 (Asai *et al.*, 2000)). This suggests that this TRAP system is using the ESR-fold for two different functions; transport and regulation of gene expression.

A TAXI ESR gene was discovered in *Rhodopirellula baltica* SH 1 with an N-terminal serine/threonine kinase fusion (Fig. 3.2, *R. baltica* SH 1). Analysis of the genome context revealed that the genes for the membrane components were not present in the same operon. When analysed using the web-based application, SignalP, no signal peptide is detected, indicating that it is a cytoplasmic protein. These properties imply that this orphan ESR no longer participates in transport and functions in signal transduction instead.

A DctP-type ESR has been discovered in *Mesorhizobium loti* that has a C-terminal fusion to a helix-turn-helix DNA binding domain (Fig. 3.2, *M. loti*). This is also an orphan ESR and not associated with any genes encoding DctQ or DctM, therefore, it is possible that this protein acts as a transcription factor in response to ligand binding. This is analogous to the previously mentioned LacI/GalR family of repressors that contain ABC ESR domains fused to a DNA binding domain; however, this is the first incidence of a TRAP ESR being employed in this manner.

Three examples of TRAP ESRs possibly being recruited for purposes other than transport have been discovered. This is a phenomenon also seen with ABC ESRs. Indeed, these proteins are well-suited for such activities as they bind ligands with high affinity which also induces a conformational change which could be recruited for a signalling function. Sequencing of more prokaryotic genomes will no doubt reveal more examples of this kind of recruitment.

3.6 UspA gene linked to TAXI TRAP transporter genes

Analysis of the genome context of TRAP transport systems held within TRAPdb has revealed genes that are regularly found in close proximity to TRAP transporters. These genes are often genes for enzymes that catabolise the substrate that the TRAP system transports into the cell, in this way, substrate specificities can be predicted like, for example, *N*-acetylneuraminic acid-specific *SiaPQM* from *Haemophilus influenzae* (Severi et al., 2005). Some genes, for example, the TRAPlinked genes (*t/g*) have been found in the same operons as TRAP transporters in many organisms with no discernable function (Kelly &

Thomas, 2001). Analysis of the TAXI TRAP systems has revealed a linkage to a gene for a universal stress protein A-like protein (UspA).

A TAXI TRAP system was discovered in *Actinobacillus succinogenes* 130Z with two genes for TAXI ESRs with the gene for the permease component (Fig. 3.2, *A. succinogenes* 130Z). The two ESRs were found to flank *dctQM* with the downstream ESR gene separated from the permease by a gene encoding a UspA-like protein.

Genes encoding the UspA protein have been found in the same operon as a number of TRAP transporters from various species, however, the reason for the correlation is unknown. This is not the first time UspA have been associated with a transporter. UspA domains have also been discovered fused to a cationic amino acid transporter from the APC (amino acid-polyamine-organocation) superfamily (Kvint *et al.*, 2003). This particular transporter has 2 UspA domains fused to it, however, the exact function of this fusion is not known. Studies on UspA have determined that its expression is induced by DNA damage under a number of conditions including starvation, UV light treatment and high salt concentration (Kvint *et al.*, 2003). The crystal structure of UspA from *H. influenzae* has been elucidated and revealed to be a homodimer (Sousa & McKay, 2001); however, the true function of UspA remains unknown.

3.7 A TRAP ESR functioning with an ABC transporter?

High resolution crystal structures are now available for a number of ABC ESRs, DctP-type TRAP ESRs and one TAXI-type TRAP ESR (see Introduction). These three families of ESRs share no sequence similarity, but have been shown to have a similar overall topology. This similarity of topology possibly suggests a shared ancestry between TRAP ESRs and ABC ESRs (Muller *et al.*, 2006). In light of this shared ancestry, it is possible that TRAP ESRs can functionally interact with ABC translocation machinery, and *vice versa*. There is an example in the literature of a glutamate-specific TRAP transporter in *Synechocystis* that uses a GlnH-like ABC ESR (Quintero *et al.*, 2001). This example and the fact that DctP-type ESRs and TAXI-type ESRs have no sequence similarity, but functionally interact with closely related homologues of the TRAP translocation machinery (DctQ and DctM) demonstrates that ESR promiscuity is not an unreasonable assumption.

Analysis of the genome context of a DctP-type TRAP ESR from *M. magneticum* AMB-1 revealed this gene to be in an operon with genes for an ABC transporter permease and ATPase subunits. In the absence of an ABC ESR gene, this suggests that this TRAP ESR may function with ABC transporter translocation machinery.

3.8 A TRAP transporter for malonate

Sinorhizobium meliloti 1021 is a symbiotic organism that resides within root nodules of the leguminous alfalfa plant (*Medicago sativa*) and has 15 predicted TRAP transporters. Curiously, one has been found to have a malonyl-CoA decarboxylase (MCD) enzyme fused to the C-terminus of the DctM component. The operon contains a DctP-type ESR gene, a *dctQ* component gene, the aforementioned *dctM*-MCD fusion gene and a gene encoding a putative malonyl-CoA synthetase. Analysis of closely related organisms revealed conservation of the association of the MCD and malonyl-CoA synthetase genes with TRAP transporter genes, although none of the other examples have this fusion protein (Fig. 3.3). One orthologous TRAP system in *Dechloromonas aromatica* RCB has a variation in that it is linked to a multisubunit MCD similar to that seen in *Pseudomonas putida* (Chohnan *et al.*, 1999), but structurally unrelated to the MCDs found with orthologous TRAP transporters. This is significant, as it demonstrates a close linkage of the TRAP transporter to a completely different type of MCD enzyme. There is also an example in *Rhizobium trifolii* where the MCD and the malonyl-CoA synthetase genes are associated with a dicarboxylate transporter (Fig. 3.3, MatC) from the Dcu family (An & Kim, 1998). Strengthening the linkage between the presence of a transporter gene and malonate catabolic genes.

Malonate is a major organic acid in the cytoplasm of the alfalfa plant cells, the bacteroids and the roots of the plants. It is thought to be an important carbon source for the bacteroids and disruption of the malonate catabolic genes has been shown to disrupt the formation of mature root nodules (An *et al.*, 2002). When taken into the bacteroid, malonate is first converted to malonyl-CoA through the action of malonyl-CoA synthetase and can then be incorporated into central metabolism as acetyl-CoA through the action of malonyl-CoA decarboxylase (An & Kim, 1998). Malonate is known to be a competitive

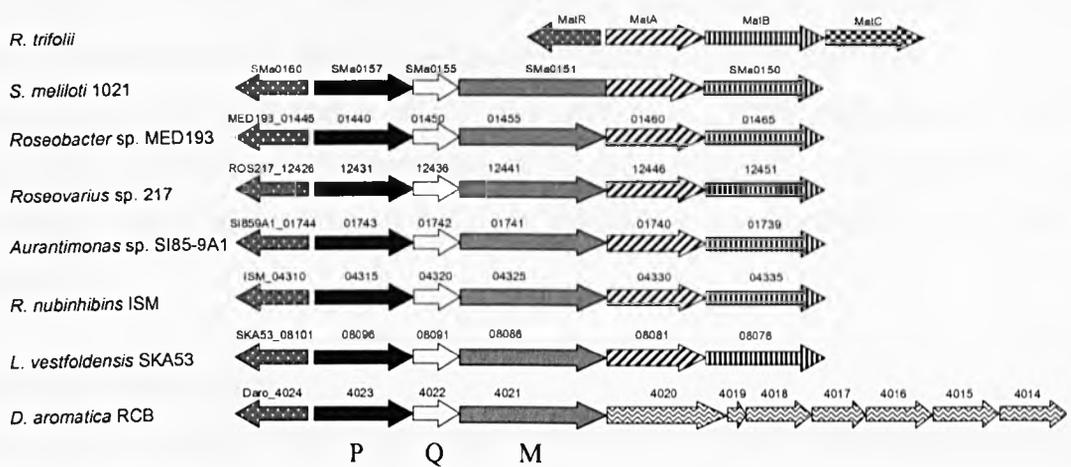


Figure 3.3 Organisation of genes predicted to encode a TRAP transporter for malonate and related malonate metabolism-related genes. The TRAP subunits are shaded as in Fig. 3.2 with the transcriptional repressor (*matR*) shaded in grey with white spots. The malonyl-CoA decarboxylase fused to DctM in *Sinorhizobium meliloti* 1021 is shaded with diagonal stripes and the malonyl-CoA synthetase with vertical stripes. In *Rhizobium leguminosarum* biovar. *trifolii* the *matC* gene that encodes a Dcu-type transporter is shaded with a checkerboard pattern. The multiple subunits of the alternative malonate decarboxylase seen linked to the orthologous TRAP system in *Dechloromonas aromatica* RCB is shaded with a horizontal wave pattern. Other species abbreviations used are: *R. nubinhibens* ISM - *Roseovarius nubinhibens* ISM; *L. vestfoldensis* SKA53 - *Loktanella vestfoldensis* SKA53.

inhibitor of succinate dehydrogenase, and so will need to be metabolised quickly to avoid any deleterious effects. It is therefore tempting to speculate that the TRAP transporter is malonate specific, but because of the potential toxicity of malonate it has to be processed quickly, necessitating the fusion of MCD to the transporter. However, one would have to wonder why the synthetase gene was not fused as it encodes the first enzyme in the catabolic pathway.

Subsequent to this finding, a study was performed which demonstrated that malonate could induce the expression of the ESR of this system (Mauchline et al., 2006), further strengthening the case that this is a malonate specific TRAP transporter.

3.9 Chapter summary

This results chapter details the compilation of a TRAP transporter-specific relational database, TRAPdb. Analysis of the data collected has provided a number of new insights. Analysis of the distribution of TRAP transporters throughout 243 different prokaryotic organisms revealed that marine-dwelling organisms have a propensity to have higher than average numbers of TRAP transporters. Where data was available for an organism (from TransportDB), the overall contribution of TRAP transporters to the total number of transporters was calculated. In the marine-dwelling organisms enriched for TRAP transporters this was as high as 10%.

Analysis of the data collected for TRAPdb also revealed a number of new TRAP gene organisations including the hitherto unseen positioning of the *dctM* gene upstream of the *dctQ* gene and the *dctQ* and *dctM* genes associated with up to 3 *dctP* genes. A number of TRAP ESR genes with the potential for non-transport functions have been identified such as a *ybdE*-like gene and a *dctP* gene fused to a serine/threonine kinase, both possibly participating in signalling processes instead of transport.

A *dctP* gene has been found associated with the genes encoding the transmembrane domains and nucleotide-binding domains of an ABC transporter, suggesting that this ESR can form a functional transporter with ABC transporter machinery.

Analysis of phylogenetic and genome context data has led to the prediction that a TRAP transporter from *S. meliloti* is specific for malonate. The *dctM* gene in

this system has also been found to be fused to a malonate decarboxylase enzyme, further strengthening this association.

Chapter 4

Expression and purification of YiaM and YiaN from *E. coli*

4.1 Cloning and expression trial of the *yiaMN* genes using pTTQ18

To characterise functionally the integral membrane proteins from the *E. coli* TRAP transporter YiaMNO, it was necessary to clone the genes *yiaM* and *yiaN* into a plasmid vector suited for overexpression of integral membrane proteins. The pTTQ18 expression vector was chosen as this has been used to overexpress a number of different integral membrane proteins, including other secondary transporters (Morrison et al., 2003, Saidijam et al., 2005, Saidijam et al., 2003, Ward, 2000).

The function of the small subunit YiaM is unknown and may be necessary for the correct folding/insertion of the large subunit, YiaN, into the membrane. To take this possibility into consideration, three different expression constructs were produced; *yiaM* alone, *yiaN* alone and the co-expression of the *yiaMN* genes.

Regions of the *E. coli* K-12 genome were amplified by PCR using primers YiaMF and YiaMR for *yiaM* amplification, YiaNF and YiaNhisR for *yiaN* amplification and YiaMF and YiaNhisR for *yiaMN* amplification. These were cloned into pTTQ18-RGSH₆ on an *EcoRI*-*PstI* fragment so that the coding sequences were in frame with the 3'-nucleotide sequence encoding the amino acids arginine, glycine serine and 6 histidines (RGSH₆). The chosen clones were named; pCM14 (YiaM alone), pCM19 (YiaN alone) and pCM20 (YiaMN co-expressed). All expression vectors were then transformed into the *E. coli* expression strain BL21 (DE3) and used in initial expression trials (Methods section 2.7.1).

A growth curve of induced and uninduced BL21 (DE3) pCM14 cultures indicated that there was a marked decrease in the growth rate when inducer was added compared to where inducer is omitted (Fig. 4.1). It is known that the overexpression of integral membrane proteins can lead to decreased growth rates and this is often a good indication of whether overexpression is occurring (Miroux & Walker, 1996).

Analysis of the steady-state levels of YiaM from the samples taken during the initial expression trial using Western blotting indicated that in the presence of

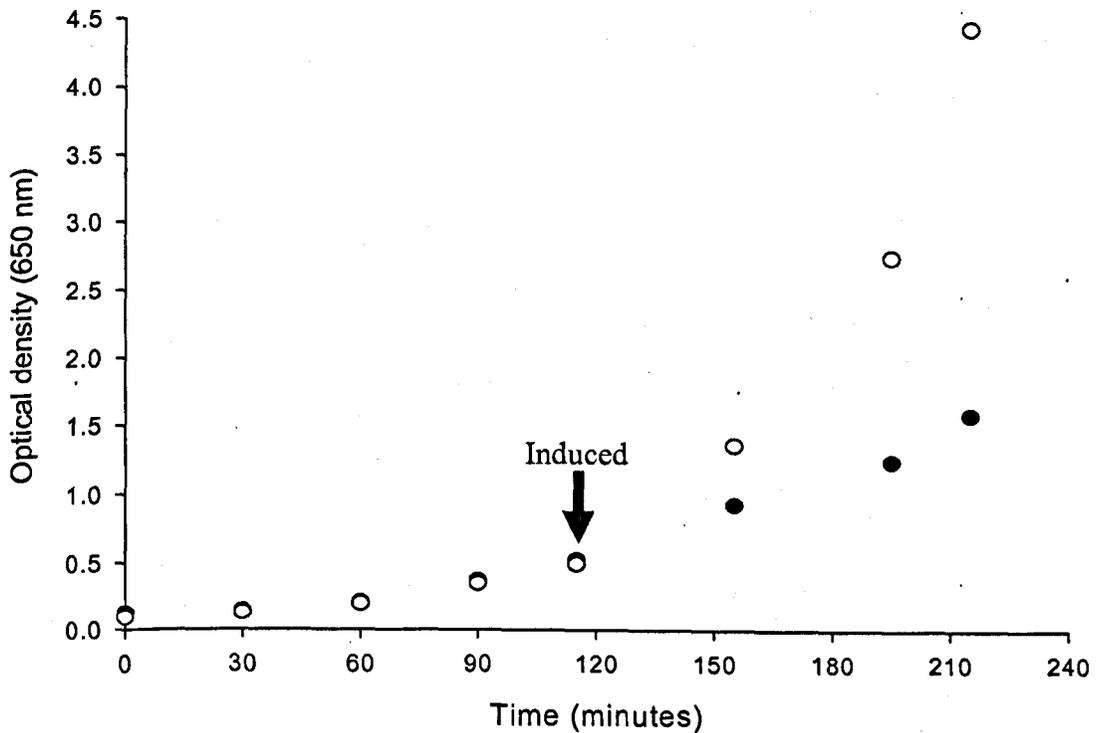


Figure 4.1. Effects of induction with IPTG on the growth of BL21 (DE3) pCM14. Optical densities (OD_{650}) of two BL21 (DE3) pCM14 cultures grown aerobically in LB (supplemented with 20 mM glycerol and 100 μ g/ml carbenicillin) to mid-log phase (OD_{650} of 0.4-0.6) and either induced with 0.5 mM IPTG (closed circles) or left uninduced (open circles) over a period of ~3.5 hours.

IPTG induction an immunoreactive protein with an apparent molecular weight of ~16-18 kDa is produced from the strain BL21 (DE3) pCM14 (YiaM) (Fig. 4.2a and b). This is in range of the predicted molecular weight of YiaM (18 kDa). This immunoreactive species was not present in the membrane sample from the uninduced culture.

A ~34 kDa immunoreactive species was present in the membrane fractions of the strains BL21 (DE3) pCM19 (YiaN alone) and BL21 (DE3) pCM20 (YiaMN), in the presence of inducer (Fig. 4.2d). This molecular weight is slightly smaller than the size predicted for YiaN (45.4 kDa), however it is commonly observed that membrane proteins migrate aberrantly fast in SDS-PAGE gels. There are 2 key differences observed between YiaN expressed on its own and YiaN co-expressed with YiaM. There are higher steady-state levels of YiaN accumulation after 3 hours growth when co-expressed with YiaM and there is also a ~27 kDa species detected when YiaM and YiaN are co-expressed which is not detected otherwise. However, the absence of this second band may be a consequence of the generally low expression from BL21 (DE3) pCM19.

It was subsequently discovered that the YiaMN construct (pCM20) used in this experiment had a mutation in *yiaN*. This mutation changed a small neutral glycine in the tenth predicted transmembrane helix into a bulkier charged aspartate residue. A wildtype YiaMN construct was subsequently produced (pCM22) and a parallel expression trial was performed on both the G→D YiaMN mutant and the wildtype YiaMN. Western blot analysis showed that the Western signal for the wildtype protein was significantly more intense than the signal for the G→D mutant (Fig. 4.2e). This result indicates that the difference in steady-state levels of YiaN in the presence or absence of YiaM co-expression is even more pronounced than previously presented (Fig. 4.2d).

These initial expression trials have shown that the pTTQ18 expression system is sufficient for the expression of YiaM and YiaN to detectable levels. YiaM, when expressed alone, produces a 16-18 kDa immunoreactive species and at longer exposure levels, higher molecular weight species are detectable, the most pronounced of which is around 30 kDa, the approximate size of a YiaM dimer. These higher molecular weight species may be the result of aggregation of YiaM protein due to the high expression levels or may be genuine oligomer formations.

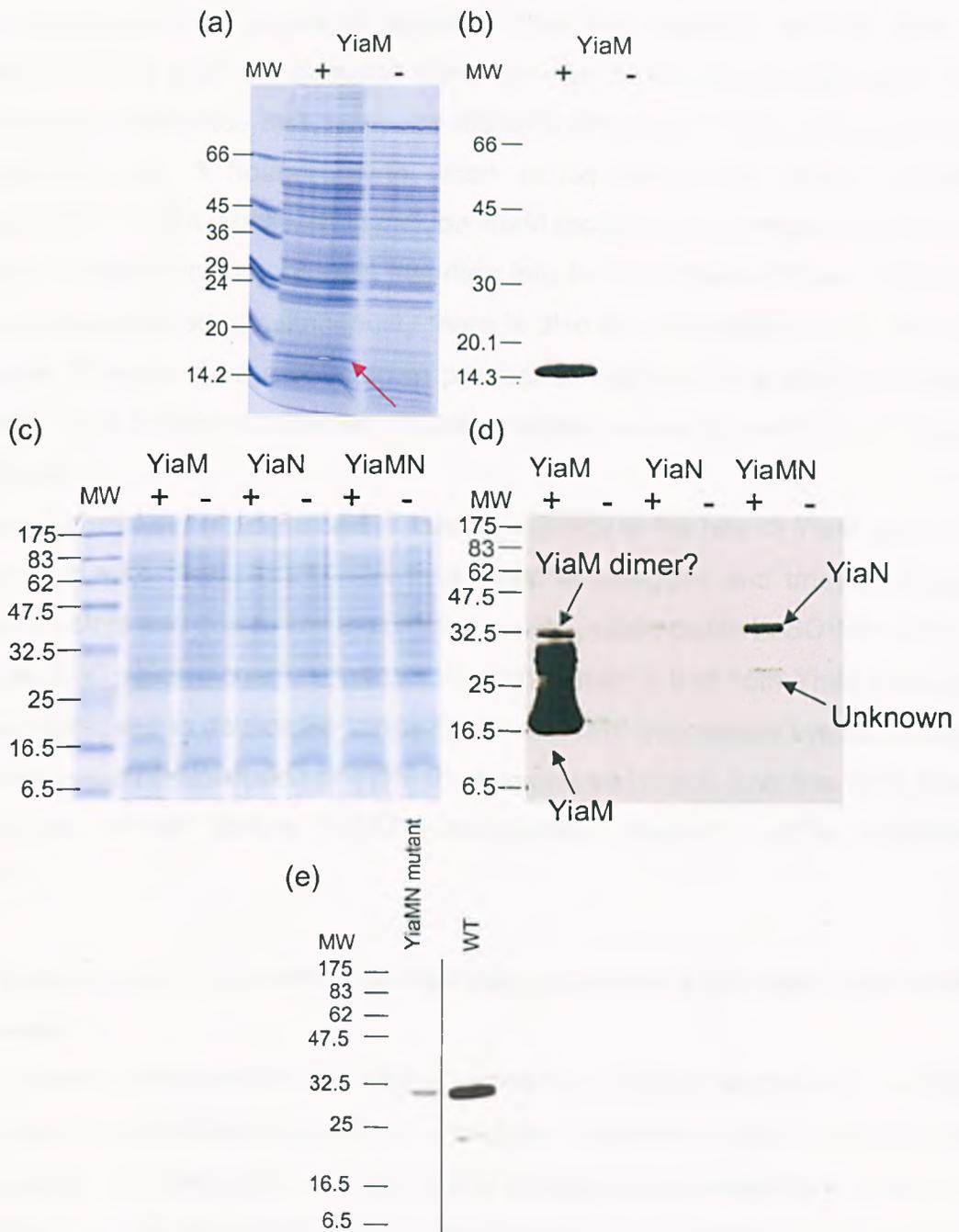


Figure 4.2 Expression of YiaM and YiaN from pTTQ18 based expression constructs in the presence (+) and absence (-) of IPTG induction. (a) Coomassie stained SDS-PAGE and (b) Western blot with anti His-tag antibody of YiaM expression from BL21 (DE3) pCM14. The identity of YiaM confirmed by peptide mass fingerprinting (red arrow). (c) Coomassie stained gel and (d) Western blot of YiaM expression from BL21 (DE3) pCM14, YiaN expression from BL21 (DE3) pCM19 and YiaMN expression from BL21 (DE3) pCM20. (e) Western blot of YiaMN expression from BL21 (DE3) pCM20 ("YiaMN mutant") and BL21 (DE3) pCM22 ("WT"). Molecular weight (MW) shown in kDa.

An immunoreactive species of approximately the predicted size of YiaN is produced when *yiaN* is expressed alone or when co-expressed with *yiaM*. The Western blot indicates that there are higher steady-state levels of YiaN in the membrane after 3 hours growth when co-expressed with YiaM. Possible explanations of this observation include YiaM increasing the stability of YiaN or increasing the efficiency of YiaN insertion into the membrane. When YiaM and YiaN are expressed simultaneously there is also the appearance of a ~27 kDa species. This could be a breakdown product of YiaN due to specific peptidase activity, or it could be a protein non-specifically interacting with the Western antibody.

Another question not addressed in this experiment is the fate of YiaM when co-expressed with YiaN. In this construct YiaM is untagged and undetectable in Western blots and due to low expression is also undetectable in SDS-PAGE.

The most significant conclusion from this experiment is that both YiaM and YiaN were expressed to detectable levels by the pTTQ18 expression system. Due to the low levels of expression of YiaN when expressed alone, only the YiaM alone (pCM14) and the YiaMN (pCM22) expressing constructs will be optimised further.

4.2 Optimisation of growth and induction conditions for YiaM and YiaMN expression

To maximize the amount of integral membrane protein produced it is often important to optimise the growth and induction conditions. A number of common parameters are taken into account, these include: the concentration of inducer used, the growth temperature, the growth medium and the length of time after induction that the cells are harvested. Each of these parameters was varied individually for the strains expressing YiaM alone or co-expressing YiaMN.

4.2.1 IPTG concentration optimisation

To assess the affect of varying the inducer concentrations on the expression of YiaM and YiaN, 6x50 ml bacterial cultures of the strains BL21 (DE3) pCM14 and BL21 (DE3) pCM22 were grown as in the initial expression trials (Methods section 2.7.1) except that various concentrations of IPTG were added; 0 mM, 0.1 mM, 0.25 mM, 0.5 mM, 0.75 mM and 1 mM.

Analysis of the steady-state levels of YiaM (Fig. 4.3a) and YiaN (Fig. 4.3b) using Western blotting indicated that varying the concentration of IPTG between 0.1 mM and 1 mM did not lead to any significant variation in the total amount of YiaM or YiaN in the membrane fractions after 3 hours of growth. There is no signal on the Western blot in the absence of inducer.

The results of this experiment reveal that any concentration between 0.1 and 1 mM IPTG will produce approximately the same amount of protein. Induction from the lactose promoter is not regarded as particularly tunable when compared to other promoters like the arabinose-inducible pBAD system. There is evidence in the literature that varying the concentration of IPTG can alter expression levels (Londer *et al.*, 2006); however, there was no dramatic difference in expression levels in the concentration range that was tested here. Increasing this range may have produced some differences in expression levels. The Western blot shows no accumulation in the absence of inducer indicating that the promoter does not have a background level of constitutive activity. This is advantageous when expressing membrane proteins as even a small amount of expression can lead to decreased growth rate or cell death. Subsequent optimisation experiments will use a concentration of 0.5 mM IPTG for both expression strains being tested.

4.2.2 Growth medium, temperature of growth and harvest time

The growth medium and the length of time after induction that the cells were harvested were the next parameters to be optimised. The initial expression trial indicated that accumulation of YiaN was lower than YiaM (Fig. 4.2d and e); therefore, in an attempt to increase the accumulation of YiaN, the temperature of growth was also optimised. This parameter was varied simultaneously with medium and harvest time for BL21 (DE3) pCM22. A combination of the signal produced by the Western blot and the optical density at the time of harvesting was used to determine the optimal growth conditions.

A standard timecourse experiment was performed for BL21 (DE3) pCM14, at 37°C, and two were performed for BL21 (DE3) pCM22, one at 37°C and the other at 25°C (Methods section 2.7.4). Mixed membrane fractions were prepared using the water lysis method and the accumulation of protein was

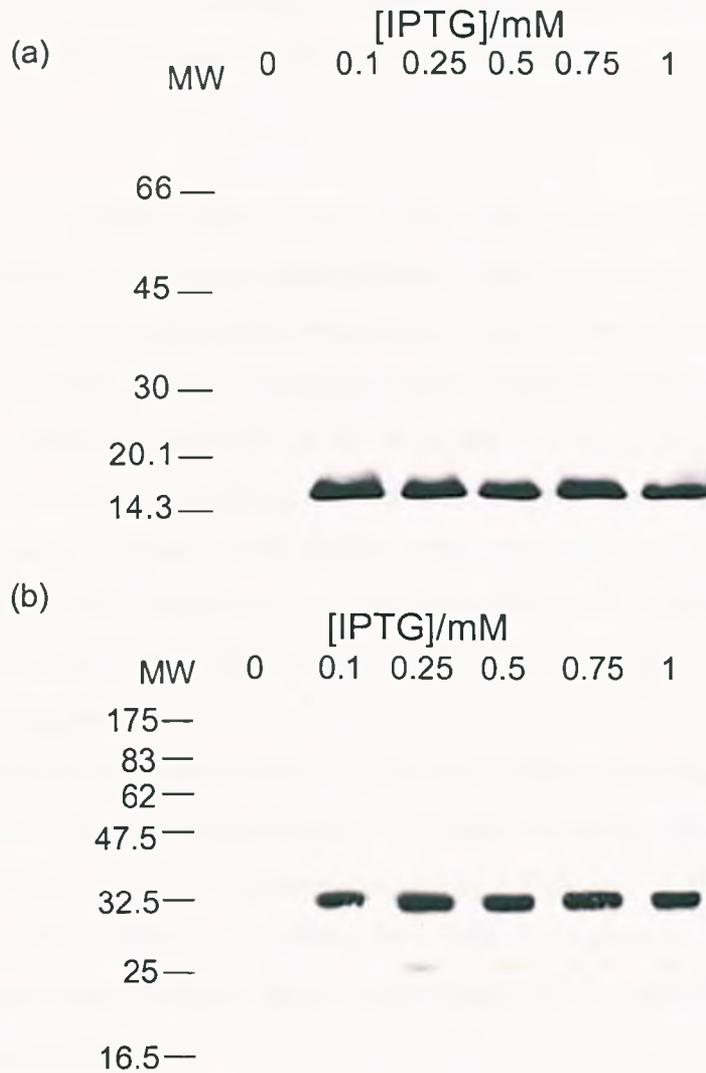


Figure 4.3 Effect of varying IPTG concentration on the accumulation of YiaM and YiaN after 3 hours growth. Western blot with anti His-tag antibody of (a) YiaM accumulation from BL21 (DE3) pCM14 and, (b) YiaN accumulation from BL21 (DE3) pCM22 in the presence of increasing concentrations of IPTG. Each culture was grown in LB, induced at mid-log phase with IPTG and harvested after a further 3 hours growth. 5 μ g of protein was used in the Western blots. Molecular weight (MW) shown in kDa.

analysed using Western blotting. Protein samples were also analysed on SDS-PAGE gels to ensure that samples had equal amounts of protein loaded (data not shown).

Analysis of the steady-state levels of YiaM using Western blotting indicated that accumulation of YiaM was approximately 2-fold higher in M9 minimal medium than LB and in both media the signals on the Western blots peaked at 3 hours after induction (Fig. 4.4a). At 3 hours after induction the OD₆₅₀ of the M9 culture was 0.882, whereas the LB culture reached an OD₆₅₀ of 1.310. Although the highest OD₆₅₀ was achieved by the LB culture grown overnight, this sample had a relatively weak signal on the Western blot, indicating that harvesting overnight culture would not result in the highest yield of YiaM attainable. These observations indicated growth on LB and harvesting after 3 hours was optimal for YiaM production.

Analysis of the steady-state levels of YiaN in the samples taken from the BL21 (DE3) pCM22 cultures using Western blotting indicated that there were very low levels of YiaN in the M9 cultures grown at 37°C (Fig. 4.4bi), and no Western blot signal was visible in samples from the LB cultures grown at 37°C (Fig. 4.4bii). These two Western blots signify that 37°C should not be used as the temperature for growth.

Analysis of the steady-state levels of YiaN when BL21 (DE3) pCM22 was grown in M9 at 25°C indicated YiaN accumulation peaked after 2 hours and then decreased steadily. There is no Western blot signal for YiaN in the sample taken after overnight growth (Fig. 4.4biii). When grown in LB at 25°C, YiaN accumulation reached its peak after 3 hours at which point the steady-state amount levelled out resulting in the samples taken at 3, 4 and 5 hours after induction having similar Western blot signal intensity (Fig. 4.4iv). The intensity of Western blot signals in these samples were higher than the intensity of the peak signal in the samples from the M9 culture and the optical densities were almost double, therefore, growing BL21 (DE3) pCM22 in LB at 25°C was identified as the growth conditions that would result in the highest accumulation of YiaN.

These experiments have highlighted the subjective nature of membrane protein expression conditions. YiaM is expressed well in M9 and LB at 37°C, however,

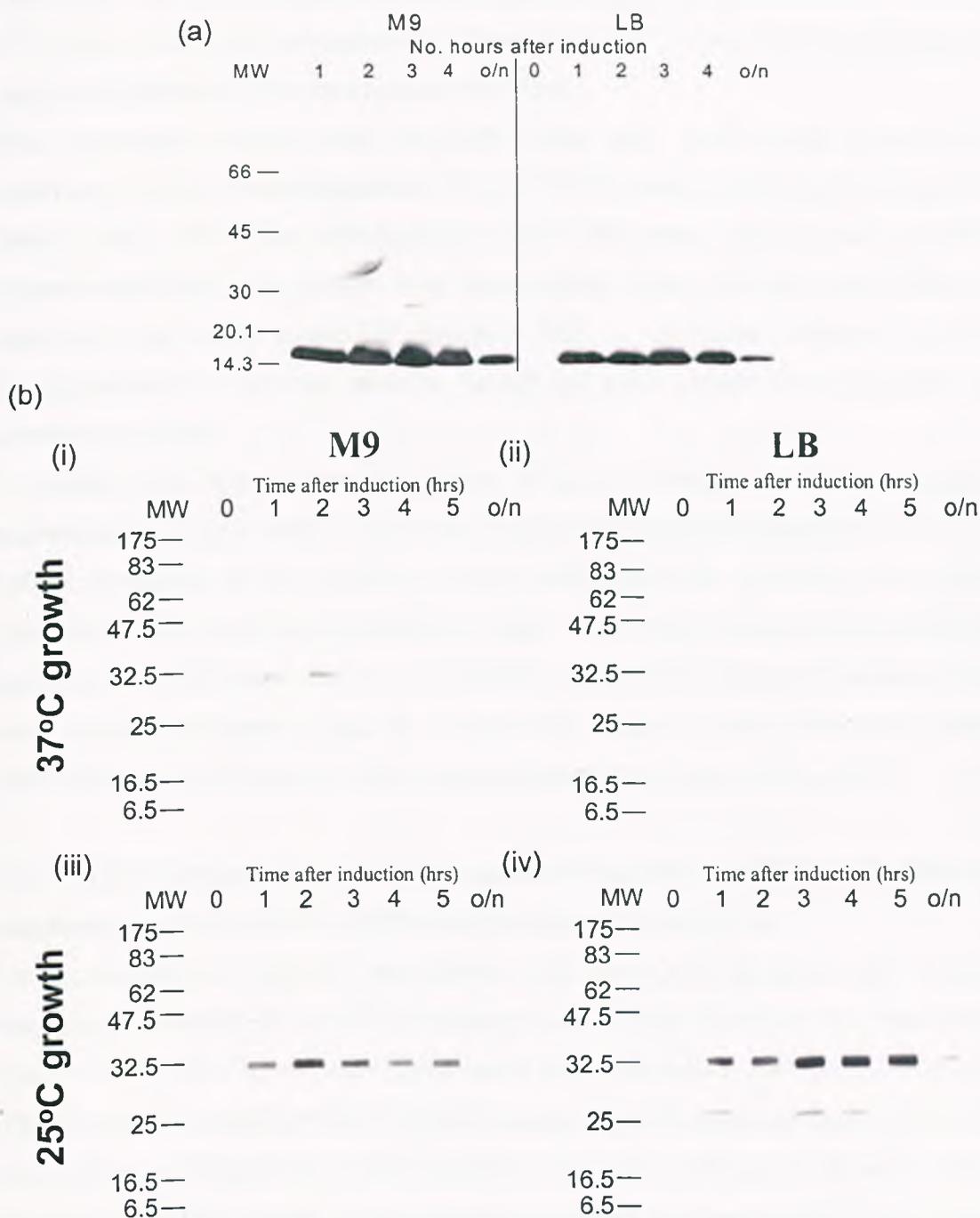


Figure 4.4 Optimisation of time of harvest after induction, medium used and temperature of growth for the production of YiaM and YiaN. Western blots of membrane fractions from samples harvested at hourly intervals after induction. **(a)** YiaM accumulation in BL21 (DE3) pCM14 when grown in LB and M9 at 37°C. **(b) (i)** YiaN accumulation from BL21 (DE3) pCM22 when grown in (i) M9 at 25°C, **(ii)** LB and 25°C, **(iii)** M9 at 37°C and **(iv)** LB at 37°C. 5 µg of protein was used for all samples in the Western blots. Molecular weight (MW) shown in kDa.

YiaN does not accumulate in the membrane to very high levels in LB or M9 at 37°C. Only when the temperature is lowered to 25°C does YiaN accumulation become detectable in the membrane fractions.

The expression timecourses for both YiaM and YiaN have shown that membrane protein overexpression from pTTQ18 peaks after only a couple of hours, after which the steady-state levels decrease quite rapidly. These experiments have both shown that steady-state levels of YiaM and YiaM are relatively low after overnight growth; this is in stark contrast to the overexpression of soluble proteins where 24 hour growth after induction is common practise.

A problem that arises from the lethal effects of integral membrane protein expression combined with the membrane proteins apparent instability is that the optical densities of the cultures when harvested are generally low. One parameter that was not investigated was the optical density at which the cultures were induced. All the experiments described were with cultures that were induced between OD₆₅₀ of 0.3 and 0.5. Higher protein yields may have been obtained if the cultures had been induced later at, say, OD₆₅₀ of 1.

4.3 Comparative expression timecourse of YiaN from pCM22 in 3 different expression strains; BL21 (DE3), C41 (DE3) and C43 (DE3).

To overcome the toxicity associated with the overexpression of integral membrane proteins, BL21 (DE3) mutants with higher tolerance for membrane protein expression have been discovered and developed. C41 (DE3) and C43 (DE3) were derived from BL21 (DE3) by selecting for lower susceptibility to the toxic effects of integral membrane protein expression (Miroux & Walker, 1996). These strains have been used a number of times in parallel with BL21 (DE3), usually in high throughput approaches, to try and increase the yield of integral membrane protein expression with varying degrees of success (Dumon-Seignovert *et al.*, 2004). Parallel timecourse experiments were performed to see if these strains could increase the yield of YiaN beyond that obtained with BL21 (DE3).

The plasmid pCM22 was transformed into C41 (DE3) and C43 (DE3) and standard timecourse experiments (Methods section 2.7.4) were performed for all 3 strains using the optimised conditions for accumulation of YiaN from BL21

(DE3) pCM22 as determined previously. Membrane fractions were prepared using the water lysis method.

Analysis of YiaN accumulation using Western blotting indicated that YiaN is expressed in all three expression strains; BL21 (DE3), C41 (DE3) and C43 (DE3). The accumulation of YiaN in BL21 (DE3) pCM22 peaked 2 hours after induction with an optical density of 0.804 (Fig. 4.5a), this follows the same pattern as seen in the optimisation experiment (Fig. 4.3 and 4.4). Accumulation of YiaN in C41 (DE3) pCM22 and C43 (DE3) pCM22 reached its maximum after one hour and the signal remained almost constant in the subsequent timepoints (Fig. 4.5b and c). The Western blot signals from the C43 (DE3) pCM22 samples were less intense than the signal obtained from BL21 (DE3) and C41 (DE3) samples indicating lower levels of accumulation. There is also an aberrant signal in the Western blot for the C43 (DE3) samples that cannot be explained; however, it is unlikely to be a genuine signal from the sample because the signal is not aligned with a lane in the gel (Fig. 4.5c).

Although the peak Western blot signals in the BL21 (DE3) and C41 (DE3) timecourse are of similar intensity, the highest optical density of C41 (DE3) pCM22 is achieved 4 hours after induction but is only 0.744, compared to an OD_{650} of 0.804 after only 2 hours growth of BL21 (DE3) pCM22. C43 (DE3) pCM22 has the highest optical densities of all 3 strains, reaching 1.902 4 hours after induction it also has the least intense signals on the Western blots.

Using C41 (DE3) and C43 (DE3) to try and increase the steady state levels of integral membrane proteins has been used in a number of different studies with varying degrees of success. The comparative expression timecourses shown here indicate that accumulation of YiaN is not significantly increased over that observed from BL21 (DE3) pCM22. As the growth and induction conditions for YiaN production have already been optimised for BL21 (DE3) pCM22 this strain will continue to be used. YiaN appears to be more stable in C41 (DE3) as the signal does not decrease during the timecourse, whereas, evidence of decreasing YiaN levels is seen after 3 hours in BL21 (DE3). Unfortunately, overnight samples from these cultures were not collected as it was assumed that steady-state levels would be low as seen previously. It would be interesting

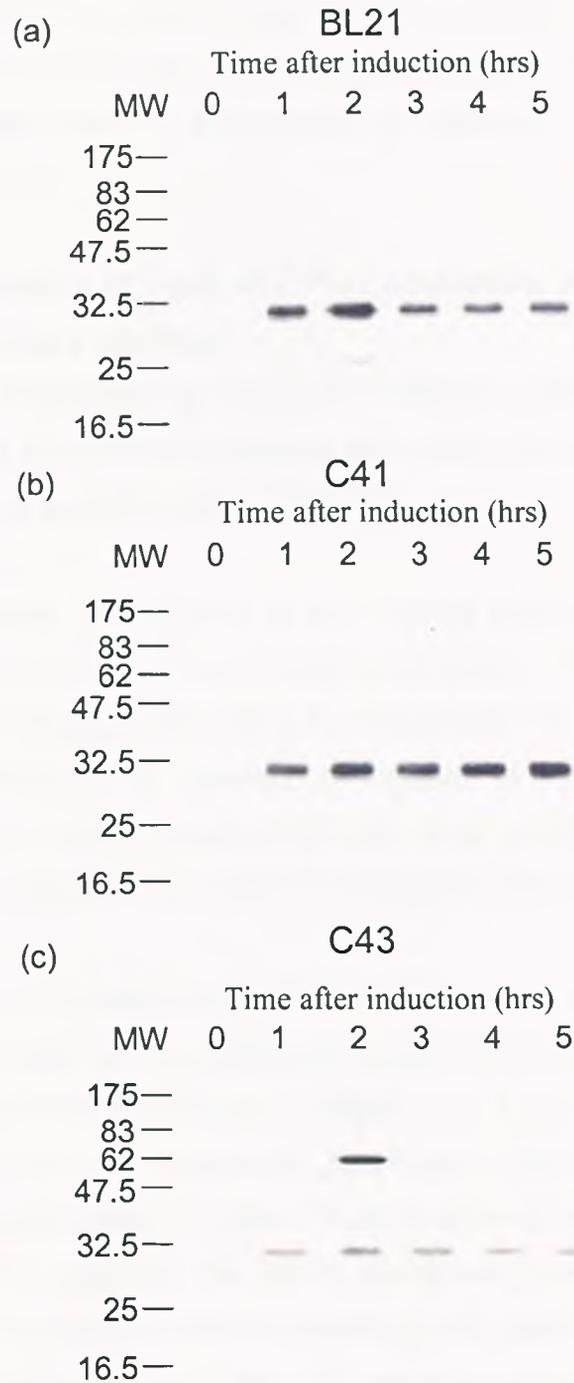


Figure 4.5 Comparative timecourses of YiaN expression from BL21 (DE3) pCM22, C41 (DE3) pCM22 and C43 (DE3) pCM22. Western blot of membrane fractions prepared from samples taken hourly after induction from the strains **(a)** BL21 (DE3) pCM22, **(b)** C41 (DE3) pCM22 and **(c)** C43 (DE3) pCM22. Cultures were grown at 25°C in LB. 5 µg of protein was used for all samples in the Western blots. Molecular weight (MW) shown in kDa.

to see whether this apparent increase in the stability of YiaN when expressed in C41 (DE3) continued into higher optical densities. If this were the case then perhaps higher yields of YiaN would be possible by using C41 (DE3) as the expression strain.

4.4 Determination of YiaM and YiaN subcellular location by separation of *E. coli* membrane fractions

As a means to confirm that YiaM and YiaN are localized to the inner membrane and also as a preliminary purification step, the inner and outer membranes were separated by a sucrose density gradient.

A larger volume, typically 5 L, of BL21 (DE3) pCM14 and BL21 (DE3) pCM22 cultures were grown and induced according to the optimised conditions mentioned previously. The mixed membrane fraction was prepared and applied to the sucrose density gradient as detailed in Methods section 2.8.4. The protein content of the mixed membranes, inner membrane and outer membrane fractions were analysed on SDS-PAGE and Western blots.

The Coomassie stained SDS-PAGE of the mixed, inner and outer membrane fractions for both of the cultures indicate significant difference in the protein content between the 3 different samples (Fig. 4.6a and c) demonstrating good separation of the inner and outer membrane. The Western blots of the same samples show signals for YiaM or YiaN in each of the membrane preparations; however, the signal in the inner membrane sample is the most intense indicating enrichment of the proteins (Fig. 4.6b and d). There is also a Western signal for YiaM and YiaN in the outer membrane fraction samples; this is most likely due to cross-contamination with inner membrane during the separation process.

Separating the inner and outer membranes has led to the confirmation that YiaM and YiaN are indeed localized to the inner membrane in *E. coli*. It has also revealed itself to be a method by which the inner membrane containing overexpressed protein can be separated from a lot of potential contaminants shown in the outer membrane lanes in the SDS-PAGE gels (Fig. 4.6a and c).

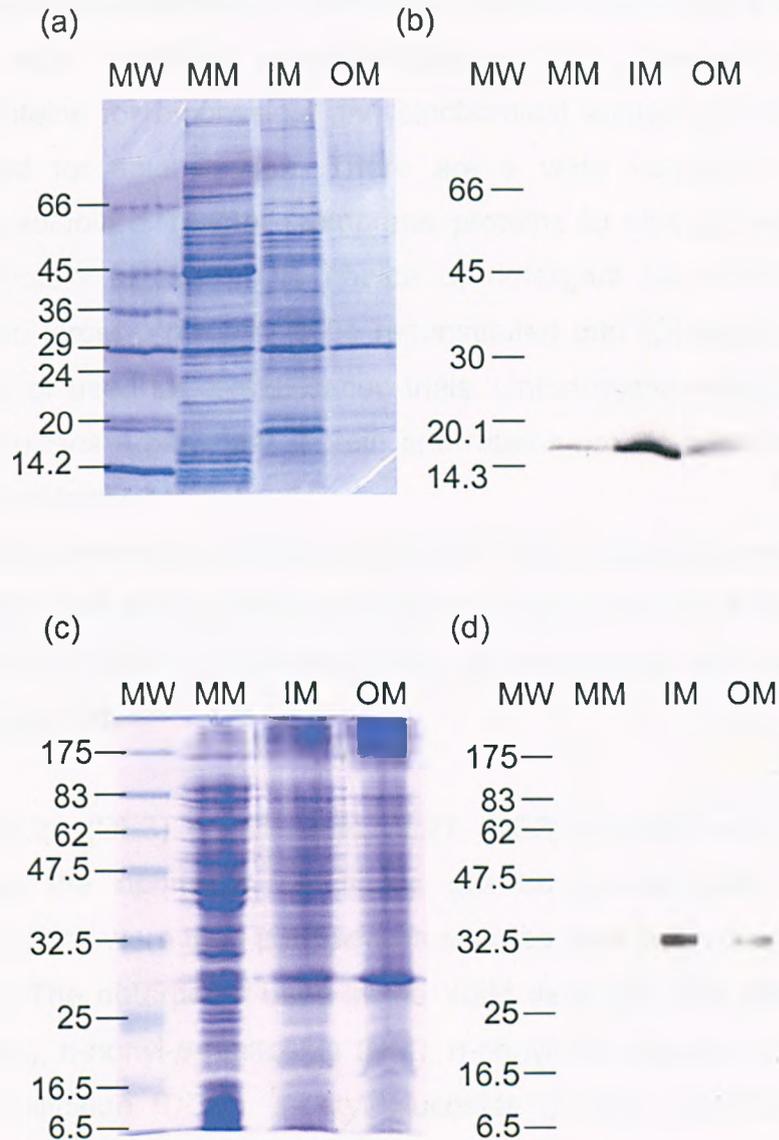


Figure 4.6 Comparison YiaM and YiaN content of mixed, inner and outer membrane fractions prepared from BL21 (DE3) pCM14 or BL21 (DE3) pCM22 cultures. (a) Coomassie stained gel and (b) Western blot of mixed (MM), inner (IM) and outer (OM) membranes produced from a culture of BL21 (DE3) pCM14 grown in LB to mid-log phase, induced with 0.5 mM IPTG and harvested after a further 3 hours growth. (c) Coomassie stained gel and (d) Western blot of mixed (MM), inner (IM) and outer (OM) membranes produced from a culture of BL21 (DE3) pCM22 grown in LB to mid-log phase, induced with 0.5 mM IPTG and harvested after a further 2 hours growth. 20 μ g of protein is applied to the Coomassie stained SDS-PAGE gels (except for the OM lane in (b) which was too dilute and had the maximum volume possible loaded) and 5 μ g of protein is applied to the Western blots. Molecular weight (MW) shown in kDa.

4.5 Differential solubilisation of YiaM and YiaN by a variety of detergents

One of the most important considerations in the expression of integral membrane proteins for biophysical and biochemical analysis is the choice of detergent used for solubilisation. There are a wide variety of detergents available that solubilise integral membrane proteins to varying degrees while maintaining protein function. The choice of detergent also becomes more important if the target protein is to be reconstituted into liposomes for *in vitro* uptake studies or used for crystallisation trials. Unfortunately, whether or not a detergent solubilises a particular protein and retains protein function has to be determined empirically.

To assess which detergents solubilise YiaM and YiaN most effectively, solubility trials were performed with a number of different detergents from the zwitterionic, alkyl glucoside and PEP (polyoxoethylenes, monodisperse and polydisperse) families of detergents.

Cultures of BL21 (DE3) pCM14 and BL21 (DE3) pCM22 were grown and induced using the optimised conditions previously elucidated. The inner membrane fractions were then purified with sucrose density gradients (Methods section 2.8.4). The detergents used in the YiaM detergent trial were; CHAPS, THESIT (C₁₂E₉), *n*-nonyl- β -maltoside (NM), *n*-nonyl- β -D-glucopyranoside (NG), dodecyl- β -D-maltoside (DDM), β -octyl glucoside (β -OG), lauryldimethylamine oxide (LDAO), Zwittergent 3-08, Triton X100 and Zwittergent 3-16. The solubilisation trial for YiaN was performed with the following detergents; CHAPS, NM, THESIT, NG, DDM, β -OG, LDAO, Zwittergent 3-08, Triton X-100 and C₈E₅. The detergent solubility screen protocol was used (Methods section 2.13).

Analysis of the soluble and insoluble fractions after solubilisation using Coomassie stained SDS-PAGE gels and Western blotting revealed that DDM, NM, Triton X-100 and THESIT were capable of solubilising YiaM (Fig. 4.7) to a satisfactory degree, whereas the other detergents tested did not. A similar pattern was observed for the YiaN solubilisation trial. In this case only NM, DDM and Triton X-100 showed satisfactory solubilisation (Fig. 4.8). DDM was particularly efficacious in the solubilisation of both YiaM and YiaN and was far superior to any other detergent tested.

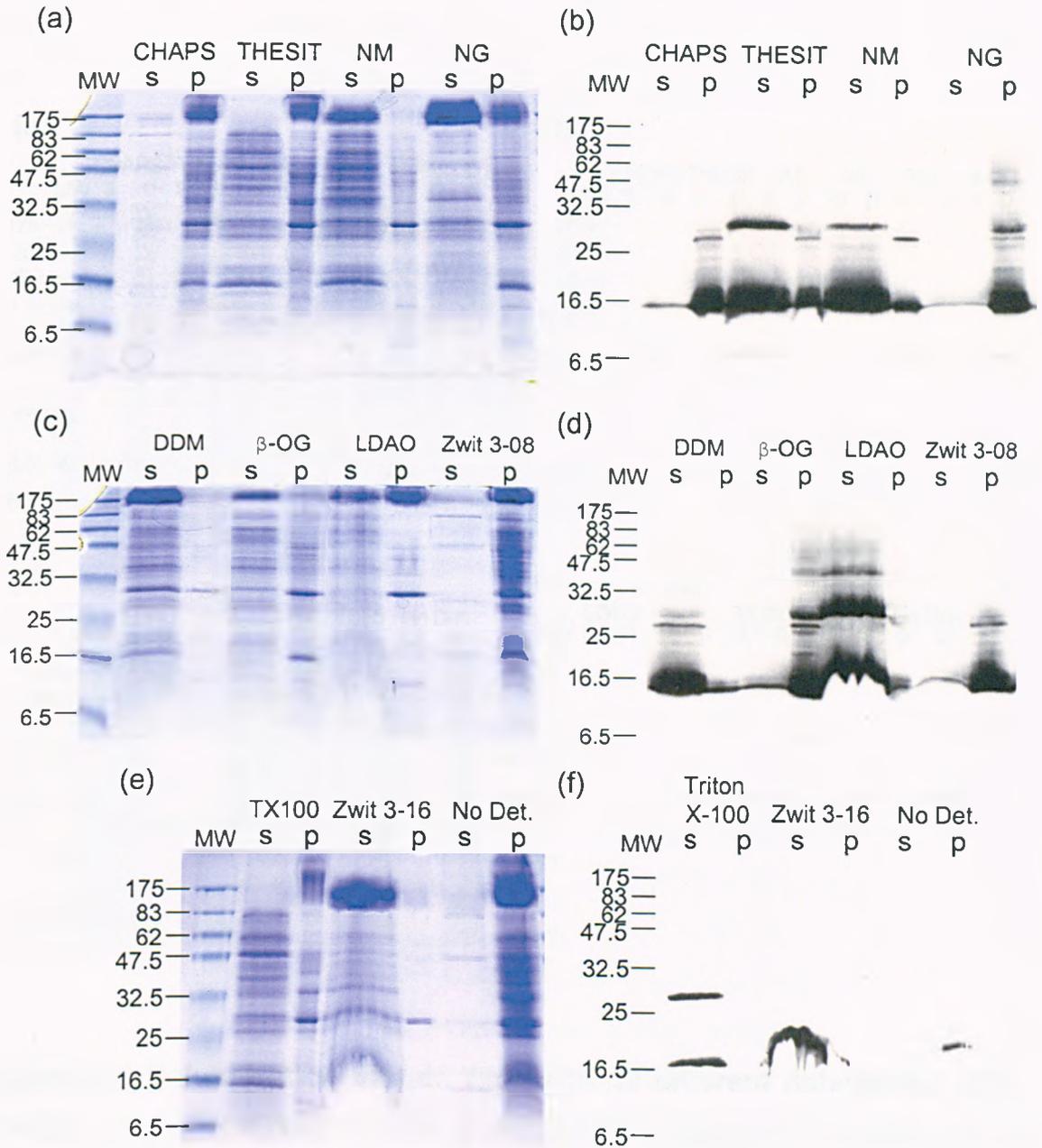


Figure 4.7 Solubilisation trial of YiaM with 10 different detergents. SDS-PAGE (a, c and e) and Western blots (b, d and f) of the soluble (s) and insoluble (p) fractions after solubilisation with various detergents. The detergents used were CHAPS, THESIT ($C_{12}E_9$), Nonyl maltoside (NM), *n*-nonyl- β -D-glucopyranoside (NG), dodecyl- β -D-maltoside (DDM), β -octyl glucoside (β -OG), Lauryldimethylamine oxide (LDAO), Zwittergent 3-08, Triton X-100 and Zwittergent 3-16. Molecular weight (MW) shown in kDa.

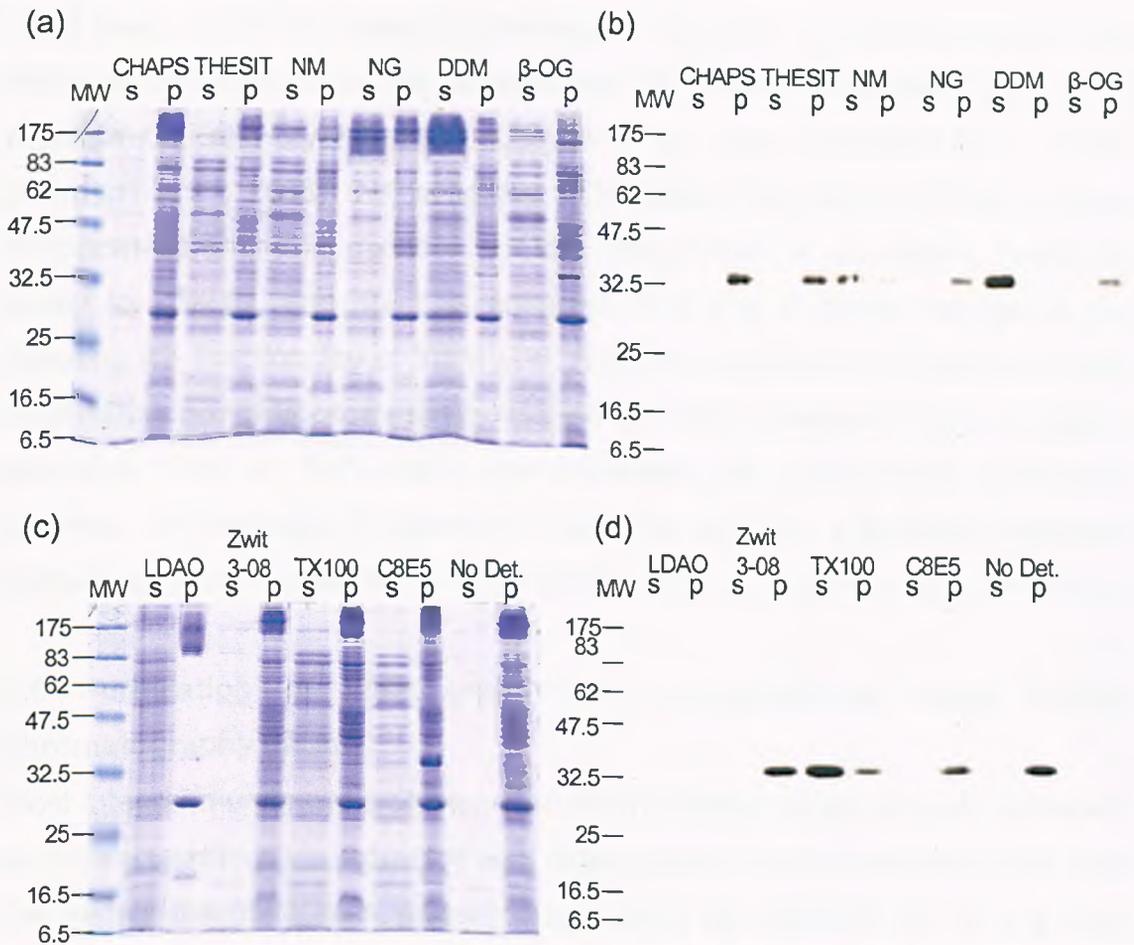


Figure 4.8 Solubilisation trial of YiaN with 10 different detergents. SDS-PAGE (a and c) and Western blots (b and d) of the soluble (s) and insoluble (p) fractions after solubilisation with various detergents. The detergents used were CHAPS, THESIT ($C_{12}E_9$), Nonyl maltoside (NM), *n*-nonyl- β -D-glucopyranoside (NG), dodecyl- β -D-maltoside (DDM), β -octyl glucoside (β -OG), Lauryldimethylamine oxide (LDAO), Zwittergent 3-08, Triton X-100 and C_8E_5 . Molecular weight (MW) shown in kDa.

The results obtained from both solubilisation trials indicate that a number of detergents could potentially be used to solubilise both YiaM and YiaN, however, DDM is by far the best at not only solubilising YiaM and YiaN, but also solubilising almost the entirety of the *E. coli* membrane proteome (Fig. 4.7c DDM lane). DDM has been the detergent of choice in many studies in the literature and is used for the solubilisation of integral membrane proteins for reconstitution and crystallisation (Curnow *et al.*, 2004, Murakami *et al.*, 2002, Screpanti *et al.*, 2006); indicating that DDM retains the function of the integral membrane proteins in question. Unfortunately there is no current functional assay for YiaMN, and so, it is unknown how any of these detergents are affecting the functionality of YiaM and YiaN, this could only be assessed after reconstitution into liposomes. Knowledge of which detergents can or cannot solubilise YiaM or YiaN could also potentially be used in the purification process, for example, a detergent could be used to selectively solubilise contaminants and separate the target protein making purification more effective.

4.6 Purification of YiaM and YiaN by immobilized metal affinity chromatography (IMAC)

Most integral membrane proteins, even when overexpressed, are not abundant and are unstable when extracted with detergents so need to be separated from the rest of the membrane proteome as quickly as possible, i.e. in one step. There is also a need to purify integral membrane proteins in one step as it has been observed that the delipidation of membrane proteins over time can negatively affect their crystallisation and reconstitution into liposomes. To overcome these problems IMAC purification is used. This method relies on the high affinity interaction of the IMAC resin with the hexahistidine tag on the target protein, which is then used to purify and enrich the protein of interest.

YiaM was purified from the inner membrane fraction prepared from BL21 (DE3) pCM14 by first extracting the protein from the membrane with DDM using the standard protocol (Methods section 2.14.1), except that 2 mg/ml (final concentration) inner membrane protein was solubilised overnight at 4°C in a 5 ml solubilisation reaction consisting of 20 mM HEPES, 10% glycerol, 300 mM NaCl and 1% DDM (w/v). YiaM was purified using a batch washing protocol with disposable columns (Methods section 2.14.2). YiaN was extracted from mixed

membranes prepared from BL21 (DE3) pCM22 using the standard protocol (Methods section 2.14.1), except that 2 mg/ml (final concentration) mixed membrane protein was solubilised overnight at 4°C in a 15 ml solubilisation reaction consisting of 20 mM HEPES, 10% glycerol, 300 mM NaCl and 1% DDM (w/v). YiaN was purified using a column and a peristaltic pump following the standard protocol (Methods section 2.14.3), except that the buffers used were the same as for the YiaM purification.

Analysis of the SDS-PAGE and Western blot of the YiaM purification fractions revealed a Western blot signal for YiaM in the inner membrane fraction and soluble fraction after detergent extraction and centrifugation, but not in the flowthrough or insoluble fractions. This indicates that almost all the YiaM is being solubilised by DDM and the soluble YiaM is binding to the Ni²⁺-NTA resin (Fig. 4.9b). SDS-PAGE of the elution fractions reveals the major band to be ~18 kDa – the molecular weight expected for YiaM (Fig. 4.9a). The Western blot of the elution fractions reveals the predominant signal to be at ~18 kDa, and also reveals signals for higher molecular weight species at regular and discreet intervals possibly signifying the formation of YiaM oligomers (Fig. 4.9b). The immunoreactive species at ~18 and ~28 kDa were confirmed to be YiaM by tandem mass spectroscopy. YiaM is estimated to be >95% pure after IMAC purification and a yield of 0.75 mg/L culture.

Western blot analysis of the YiaN purification fractions indicates that it is present in the mixed membrane fraction and soluble fraction after solubilisation and centrifugation. As with the YiaM purification there is no Western blot signal in either the insoluble fraction or in the flowthrough indicating that YiaN was successfully solubilised and it was completely bound to the Ni²⁺-NTA resin (Fig. 4.9d). The Western blot indicates that there is an immunoreactive species in the second elution fraction with a molecular weight of ~31 kDa; which was confirmed to be YiaN by peptide mass fingerprinting. There is also a higher molecular weight immunoreactive species visible in the Western blot at ~63 kDa, of which the identity is unknown (Fig. 4.9d). A number of low intensity bands are visible in the second elution fraction on the Coomassie stained SDS-PAGE gel with the most abundant at ~83 kDa (Fig. 4.9c). To identify what these contaminants were, the second elution fraction was concentrated using a 5000 MWCO centrifugal concentrators (Vivascience) and analysed by SDS-PAGE

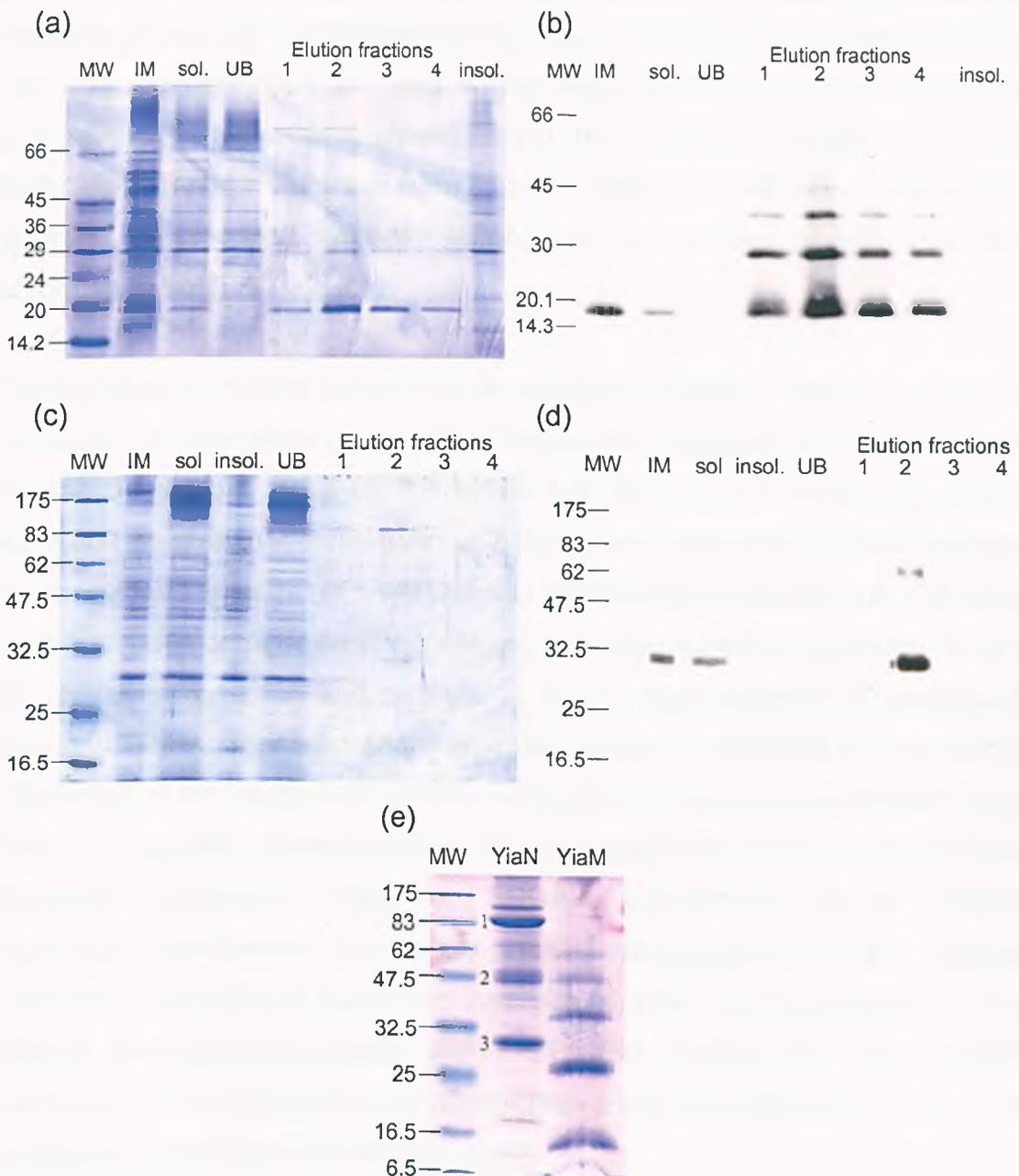


Figure 4.9 IMAC purification of YiaM and YiaN from inner membrane. (a) SDS-PAGE and (b) Western blot of YiaM purification and (c) SDS-PAGE and (d) Western blot of YiaN purification. SDS-PAGE and Western blots show the following purification samples; inner membrane fraction (IM) or mixed membrane fraction (MM), the soluble fractions (sol) and insoluble fractions (insol) after detergent extraction and centrifugation, the flowthrough (UB) and the elution fractions. (e) SDS-PAGE of concentrated elution fractions from YiaM and YiaN IMAC purifications. Proteins identified by peptide mass fingerprinting as ¹ AcrB, ² GlgA and ³ YiaN. Molecular weight (MW) shown in kDa.

along with YiaM elution fraction and the 3 most intense bands were excised and identified by peptide mass fingerprinting (Fig. 4.9e). The most intense band at ~83 kDa was identified as AcrB which has been observed as a common contaminant in other laboratories using the pTTQ18 system and IMAC purification (Peter Henderson, unpublished data), the ~47 kDa species was identified as glycogen synthase (GlgA) and the species at ~31 kDa was confirmed to be YiaN.

The purification of YiaM from the inner membrane fraction yielded a ladder of bands on the Coomassie stained SDS-PAGE and Western blot (Fig 4.9a, b and e). The bands that make up the ladder are discreet and have approximate molecular weights that correspond with them being oligomers of YiaM. Indeed, the monomeric species (at ~18 kDa) and potential dimer species (at ~28 kDa) have both been confirmed to be YiaM using tandem mass spectroscopy. This is the first overexpression and purification of an integral membrane component from any TRAP transporter and due to the scarcity of structural and functional knowledge of the integral membrane components it is unknown whether these potential oligomeric states are of functional significance or an artefact of the purification procedure. YiaM has separate polypeptides for its integral membrane components, but many TRAP transporters have the integral membrane components fused into one polypeptide thus suggesting that the integral membrane components are in a 1:1 ratio. If this is the case and the functional YiaM component is a multimer then one would expect YiaN to also be a multimer when co-expressed with YiaM.

The Western blot of the purification of YiaN revealed 2 signals; one at ~30 kDa and the second at ~62 kDa. The species at 30 kDa was confirmed to be YiaN, and although the higher molecular weight species has not been identified the molecular weight is the approximate size one would predict for a dimeric YiaN. The construct used to express YiaN also should also produce an untagged YiaM that will not be visible on the Western blot, but may be discernable on the SDS-PAGE gel. The SDS-PAGE of the elution fractions revealed very weak bands indicating dilute levels of protein in the eluate. To analyse the protein content more thoroughly the second elution fraction was concentrated and

analysed again on SDS-PAGE (Fig. 4.9e). This revealed a number of proteins that were not apparent previously but could be identified by tandem mass spectroscopy. The most intense band was identified as AcrB with a molecular weight of 83 kDa (Fig. 4.9e). AcrB is a component of the *E. coli* acridine efflux pump and has been found to be a common contaminant in other laboratories using the pTTQ18 system and IMAC purification (Peter Henderson, unpublished data); the second most intense was YiaN at ~30 kDa (Fig. 4.9e) and the third most intense was *E. coli* glycogen synthase (GlgA) at ~47 kDa (Fig. 4.9e). Glycogen synthase is a natural metal-binding protein and is reported to be a common contaminant of IMAC purifications (Bolanos-Garcia & Davies, 2006). GlgA has a higher affinity for Mg^{2+} than Ni^{2+} so including some Mg^{2+} in the washing stages may be a method by which this contaminant could be removed. There are a number of reasons that the YiaN purification is not as successful as the YiaM purification. The first is that the resin volume used to purify YiaN was too large and therefore did not specifically select for the highest affinity motifs, i.e. the hexahistidine tag. This allowed a higher capacity for lower affinity interactions to occur. Unlike the YiaM purification protocol, the solubilised membrane proteins were not incubated with the resin for a prolonged period of time which may help the resin bind affinity tags that may be buried or less accessible. YiaM was purified from inner membrane, whereas YiaN was purified from mixed membranes; however, the major contaminants identified were all inner membrane proteins so the retention of the outer membrane was not detrimental to the overall purity. The purification of YiaN could be made more effective by the optimisation of purification conditions or the inclusion of a secondary purification step, for example ion exchange chromatography or size exclusion chromatography.

4.7 Treatment of purified YiaM with heat denaturation and reducing agent

Purification of YiaM by IMAC yielded a ladder of immunoreactive species of which the 2 lowest molecular weight species were confirmed as YiaM using tandem mass spectroscopy. YiaM has 3 cysteine residues that are predicted to be at the centre of transmembrane helices I, II and IV (using TMHMM) and it is unknown whether these contribute to oligomer formation.

To demonstrate that these higher molecular weight species were indeed oligomers of YiaM an experiment was designed in which purified YiaM was treated with boiling and the reducing agent β -mercaptoethanol to try and obtain only the monomeric species.

YiaM was purified using IMAC and treated with combinations of 0%, 5% or 20% β -mercaptoethanol with no boiling, 1 minute boiling or 5 minute boiling at 95°C.

In the presence of no β -mercaptoethanol (0%) there are 4 major species at ~18, ~28, ~46 and ~100 kDa on the Coomassie stained SDS-PAGE and Western blot (Fig 4.10a and b). The predominant species is the 28 kDa "dimer" species in the SDS-PAGE and boiling for 0, 1 or 5 minutes does not significantly change the appearance of the ladder. With the addition of 5% β -mercaptoethanol the ~100 kDa species is no longer visible on the Western blot and the ~18 kDa species becomes the predominant species in both the SDS-PAGE and the Western blot. After 5 minutes boiling the only species visible on the Western blot were the monomer and the dimer. The monomer was the predominant species after the addition of 20% β -mercaptoethanol with no boiling. After boiling for 1 minute, the monomer is visible and a very weak signal for the dimer, there is also evidence of aggregation where the protein did not leave the wells and are visible at the top of the stacking gel in the Western blot (Fig. 4.10b). After 5 minutes boiling in the presence of 20% β -mercaptoethanol there were no immunoreactive species as seen in the previous lanes, there was however evidence of aggregate in the stacking gel, suggesting that this treatment was too harsh and the protein completely aggregated/degraded.

Treatment of the YiaM ladder with β -mercaptoethanol leads to a reduction in the number and intensity of high molecular weight species. β -mercaptoethanol treatment in combination with heat denaturation led to a further decrease in the number of higher molecular weight species. At the more extreme conditions with 20% β -mercaptoethanol and boiling for 1 or 5 minutes significant aggregation of protein was observed. These observations and the fact that the 2 lowest molecular weight species of the ladder were identified as YiaM suggest that YiaM is forming higher order oligomers. This does not mean that the

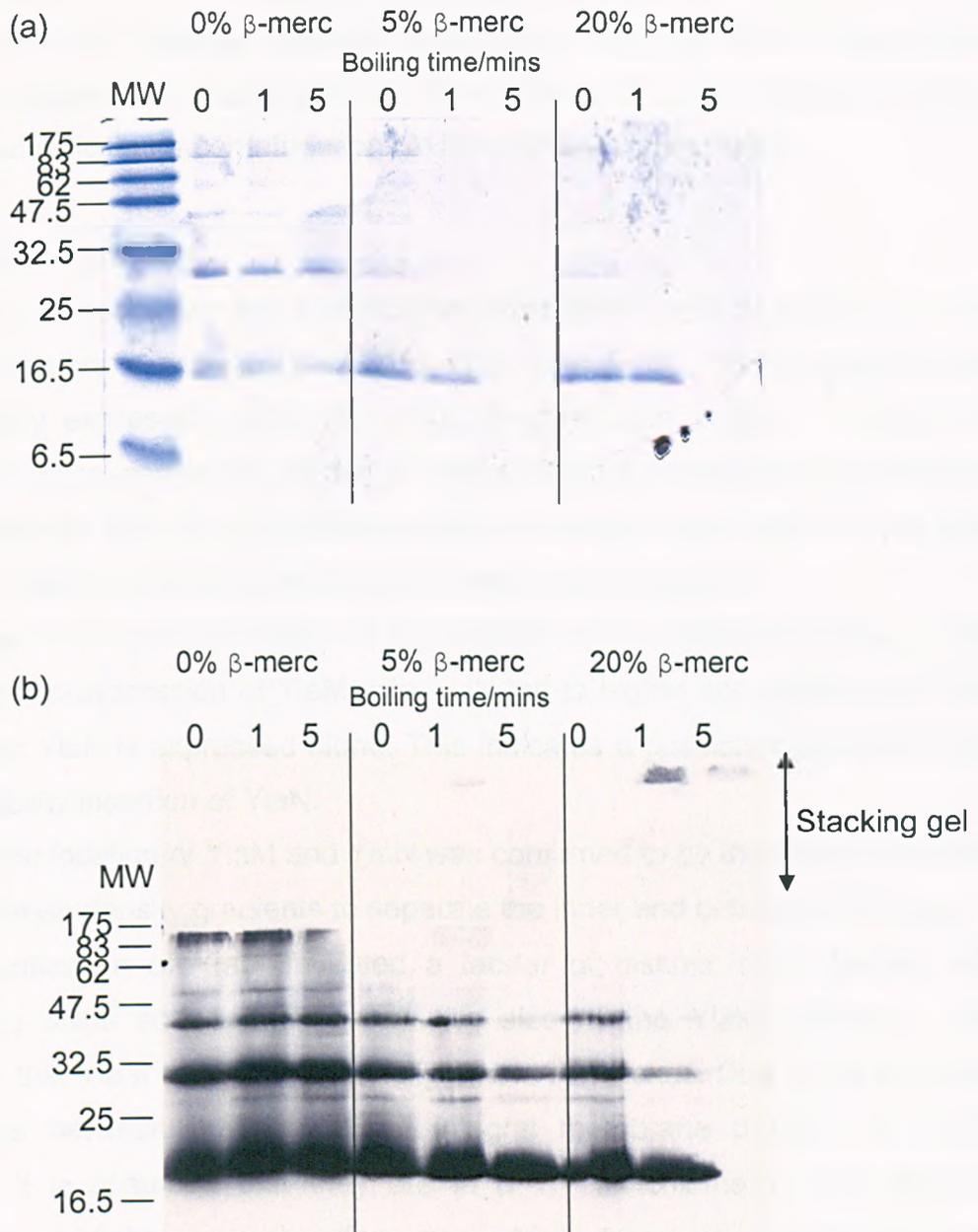


Figure 4.10 Effect of β -mercaptoethanol treatment and heat denaturation on the apparent YiaM ladder after IMAC purification. **(a)** SDS-PAGE and **(b)** Western blot of the effects of the addition of 0%, 5% or 20% β -mercaptoethanol in combination with no boiling, 1 minute or 5 minute boiling. The stacking gel is indicated in **(b)**. The sample for 5% β -mercaptoethanol with 5 minutes boiling was not loaded in **(a)**. Molecular weight (MW) shown in kDa.

oligomers formed are functionally significant and they may just be a purification artefact with the cysteine residues interacting aberrantly. More experiments would be required to assess whether these oligomers are functionally relevant such as analytical ultracentrifugation (AUC) or *in vivo* crosslinking.

4.8 Chapter summary

This chapter describes the homologous expression and purification of the integral membrane proteins YiaM and YiaN from *E. coli*. Both proteins were successfully expressed using the pTTQ18 expression vector. A number of growth/induction conditions known to affect integral membrane accumulation were optimised including; the concentration of inducer, the length of time after induction before harvesting the cells and the medium of growth.

YiaM was found to accumulate to high levels when expressed alone. It was found that co-expression of YiaM with YiaN led to higher accumulation of YiaN than when YiaN is expressed alone. This indicates a possible function of YiaM in the stability/insertion of YiaN.

The cellular location of YiaM and YiaN was confirmed to be the inner membrane using sucrose density gradients to separate the inner and outer membranes.

IMAC purification of YiaM revealed a ladder of distinct YiaM species with increasing sizes corresponding with the size of the YiaM monomer. This suggests that YiaM can form higher oligomeric complexes. Due to the existence of fusions between the two TRAP integral membrane proteins in certain systems it is assumed that they are in a 1:1 stoichiometry. The apparent oligomers of YiaM would therefore mean YiaN forms oligomeric structures. When *yiaM* and *yiaN* were co-expressed and purified using the hexa-histidine tag on the YiaN subunit, a ladder of YiaN species was not observed and a band for YiaM was not identified, so it is currently unknown what the subunit stoichiometry is. The exact nature of the apparent YiaM complexes in solution is not known and may be an artefact of the expression and purification procedures.

Chapter 5

Expression and purification of SiaQM from *H. influenzae*

5.1 Cloning and expression trial of the *siaQM* genes using pTTQ18

To characterise the integral membrane protein from the *H. influenzae* TRAP transporter SiaPQM, the gene encoding this protein, *SiaQM*, was cloned into the pTTQ18-RGSH₆ expression vector for overexpression and purification. To do this the region of the *H. influenzae* Rd KW20 genome encoding *siaQM* was amplified by PCR using primers SiaQMpTTQHisF and SiaQMpTTQHisR and cloned into pTTQ18-RGSH₆ on an *EcoRI-PstI* fragment so that the coding sequence was in frame with the 3'-nucleotide sequence encoding the amino acids arginine, glycine, serine and 6 histidines (RGSH₆). The chosen clone was named pCM23.1. This expression vector was transformed into *E. coli* BL21 (DE3) and used in expression trials (Methods section 2.7.1). The initial expression trial was performed in both LB and M9 minimal media. As observed in the expression optimisation of BL21 (DE3) pCM22 (Chapter 4), growth at 37°C can have deleterious effects on the steady state levels of overexpressed membrane proteins. It was therefore decided to perform the initial expression trial at 25°C.

As observed for the expression of YiaM and YiaN in pTTQ18-RGSH₆, the overexpression of *SiaQM* in pTTQ18 caused a significant decrease in the growth rate when inducer was added compared to the absence of inducer (data not shown) suggesting expression of an integral membrane protein. Analysis of the steady-state levels of *SiaQM* using Western blotting revealed that two immunoreactive proteins were produced from BL21 (DE3) pCM23.1 in the presence of IPTG when grown in either LB or M9 (Fig. 5.1). The immunoreactive species have apparent molecular weights of ~50 kDa and ~28 kDa and were not present in the membrane samples from the uninduced cultures. The predicted molecular weight of *SiaQM* is ~76 kDa and taking into account the aberrant migration of integral membrane proteins in SDS-PAGE, the ~50 kDa species was predicted to be *SiaQM*. The identity of the immunoreactive species at ~28 kDa was unknown. This ~28 kDa species may be a breakdown product of *SiaQM* or may be caused by the binding of the anti His-tag antibodies to an *E. coli* protein. Therefore, this initial expression experiment has shown that the pTTQ18 expression system is sufficient for the expression of *SiaQM* to detectable levels.

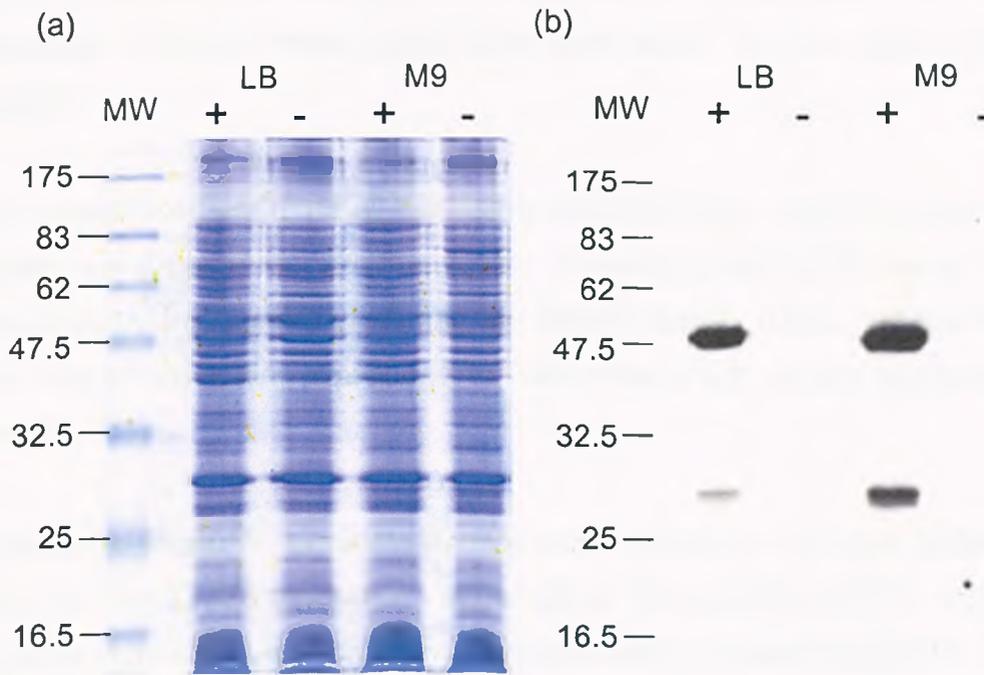


Figure 5.1 Expression of SiaQM from pTTQ18 based expression constructs in the presence (+) and absence (-) of IPTG induction. Coomassie stained SDS-PAGE (a) and Western blot (b) with anti His-tag antibody of SiaQM expression from BL21 (DE3) pCM23.1 grown in either LB or M9 minimal media to mid-log phase, induced with 0.5 mM IPTG and harvested after a further 2 hours growth. 20 μ g of protein was applied to the Coomassie stained SDS-PAGE and 5 μ g was applied to the Western blots. Molecular weight (MW) shown in kDa.

5.2 Optimisation of growth and induction conditions for SiaQM expression

As with YiaM and YiaN expression, it was important to optimise the growth and induction conditions for the production of SiaQM. A number of common parameters were taken into account, including the growth temperature, the growth medium and the length of time after induction that the cells are harvested. Each of these parameters was varied for the strain expressing SiaQM.

The growth medium, time of harvesting after induction and the temperature of growth were optimised simultaneously. A combination of the signal intensity produced by the Western blot and the optical density (OD₆₅₀) of the culture at the time of harvesting was used to determine which growth conditions were optimal.

Standard timecourse experiments (Methods section 2.7.4) were performed for BL21 (DE3) pCM23.1 grown on either LB or M9 at 25°C or 37°C, producing 4 cultures in total. Mixed membrane fractions were prepared using the water lysis method and the accumulation of SiaQM was analysed using Western blotting. Protein samples were also analysed on SDS-PAGE gels to ensure that samples had equal amounts of protein loaded (data not shown).

Analysis of steady-state levels of SiaQM using Western blotting indicated that SiaQM was produced by BL21 (DE3) pCM23.1 in all 4 of the different conditions (Fig. 5.2a-d). When grown at 37°C in M9 the peak signal was achieved after 1 hour of growth post-induction, the steady-state levels then gradually decreased until only a very weak signal remained on the Western blot. There was no visible Western blot signal in the sample taken from bacteria grown overnight (Fig. 5.2a). When grown at 37°C in LB, the signal increased and peaked after 5 hours growth and there was no Western signal from the sample taken after overnight growth (Fig. 5.2b).

Samples taken from BL21 (DE3) pCM23.1 grown at 25°C in M9 had the most intense Western blot signals of all the conditions tested (Fig. 5.2c). The accumulation of SiaQM did not peak in these conditions until after 6 hours

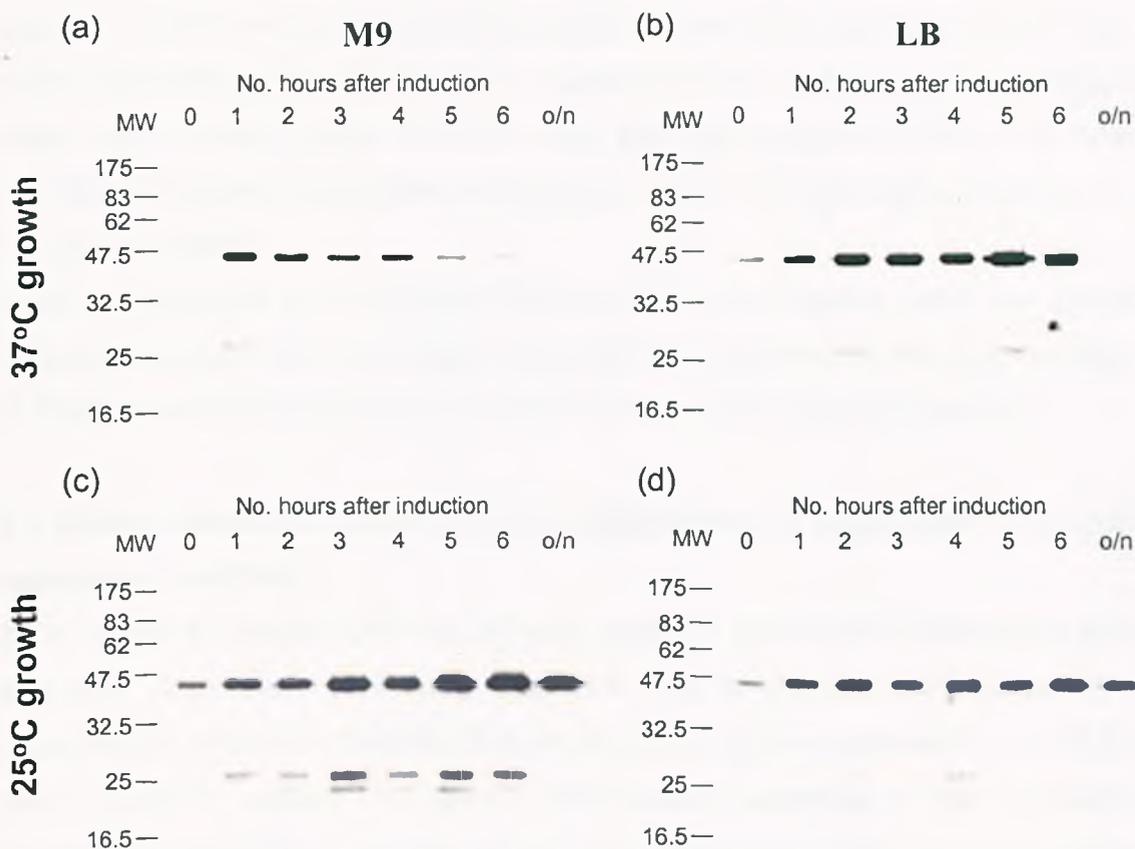


Figure 5.2 Optimisation of time of harvest after induction, growth medium used and temperature of growth for the production of SiaQM. Western blots with anti His-tag antibody of membrane fractions from BL21 (DE3) pCM23.1 samples harvested at hourly intervals after induction. BL21 (DE3) pCM23.1 grown in M9 at 37°C (a), LB at 37°C (b), M9 at 25°C (c) and LB at 25°C (d). All samples were induced at mid-log phase with 0.5 mM IPTG. 5 µg of protein was used for all samples in the Western blots. Molecular weight (MW) shown in kDa.

growth and, unlike the cultures grown at 37°C, there was also significant Western blot signal in the sample taken from culture grown overnight. A similar accumulation pattern was observed for samples taken from cells grown in LB at 25°C; the peak accumulation occurred after 6 hours growth and significant levels of SiaQM were observed in the sample taken after overnight growth (Fig. 5.2d). Even though the Western blot signals are more intense in the samples taken from cultures grown at 25°C in M9, the cells grown at 25°C in LB have significantly higher optical densities (OD₆₅₀) of 6.7, ~10 fold higher than any of the other samples.

Using the intensity of the Western blot signals in combination with the optical densities obtained for the samples, the optimal conditions for the accumulation of SiaQM are growth at 25°C in LB and harvested after overnight growth.

5.3 Determination of SiaQM cellular sublocation by separation of *E. coli* membrane fractions

As a means to confirm that SiaQM was localized to the inner membrane and also as a preliminary purification step, the inner and outer membranes were separated by a sucrose density gradient. A larger volume, typically 5 L, of BL21 (DE3) pCM23.1 culture was grown and induced according to the optimised conditions mentioned previously. The mixed membrane fraction was prepared and applied to the sucrose density gradient as detailed in Methods section 2.8.4. The protein contents of the mixed membranes, inner membrane and outer membrane fractions were analysed on SDS-PAGE and Western blots.

Analysis of the protein content of the inner and the outer membrane fractions by SDS-PAGE indicated some differences in the protein content between the 3 different samples (Fig. 5.3a) demonstrating that there had been some separation of the inner and outer membrane. The Western blot of the same samples shows a signal for SiaQM in each of the membrane preparations, however, the signal in the inner membrane sample is the most intense indicating enrichment of the protein in this fraction (Fig. 5.3b). There is also a very weak Western blot signal for SiaQM in the outer membrane fraction samples; this is most likely due to cross-contamination with inner membrane during the separation process.

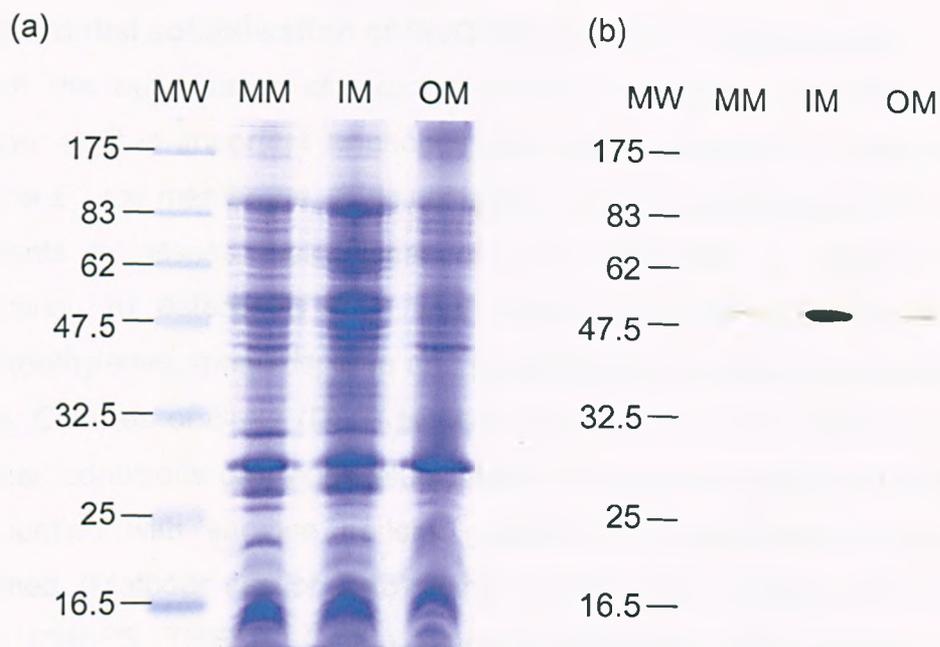


Figure 5.3 Comparison SiaQM content of mixed, inner and outer membrane fractions prepared from BL21 (DE3) pCM23.1. Coomassie stained SDS-PAGE (a) and Western blot (b) of mixed (MM), inner (IM) and outer (OM) membranes produced from a culture of BL21 (DE3) pCM23.1 grown in LB to mid-log phase, induced with 0.5 mM IPTG and harvested after overnight growth. 20 μ g of protein was applied to the Coomassie stained SDS-PAGE and 5 μ g of protein is applied to the Western blots. Molecular weight (MW) shown in kDa.

Separating the inner membrane and the outer membrane has led to the confirmation that SiaQM is indeed localised to the inner membrane in *E. coli*. It has also revealed itself to be a method by which the inner membrane containing overexpressed protein can be separated from a lot of potentially contaminating protein shown in the outer membrane lanes in the SDS-PAGE gels (Fig. 5.3a).

5.4 Differential solubilisation of SiaQM by a variety of detergents

As with the optimisation of expression and purification of YiaM and YiaN (Chapter 4), it is important to choose the correct detergent to extract SiaQM from the *E. coli* membrane. A solubilisation screen is performed with different detergents to assess their efficacy for the extraction of SiaQM from the membrane. 10 detergents from the zwitterionic, alkyl glucoside and PEP (polyoxoethylenes, monodisperse and polydisperse) families are included in the screen. Cultures of BL21 (DE3) pCM23.1 were grown and induced using the optimised conditions previously elucidated. The inner membrane fractions were then purified with sucrose density gradients. A solubilisation screen was performed (Methods section 2.13) using the detergents; dodecyl- β -D-maltoside (DDM), CHAPS, THESIT (C₁₂E₉), *n*-nonyl- β -maltoside (NM), β -octyl glucoside (β -OG), lauryldimethylamine oxide (LDAO), Zwittergent 3-08, Triton X-100, C₈E₅, C₁₃E₁₀ and octyl- β -D-thioglucopyranoside (OTG).

Analysis of the soluble and insoluble fractions after solubilisation using Coomassie stained SDS-PAGE gels and Western blotting revealed that DDM, NM, β -OG were capable of solubilising SiaQM to a satisfactory degree, whereas the other detergents tested did not (Fig. 5.4). As seen in the detergent trial for YiaM and YiaN, DDM was particularly efficacious in the solubilisation of SiaQM and was far superior to any other detergent tested.

The results obtained from this solubilisation trial indicated that a number of detergents could potentially be used to solubilise SiaQM, however, as seen for YiaM and YiaN, DDM was by far the best. Knowledge of which detergents can or cannot solubilise SiaQM could also be used potentially in the purification process, for example, a detergent could be used to selectively solubilise contaminants and separate the target protein making purification more effective.

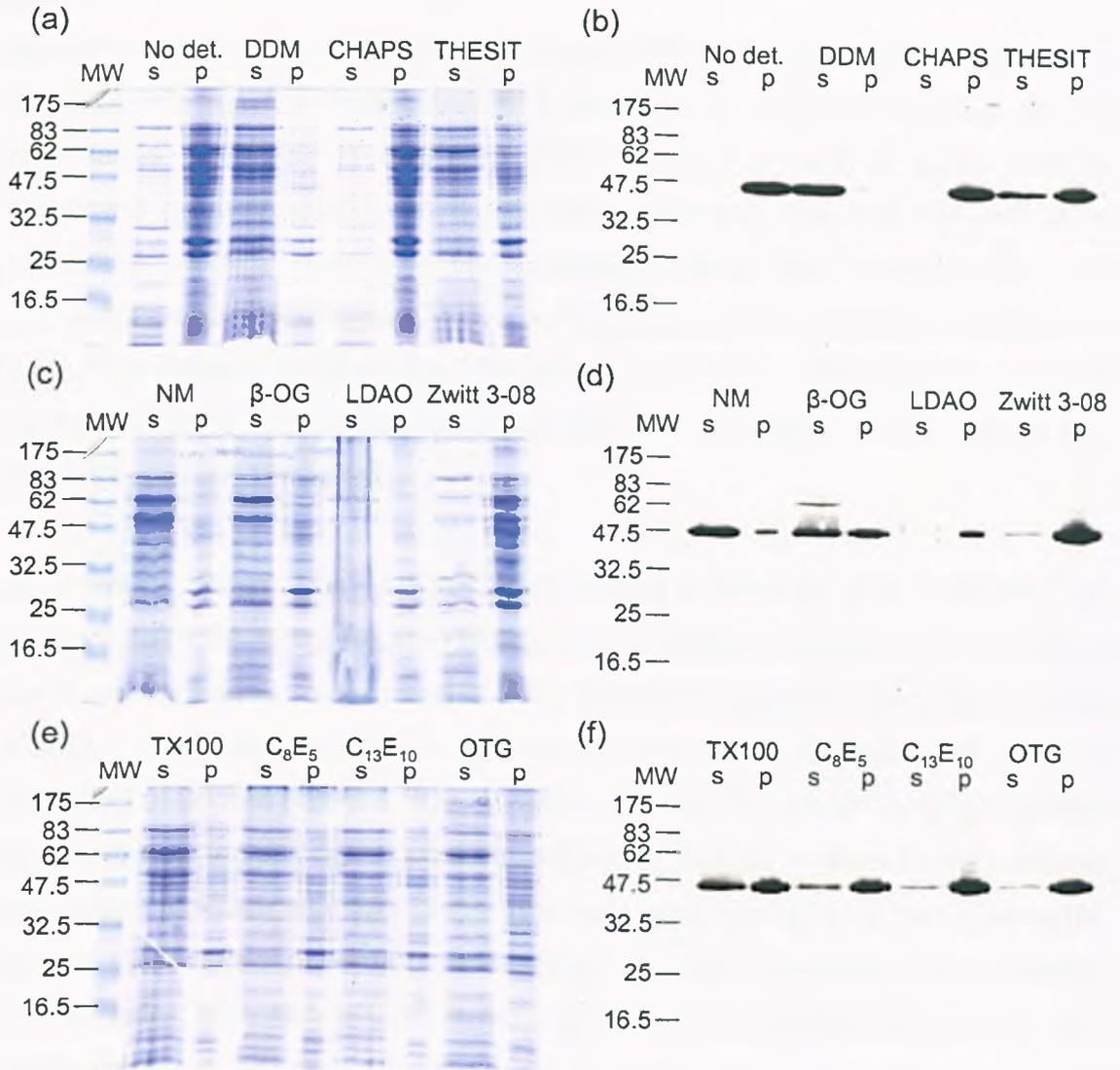


Figure 5.4 Solubilisation trial of SiaQM with 10 different detergents. Coomassie stained SDS-PAGE (a, c and e) and Western blots (b, d and f) of the soluble (s) and insoluble (p) fractions after solubilisation with various detergents. The detergents used were dodecyl- β -D-maltoside (DDM), CHAPS, THESIT ($C_{12}E_9$), *n*-nonyl- β -maltoside (NM), β -octyl glucoside (β -OG), lauryldimethylamine (LDAO), Zwittergent 3-08, Triton X-100, C_8E_5 , $C_{13}E_{10}$ and octyl- β -D-thioglucopyranoside (OTG). Molecular weight (MW) shown in kDa.

5.5 Purification of SiaQM by immobilized metal affinity chromatography (IMAC)

Most integral membrane proteins, even when overexpressed, are not abundant and are unstable when extracted with detergents so need to be purified as well as possible preferably in one step. There is also a need to purify integral membrane proteins quickly as it has been observed that the delipidation of membrane proteins over time can negatively affect their crystallisation and reconstitution into liposomes. To overcome these problems IMAC purification is used. This method relies on the high affinity interaction of the IMAC resin with the hexahistidine tag on the target protein, which is then used to purify and enrich the protein of interest.

SiaQM was purified from the inner membrane fraction by first extracting the protein with DDM using the standard protocol (Methods section 2.14.1), except that 7 mg/ml (final concentration) of inner membrane prepared from BL21 (DE3) pCM23.1 was solubilised for 1 hour at 4°C in a 9 ml solubilisation reaction consisting of 20 mM HEPES, 10% glycerol, 300 mM NaCl and 1% DDM. SiaQM was purified using the batch washing method (Methods section 2.14.2), except that, after the batch washing stage, the resin was applied to a non-disposable column and eluted according to Methods section 2.14.3 with the following elution buffer; 10 mM NaPi, pH 8, 10% glycerol, 250 mM imidazole, pH 8, and 0.05% DDM.

Analysis of the SiaQM purification fractions with Coomassie-stained SDS-PAGE and Western blotting revealed that a strong signal on the Western blot was visible in the inner membrane and supernatant fractions, and a very weak signal was present in the flowthrough and insoluble fractions (Fig. 5.5a and b). The weak signal in the insoluble fraction may be due to the membrane fraction not being fully solubilised or could be a contamination of the soluble fraction. The nature of batch washing means that a small amount of resin (bound to protein) is invariably still suspended after centrifugation and this leads to the “unbound fraction”, which should be devoid of any SiaQM and resin, containing a small amount of both, which causes the signal in the Western blot. Regardless of

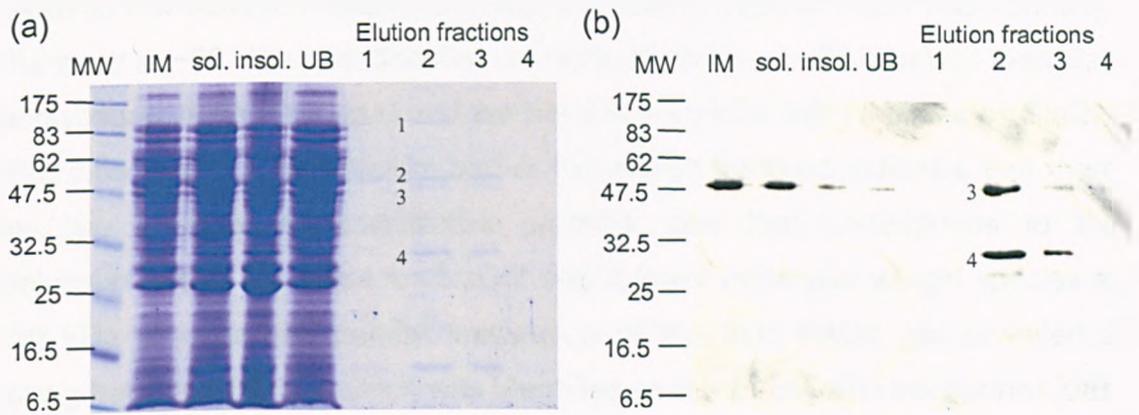


Figure 5.5 IMAC purification of SiaQM from inner membrane. Coomassie stained SDS-PAGE (a) and Western blot (b) of IMAC purification of SiaQM from inner membrane prepared from BL21 (DE3) pCM23.1. The following purification samples were analysed; inner membrane fraction (IM), the soluble fraction (sol) and insoluble fraction (insol) after detergent extraction and centrifugation, the flowthrough (UB) and the elution fractions. Proteins identified by peptide mass fingerprinting as 1) AcrB, 2) GlgA, 3) SiaQM and 4) ZitB. Molecular weight (MW) shown in kDa.

these technicalities, the Western blot indicates that the majority of SiaQM is solubilised and is bound to the resin.

Analysis of the elution fractions with Coomassie-stained SDS-PAGE revealed several bands of equal intensity indicating that the IMAC purification was not entirely successful (Fig. 5.5a). To aid in rationally designing further purification methods and to assess whether SiaQM was in the eluate, a number of the bands in the elution fraction were identified using peptide mass fingerprinting. The band at ~90 kDa was identified as AcrB, the band at ~50 kDa was identified as glycogen synthase (GlgA) and the band at ~45 kDa was identified as SiaQM (Fig. 5.5a and b). The Western blot of the elution fractions indicates that there are two strongly immunoreactive proteins; one that corresponds to the molecular weight expected for SiaQM and a lower molecular weight species at ~30 kDa of unknown identity. Inspection of the SDS-PAGE gel revealed a strong band at ~30 kDa which was identified as the *E. coli* zinc transporter ZitB. The contaminant at ~10 kDa was not identified.

Analysis of the preliminary purification of SiaQM using SDS-PAGE revealed that in the conditions used it was not possible to purify SiaQM to homogeneity. Using a combination of Western blotting and peptide mass fingerprinting however, it was confirmed that SiaQM was bound by the Ni²⁺-NTA resin and was a major constituent of the eluate. In an effort to design a purification protocol that would result in homogeneous SiaQM, the contaminants were identified. Two of the contaminants identified, AcrB and GlgA, were also major contaminants of the YiaN purification by IMAC (Chapter 4). The same resin was used in both and so it is likely that these proteins are weakly bound by non-specific interactions with the resin. GlgA is a common contaminant of IMAC because it is a natural metal binding protein (Bolanos-Garcia & Davies, 2006).

Analysis of the amino acid sequence of AcrB revealed the C-terminus to be histidine rich with 4 out of the 8 terminal residues being histidines. Analysis of the amino acid sequence of the contaminant ZitB revealed that it is a histidine-rich protein with a sequence of H-S-H-H-H-H at the C-terminus, forming a functional, native histidine tag. This histidine tag is the cause of the ~30 kDa immunoreactive species in the Western blot and facilitates ZitB binding to the

Ni²⁺-NTA resin. There are a number of methods available to purify SiaQM further including increasing the imidazole concentration of the washing steps or implementing an imidazole gradient to elute the more weakly associated contaminants. ZitB does not have a full hexahistidine tag unlike SiaQM so may bind more weakly. It may also be necessary to use another purification technique such as size exclusion chromatography or anion exchange chromatography to remove ZitB.

5.6 Further purification of SiaQM from inner membrane using anion exchange chromatography followed by IMAC

Following the identification of the IMAC purification contaminants, a number of modifications to the IMAC purification protocol were implemented to try and remove the contaminants. The imidazole concentration in the washing step was increased from 20 mM to 50 mM followed by elution with 250 mM imidazole. This resulted in SiaQM eluting from the column during the washing step and ZitB eluting in the 250 mM elution step. This indicated that ZitB actually has a higher affinity for the Ni²⁺-NTA than SiaQM. It was considered, however, it may still be possible to purify SiaQM by washing at a lower concentration of imidazole to remove all contaminants except ZitB and then increase the imidazole concentration and elute SiaQM without eluting ZitB. To assess whether this was possible, a step gradient of imidazole from 10-50 mM in 10 mM increments followed by elution at 250 mM imidazole was implemented. This revealed that SiaQM eluted at ~30 mM imidazole, but still unfortunately contained unacceptably high levels of ZitB. The detergent trial was revisited to see if ZitB could be selectively solubilised and separated from SiaQM, but the appearance of the ZitB signal in the detergent trial is not consistent even after long exposure times (Fig. 5.4) making conclusions difficult to draw. It was finally decided that anion exchange chromatography would be used as a way of separating SiaQM from ZitB. Preliminary anion exchange chromatography revealed that SiaQM and ZitB had different elution profiles over a linear gradient of NaCl between 5-200 mM. Elution of SiaQM was observed to occur between ~56-130 mM NaCl and ZitB eluted between 110-130 mM (data not shown). Using these data, the anion exchange protocol was refined and the fractions containing only SiaQM could be purified further using IMAC.

Using the preliminary anion exchange data, the standard anion exchange protocol (Methods section 2.14.6) was refined so that a shallower gradient was applied over the NaCl concentration range SiaQM and ZitB eluted. SiaQM was extracted from the inner membrane fraction using the standard protocol (Methods section 2.14.1), except that 6.9 mg/ml (final concentration) of inner membrane prepared from BL21 (DE3) pCM23.1 was solubilised for 1 hour at 4°C in a 12 ml solubilisation reaction consisting of 20 mM HEPES, 10% glycerol, 5 mM NaCl and 1% DDM. The refined anion exchange protocol involved a steep linear gradient from 5-30 mM of NaCl over 2 column volumes followed by a shallow linear gradient from 30-130 mM over 16 column volumes collected in multiple small fractions. This was followed by a linear gradient from 130-200 mM over 4 column volumes. SiaQM was further purified from the anion exchange eluate using 7 ml Ni²⁺-NTA resin (Qiagen) with a non-disposable column using the standard protocol (Methods section 2.14.3).

The absorbance (A₂₈₀) trace of the elution profile revealed a number of small contaminant elution peaks during the 0-30 mM NaCl wash. The trace also revealed a large elution peak between 56 and 103 mM NaCl, during the shallow NaCl gradient, followed by a steady elution of protein throughout the 130-200 mM NaCl gradient (Fig. 5.6). Analysis of the protein content of these fractions using Western blotting revealed a signal for SiaQM in elution fractions 30-50 which corresponded to between 60 and 100 mM NaCl (Fig. 5.7). The samples on the Western blot that contained a SiaQM signal did not contain a signal for ZitB, which was expected because the preliminary anion exchange purification revealed that ZitB eluted between 110-130 mM NaCl. Analysis of the protein content of the elution fractions using SDS-PAGE revealed a large proportion of contaminants eluted before and after the SiaQM elution peak (Fig. 5.7a and c), however a Western blot signal for ZitB was not visible (Fig. 5.7d) but this may be due to dilution of the protein. The elution fractions containing a signal for SiaQM on the Western blot (Fig. 5.7c and d, fractions 22-44) were collated and applied to IMAC resin.

Analysis of the IMAC purification samples using Western blotting revealed a

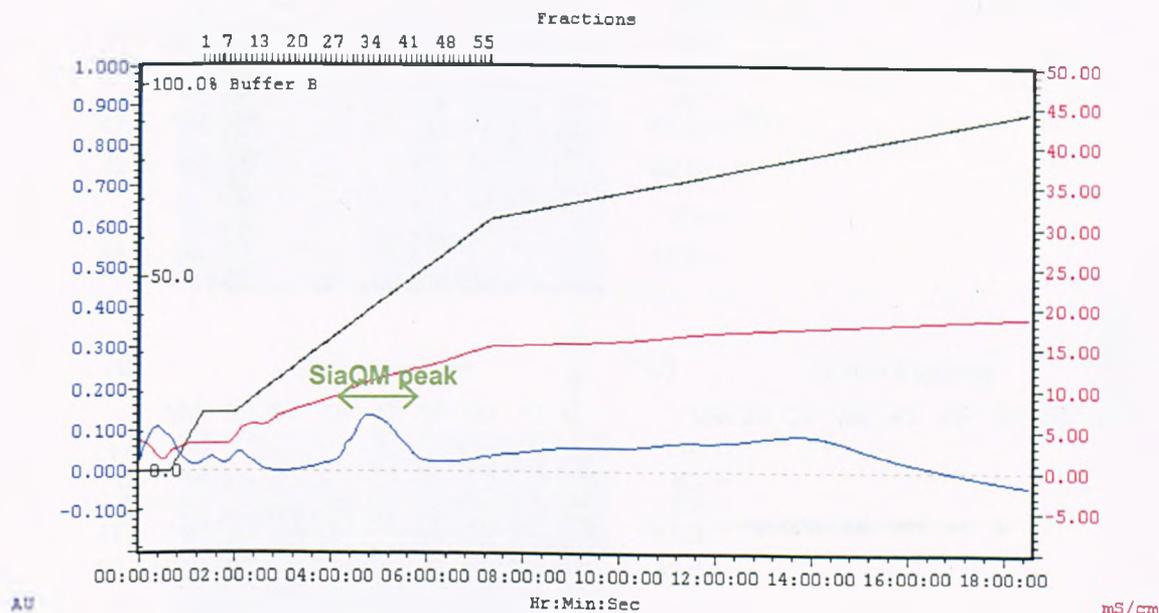


Figure 5.6 Anion exchange chromatography of SiaQM from inner membrane. Elution profile from DEAE sepharose resin (Sigma) monitored by absorbance (A_{280}) of the inner membrane fraction prepared from BL21 (DE3) pCM23.1 grown in LB to mid-log phase, induced with 0.5 mM IPTG and harvested after overnight growth. SiaQM peak indicated as identified from Western blotting (Fig. 7). A_{280} (blue line), conductance (red line) and percentage of 200 mM NaCl buffer (black line).

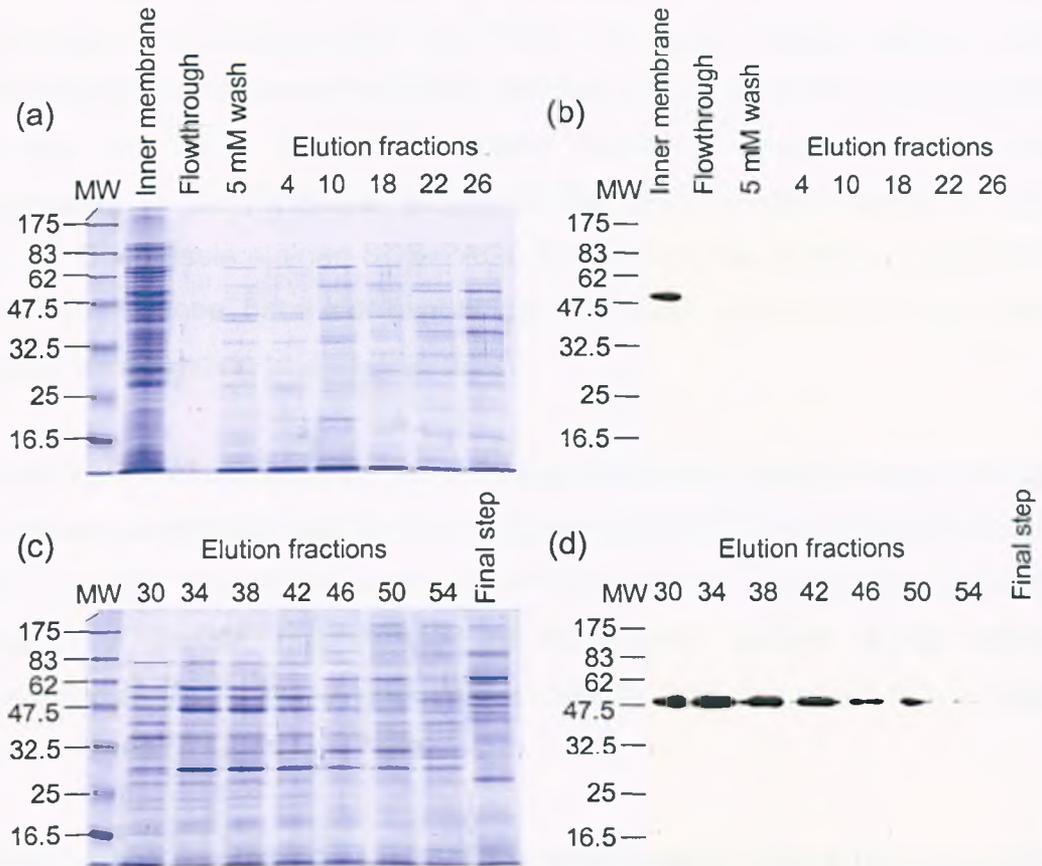


Figure 5.7 Analysis of fractions from anion exchange chromatography of SiaQM from inner membrane. Coomassie stained SDS-PAGE (a and c) and Western blot (b and d) of the purification of SiaQM with anion exchange chromatography. The samples included are the inner membrane fraction, the flowthrough from the column, the initial 5 mM wash and the numbered elution fractions. 20 μ g of protein was applied to the Coomassie stained SDS-PAGE and 5 μ g of protein is applied to the Western blots. Molecular weight (MW) shown in kDa.

signal for SiaQM in the anion exchange eluate and in elution fraction 2 and 3 (Fig. 5.8b). There were no signals for SiaQM in the flowthrough or the 20 mM wash fractions indicating that SiaQM was binding to the resin and remained bound during the washing step (Fig. 5.8b). The most intense signal in the elution fractions corresponds to SiaQM; there is also evidence of degradation. There are two more signals for lower molecular weight species; one corresponding to the molecular weight of ZitB and the other unknown (Fig. 5.8b). The Coomassie stained SDS-PAGE reveals only two bands in the elution fractions; an intense band corresponding to SiaQM and a very weak band probably corresponding to ZitB (Fig. 5.8a).

The combination of anion exchange chromatography and IMAC to purify SiaQM was relatively successful resulting in a preparation that contains SiaQM as the dominant species and ZitB as a very minor contaminant. Despite the successful purification of SiaQM, quantification of the protein content of the elution fractions revealed that the overall yield of SiaQM was low at ~0.15% of total inner membrane (0.12 mg/L culture).

5.7 Cloning and expression trials of the *siaQM* genes using the arabinose inducible pBAD vector in *E. coli* and the nisin inducible pNZ8048 vector in *L. lactis*.

One of the major reasons for the overexpression and purification of SiaQM was to reconstitute SiaQM into liposomes and characterise the transport with *in vitro* *N*-acetylneuraminic acid uptake assays (in collaboration with Professor Bert Poolman's laboratory in Groningen). Unfortunately, there were problems associated with the production of SiaQM from BL21 (DE3) pCM23.1 that were insurmountable in the timeframe of this project. The overall yield of SiaQM was low, which meant that a large volume of culture was required to obtain sufficient amounts of protein. Another obstacle was that when integral membrane proteins were extracted from the native membrane with detergents, delipidation occurs over time. This process, where the native lipids are stripped away as a function of time is thought to negatively affect the function/stability of integral membrane proteins (E.R. Geertsma, personal communication), therefore, the requirement for two purification techniques to obtain homogeneous SiaQM took

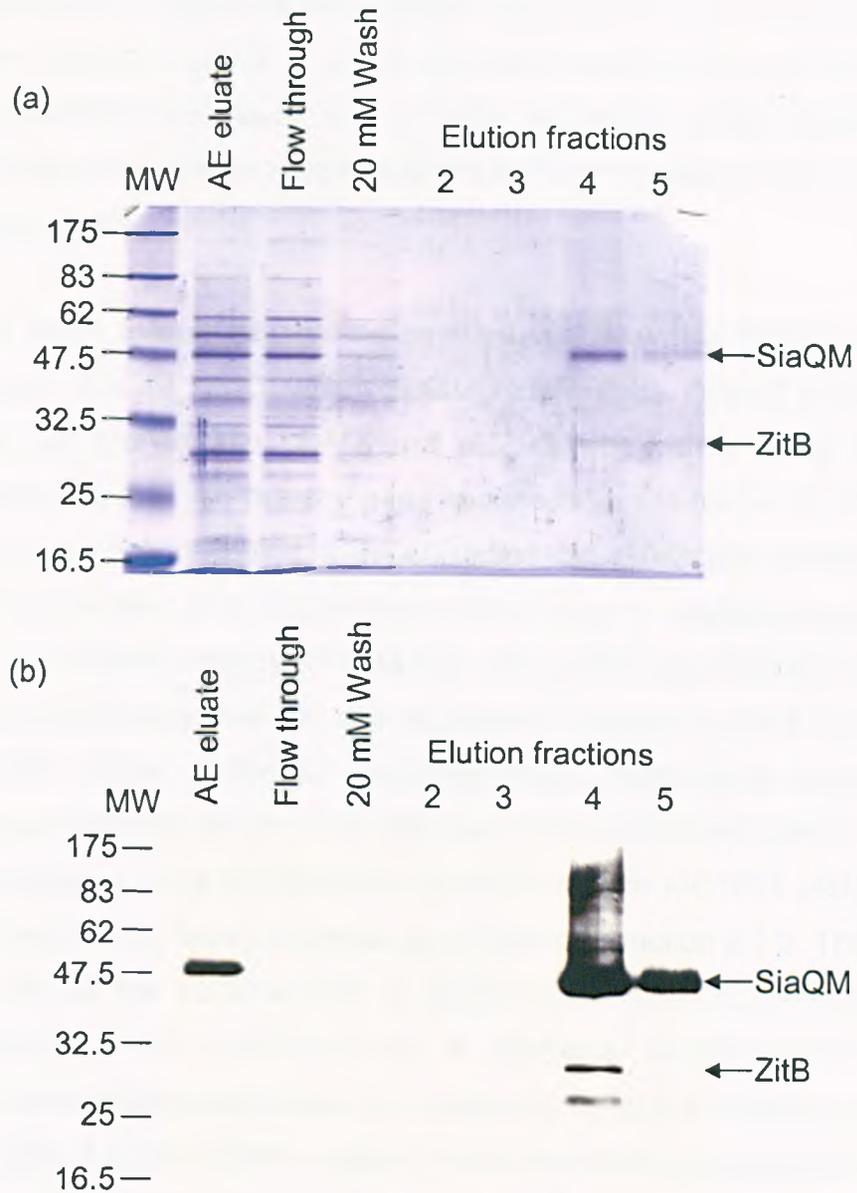


Figure 5.8 IMAC purification of SiaQM from anion exchange eluate. Coomassie stained SDS-PAGE (a) and Western blot (b) of IMAC purification of SiaQM from anion exchange eluate. The following purification samples were analysed; anion exchange (AE) eluate, the flowthrough, the 20 mM wash fraction and the elution fractions. 20 μ g of protein was applied to the Coomassie stained SDS-PAGE and 5 μ g of protein is applied to the Western blots. Molecular weight (MW) shown in kDa.

too much time. Another setback was that the purification protocol could not be reproduced during my research visit to Groningen.

These reasons led to trying to express SiaQM using 2 more expression vectors; the arabinose inducible pBAD in an *E. coli* expression strain and the nisin inducible pNZ8048-derived vector in a *L. lactis* expression strain. SiaQM was cloned into these 2 expression vectors and the level of expression was tested in initial expression trials.

The region of the *H. influenzae* genome encoding *siaQM* was amplified by PCR using primers SiaQM_Nfor with SiaQM_Nrev and SiaQM_Cfor with SiaQM_Crev and cloned into pBAD and pNZ8048-derivative using ligation independent cloning (LIC) so that the gene sequence was in-frame with a 5'- or 3'-nucleotide sequence encoding a decahistidine tag (Methods section 2.6.5 and 2.6.6). This resulted in 4 different expression vectors expressing an N- or C-terminally decahistidine tagged SiaQM from the pBAD and pNZ8048 vectors. The vectors were transformed into the expression strains *E. coli* MC1061 or *L. lactis* NZ9000. Three clones for each expression vector were included in initial expression trials (except for nLIC-QM where only 2 were selected).

The initial expression trials for SiaQM accumulation from MC1061 pBADnQM and MC1061 pBADcQM were performed as in Methods section 2.7.2. The initial expression trial for the accumulation of SiaQM from NZ9000 pNZnQM and NZ9000 pNZcQM were performed as in Methods sections 2.7.3. The membranes from all expression trials were prepared using the Fastprep method (Methods section 2.8.2) and the resultant mixed membrane preparations were analysed using Western blotting.

Analysis of SiaQM accumulation in the three clones of MC1061 pBADcQM and MC1061 pBADnQM using Western blotting revealed the presence of an immunoreactive species at the predicted molecular weight of SiaQM (~50-60 kDa) in two of the pBADcQM clones and in all three pBADnQM clones (Fig. 5.9a). There is a strong Western blot signal in the membrane fraction prepared from the MC1061 pBADcQM 1 culture compared to a much weaker signal from the MC1061 pBADcQM 2 culture (Fig. 5.9a). There is no SiaQM Western blot signal in the sample prepared from MC1061 pBADcQM 3 culture. The intensity

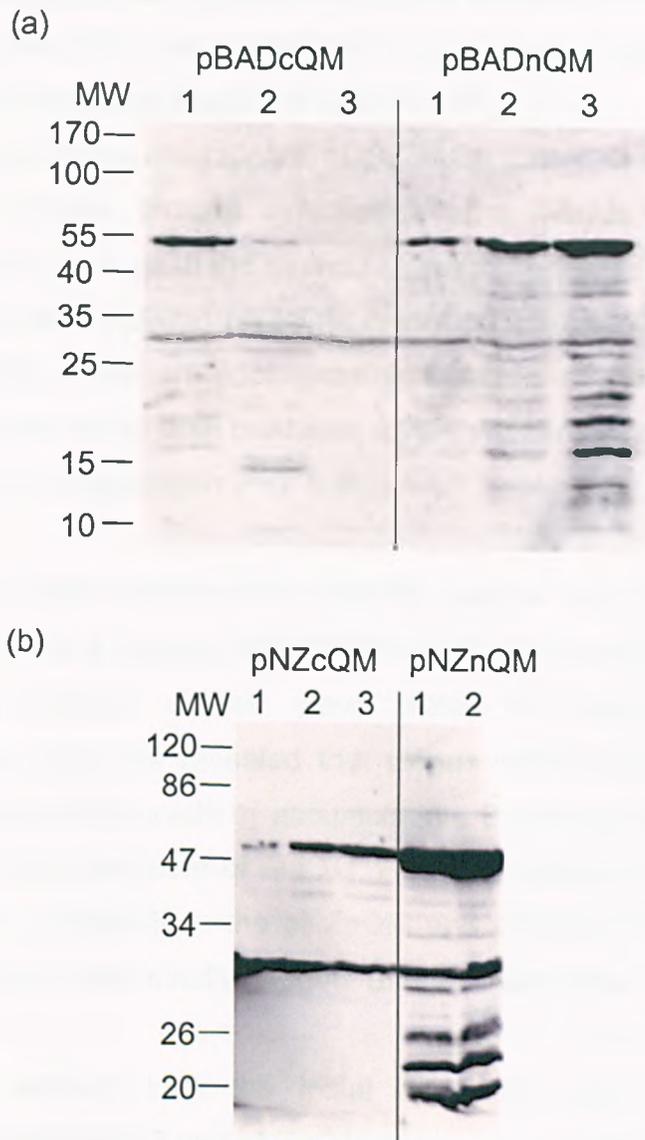


Figure 5.9 Expression of SiaQM from the *E. coli* pBAD expression vector and *L. lactis* pNZ8048 expression vector. (a) Western blot of SiaQM accumulation from 3 independent clones of MC1061 pBADcQM and MC1061 pBADnQM. Cultures were grown to mid-log phase at 37°C, induced with 0.005% L-arabinose and harvested after a further 2 hours growth. **(b)** Western blot of SiaQM accumulation from 3 independent clones of NZ9000 pNZcQM and 2 independent clones of NZ9000 pNZnQM. Cultures were grown to mid-log phase at 30°C, induced with 1/5000 dilution of nisin and harvested after a further 2 hours growth. Equal amounts of protein were added for each sample by normalising by the optical density when harvesting the cells. Molecular weight (MW) shown in kDa.

of the SiaQM signal also differs in intensity between the MC1061 pBADnQM clones with the third one being the most intense; however, with the more intense signal more degradation is apparent (Fig. 5.9a).

Analysis of SiaQM accumulation in the NZ9000 pNZnQM and NZ9000 pNZcQM clones by Western blotting revealed all the clones tested produced an immunoreactive species at the molecular weight expected SiaQM (Fig. 5.9b). In this case, the first NZ9000 pNZcQM clone accumulated less SiaQM than the second or third, which are approximately equal in intensity. The two pNZnQM clones that were tested both produced a very intense signal on the Western blot and evidence of degradation (Fig. 5.9b).

The Western blots indicate that SiaQM can be expressed to measurable amounts using all 4 vectors; pBADnQM, pBADcQM, pNZnQM and pNZcQM. A number of different clones were tested for their expression levels simultaneously and this revealed that clones within the same transformation reaction had differing levels of accumulation. By testing more than one clone, the most effective producer of SiaQM can be selected. The Western blots also indicated that, especially in the pNZnQM and pNZcQM trials, there was more accumulation of N-terminally tagged SiaQM than when the tag is at the C-terminus.

The clones chosen from the initial expression trial were pBADnQM 3, pBADcQM 1, pNZcQM 3 and pNZnQM 2.

5.8 Optimisation of growth and induction conditions for SiaQM expression from MC1061 pBADcQM and MC1061 pBADnQM

The optimisation of growth and induction conditions for the expression of SiaQM from MC1061 pBADnQM and MC1061 pBADcQM was performed to maximize the overall accumulation of SiaQM whilst minimising degradation.

Using the observed accumulation levels in the initial expression trials and previous studies with these expression vectors (Geertsma, 2005) the temperature of growth, the concentration of inducer and the optical density (OD_{660}) of induction were varied.

SiaQM accumulation was optimised in MC1061 pBADnQM and MC1061 pBADcQM by growing the cultures as in the initial expression trials in combinations of the following conditions; at either 37°C or room temperature, induced with either 0.005 or 0.0005 % L-arabinose at optical densities (OD_{660}) of 0.5 or 1.

Analysis of SiaQM accumulation in the membrane fractions prepared from cultures of MC1061 pBADcQM using Western blotting revealed very low levels of immunoreactive material. At the highest exposure, faint bands were present at the estimated molecular weight for SiaQM in only 2 lanes; the sample from cells grown at 37°C, induced at OD_{660} of 0.5 with 0.005% L-arabinose and the sample from cells grown at 37°C induced at OD_{660} of 1 with 0.005% L-arabinose (Fig. 5.10a). In contrast to this, a strong SiaQM Western blot signal is present in each of the membrane fractions prepared from cultures of MC1061 pBADnQM. The SiaQM Western blot signals were of similar intensity for membrane fractions prepared from cells grown at 37°C and induced with either 0.005 or 0.0005% L-arabinose at an OD_{660} of 0.5 (Fig. 5.10b). The OD_{660} of these samples when harvested were 0.584 and 0.804, respectively. The most intense signal was obtained from the membrane fraction prepared from cells grown at 37°C to an OD_{660} of 1 and induced with 0.005% L-arabinose, this had a final OD_{660} of 1.028. In the same conditions, but with a lower concentration of L-arabinose there is a distinctly lower intensity in the SiaQM signal and a higher OD_{660} of 1.236. Samples from cells grown at room temperature and induced with either 0.005% or 0.0005% L-arabinose at an OD_{660} of 0.5 had the lowest OD_{660} values at 0.324 and 0.396, respectively. These optical densities were so low that there was not enough of the sample to load normalised amounts of protein on the gel, and so there is actually approximately half the amount of protein loaded compared to the other fractions on the gel. This was confirmed when the gel was stained after transfer during Western blotting. If the same amount of protein was added then the cells grown at room temperature would probably have higher Western blot signal intensity, but would still have very low optical densities, so would not be as effective at SiaQM production as growing the cells at 37°C and inducing them at an OD_{660} of 1.

(a)

OD ₆₆₀ of induction:	0.5	0.5	1	1	0.5	0.5
Temp. of growth:	37	37	37	37	RT	RT
Conc. of inducer:	10 ⁻³	10 ⁻⁴	10 ⁻³	10 ⁻⁴	10 ⁻³	10 ⁻⁴



(b)

OD ₆₆₀ of induction:	0.5	0.5	1	1	0.5	0.5
Temp. of growth:	37	37	37	37	RT	RT
Conc. of inducer:	10 ⁻³	10 ⁻⁴	10 ⁻³	10 ⁻⁴	10 ⁻³	10 ⁻⁴

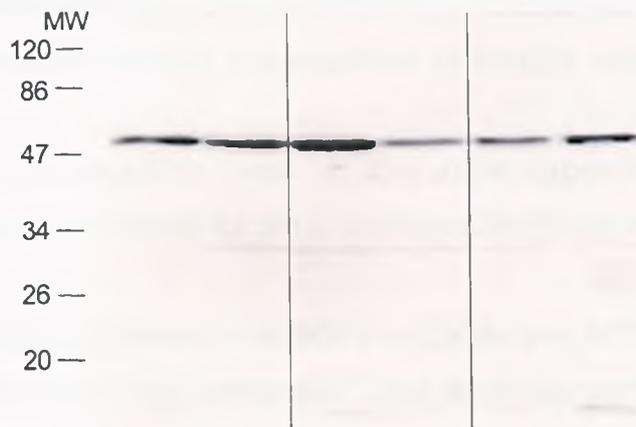


Figure 5.10 Optimisation of temperature of growth, inducer concentration and optical density (OD₆₆₀) of induction for the production of SiaQM from MC1061 pBADcQM and MC1061 pBADnQM. Western blots with anti His-tag antibody of membrane fractions from 5 ml MC1061 pBADcQM (a) and MC1061 pBADnQM (b) cultures grown at either room temperature or 37°C and induced at either OD₆₆₀ of 0.5 or 1 with either 0.005% (10⁻³) or 0.0005% (10⁻⁴) L-arabinose. All samples were harvested after a further 2 hours growth after induction. Equal amounts of protein were added for each sample, except for the last two lanes in (b) which have approx. half the amount of protein. Molecular weight (MW) shown in kDa.

The optimisation of SiaQM accumulation in MC1061 pBADnQM and MC1061 pBADcQM has revealed that N-terminally tagged SiaQM accumulates to a higher degree than the C-terminally tagged version. This is not just a SiaQM-specific observation as this difference in expression levels has been observed for other proteins expressed in the pBAD system (Geertsma E.R., personal communication); however, the reason for this is unknown. This observation indicated that MC1061 pBADnQM should be used for the production SiaQM in *E. coli*. A number of conditions were varied to maximise the output of SiaQM from MC1061 pBADnQM while at the same time minimising degradation. The optimal conditions were identified to be growth at 37°C and induction at an OD₆₆₀ of 1 with 0.005% L-arabinose.

5.9 Optimisation of growth and induction conditions for SiaQM expression from NZ9000 pNZcQM and NZ9000 pNZnQM

The optimisation of the growth and induction conditions for the expression of SiaQM from *L. lactis* NZ9000 pNZnQM and NZ9000 pNZcQM was performed in an attempt to maximise the overall accumulation of SiaQM whilst minimising degradation.

Using the observed accumulation levels in the initial expression trials the concentration of inducer was varied for the *L. lactis* pNZ8048-derived vectors.

SiaQM accumulation was optimised in NZ9000 pNZnQM and NZ9000 pNZcQM by growing cultures as in the initial expression trials (Methods section 2.7.3) and inducing NZ9000 pNZnQM with 1/5000, 1/10000, 1/20000 or 1/50000 dilution of nisin and inducing NZ9000 pNZcQM with either 1/1000 or 1/5000 dilution of nisin.

Analysis of the steady-state levels of SiaQM from cultures of NZ9000 pNZnQM using Western blotting revealed that when NZ9000 pNZnQM is treated with a nisin dilution of 1/5000 there is an intense SiaQM Western blot signal and some degradation, but as the dilution gets larger to 1/10000, 1/20000 and 1/50000, the SiaQM signal intensity decreases along with the appearance of degradation products (Fig. 5.11). In addition, there is an increase in the OD₆₀₀ values for each culture corresponding to the decreasing amount of induction and

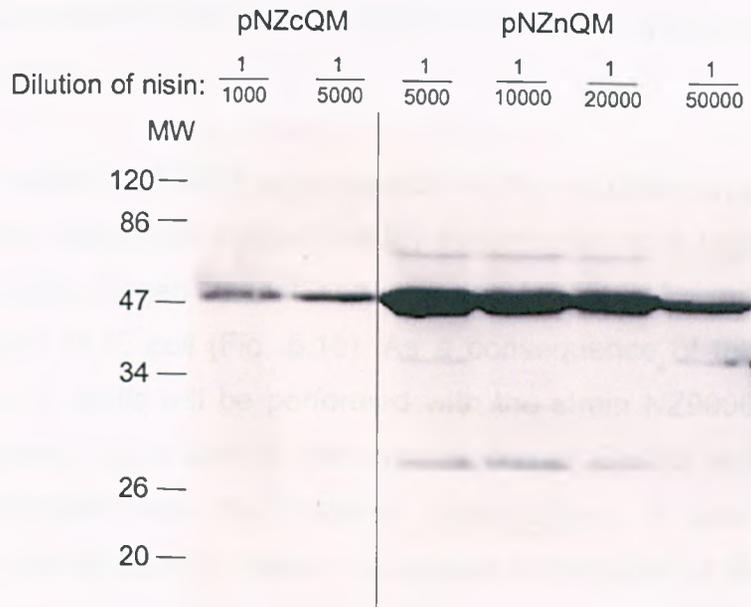


Figure 5.11 Optimisation of inducer concentration for the production of SiaQM from NZ9000 pNZcQM and NZ9000 pNZnQM. Western blot with anti His-tag antibody of membrane fractions from 10 ml NZ9000 pNZcQM and NZ9000 pNZnQM cultures grown at 30°C and induced at OD₆₀₀ of 0.5 with nisin dilutions of 1/1000, 1/5000, 1/10000, 1/20000 and 1/50000. All samples were harvested after a further 2 hours growth after induction.

subsequent lower steady-state levels of SiaQM (Fig. 5.11). The steady-state levels of SiaQM from cultures of NZ9000 pNZcQM did not differ so greatly between 1/1000 and 1/5000 dilution of nisin (Fig. 5.11).

This experiment confirmed that the accumulation of N-terminally tagged SiaQM is far greater than the C-terminally tagged variant as originally seen in the initial expression trial.

The optimisation of SiaQM accumulation in the pNZ8048 system in *L. lactis* revealed that N-terminal tagged SiaQM accumulates to a higher degree than the C-terminally tagged variant as was seen for SiaQM expression from the pBAD system in *E. coli* (Fig. 5.10). As a consequence of this, production of SiaQM from *L. lactis* will be performed with the strain NZ9000 pNZnQM. The only condition to be varied in the optimisation of SiaQM accumulation from NZ9000 pNZnQM was the inducer concentration. It was revealed that decreasing the amount of inducer decreased the amount of SiaQM produced, and also increased the optical density. This phenomenon was also observed in the pBAD system and clearly indicates that it is the production of the integral membrane protein that is decreasing bacterial growth rate.

The optimal conditions were identified as 30°C growth and induction at an OD₆₀₀ of 0.5 with a 1/20000 dilution of nisin. This particular dilution was selected as it is a compromise between SiaQM production and low degradation.

Analysis of steady-state levels of SiaQM from cultures of NZ9000 pNZcQM and NZ9000 pNZnQM with different levels of induction revealed decreasing the concentration of inducer decreased the level of SiaQM accumulation and the appearance of degradation products.

5.10 Purification of N-terminally decahistidine tagged SiaQM by immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC)

IMAC is the purification method of choice for membrane proteins as it typically purifies target proteins to >95% homogeneity and is a relatively short procedure compared to other purification techniques. SEC will be used to purify SiaQM to homogeneity and the elution profile will be used to assess the monodispersity of

the protein produced. Monodisperse SiaQM will have a symmetrical elution peak that is greater in magnitude than any aggregate elution peak that is formed.

Larger scale cultures of MC1061 pBADnQM and NZ9000 pNZnQM, typically 2-5 L, were grown and induced according to the optimised conditions previously established. SiaQM was purified from mixed membrane preparations from each culture by first extracting SiaQM from the membranes with DDM using the standard conditions (Methods 2.14.1). SiaQM was purified using Ni²⁺-affinity chromatography using disposable gravity-flow columns (Methods section 2.14.4). Size exclusion chromatography was performed using Superdex 20 10/300 GL column (Methods section 2.14.7).

Analysis of the fractions taken during the IMAC purification of SiaQM from *E. coli* vesicles using Coomassie stained SDS-PAGE revealed a major band in elution fraction 2 with an apparent molecular weight estimated for SiaQM (~47-50 kDa) (Fig. 5.12a). SiaQM is not abundant enough to be visible in the SDS-PAGE of the mixed membrane sample or the soluble fraction after detergent extraction and ultracentrifugation (Fig. 5.12a). There are a number of minor contaminants in the elution fraction, but the band at ~47-50 kDa is the most abundant.

Homogeneous SiaQM was attained by applying elution fraction 2 to a Superdex 20 10/300 GL size exclusion chromatography (SEC) column. The A₂₈₀ elution profile of the SEC purification revealed a minor peak eluting immediately after the void volume which was presumed to be aggregate (Fig. 5.12b). A major peak over fractions 10 and 11 was evident which, when analysed using SDS-PAGE was revealed to be the ~47-50 kDa band that is presumed to be SiaQM (Fig. 5.12b and c). The SDS-PAGE shows SiaQM to be homogeneous and abundant in fractions 10 and 11 (Fig. 5.12c).

Analysis of the SDS-PAGE and elution profile for the purification of SiaQM from *L. lactis* vesicles revealed that the purification proceeds almost exactly as the purification from *E. coli* vesicles (Fig. 5.13a, b and c). There were some differences between the two purifications; the IMAC elution fraction from the purification of SiaQM expressed in *L. lactis* contained less contaminants and the SiaQM band was more intense than the IMAC elution from the purification of

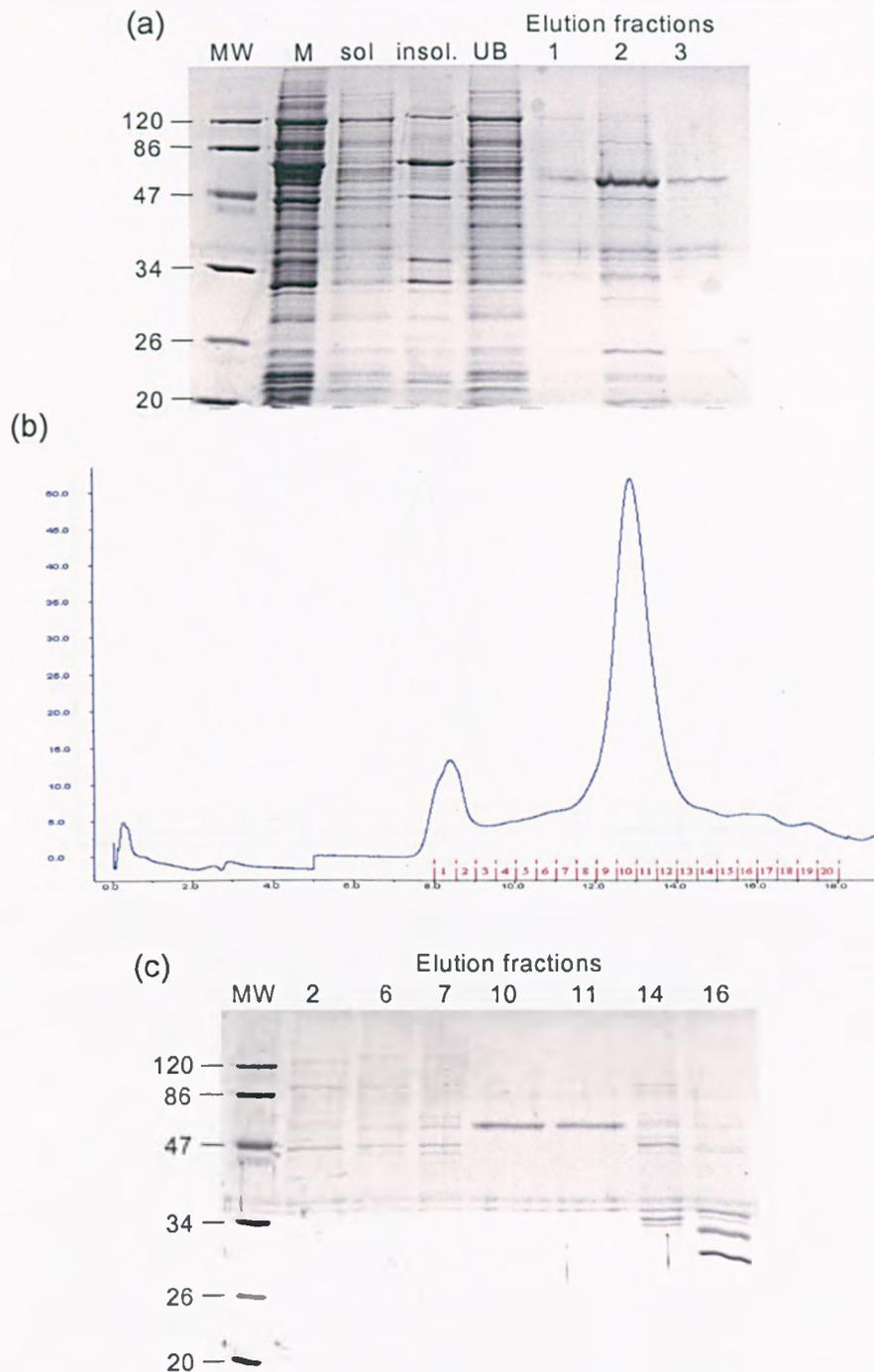


Figure 5.12 IMAC followed by SEC of SiaQM from membrane preparation from MC1061 pBADnQM. **(a)** Coomassie stained SDS-PAGE of SiaQM purification with IMAC showing membrane fraction (M), the soluble fraction (sol) and insoluble fraction (insol) after detergent extraction and centrifugation, the flowthrough (UB) and the elution fractions. **(b)** Elution profile of SEC purification of SiaQM from IMAC elution fraction 2 monitored by absorbance (A_{280}). **(c)** Coomassie stained SDS-PAGE of the fractions taken from SEC purification of SiaQM from IMAC elution fraction 2. Molecular weight (MW) shown in kDa.

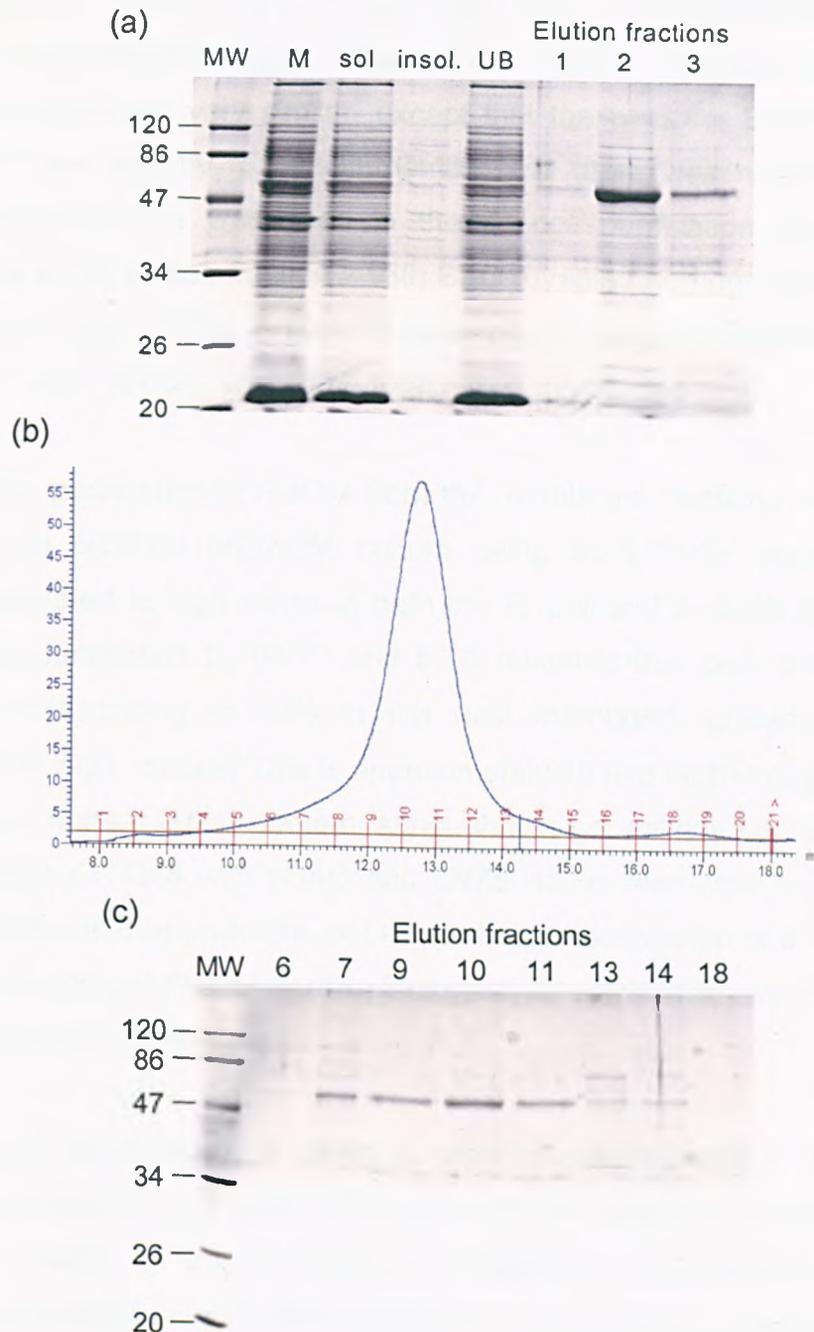


Figure 5.13 IMAC followed by SEC of SiaQM from membrane preparation from NZ9000 pNZnQM. **(a)** Coomassie stained SDS-PAGE of SiaQM purification with IMAC showing membrane fraction (M), the soluble fraction (sol) and insoluble fraction (insol) after detergent extraction and centrifugation, the flowthrough (UB) and the elution fractions. **(b)** Elution profile of SEC purification of SiaQM from IMAC elution fraction 2 monitored by absorbance (A_{280}). **(c)** Coomassie stained SDS-PAGE of the fractions taken from SEC purification of SiaQM from IMAC elution fraction 2. Molecular weight (MW) shown in kDa.

SiaQM expressed in *E. coli* (Fig. 5.13a). The SEC purification elution profiles monitored by absorbance (A_{280}) revealed that SiaQM produced by the two expression strains were very similar, except that the peak for SiaQM purified from *L. lactis* was slightly more symmetrical and there was no evidence of aggregate formation, as was seen in the *E. coli* purification (Fig. 5.13b). Purification of IMAC elution fraction 2 with SEC revealed homogeneous SiaQM in fractions 9, 10 and 11 (Fig. 5.13c). The overall yield of pure SiaQM from both *E. coli* and *L. lactis* vesicles was 2-3% total mixed membrane

Analysis of the purification of SiaQM from the membrane fractions of MC1061 pBADnQM and NZ9000 pNZnQM culture using SDS-PAGE revealed that SiaQM accumulated to high levels in both the *E. coli* and *L. lactis* expression systems. The purification by IMAC and SEC revealed that both proteins are abundant, corresponding to ~3% of the total membrane proteome, which equates to ~1.5 mg/L culture. This is approximately 10 fold higher accumulation of SiaQM than in the pTTQ18 system, although it is not entirely fair to compare expression from pTTQ18 with pBAD and pNZ8048 as there are a number of differences between these vectors, not least of all is the location of the histidine tag. SEC purification of SiaQM from *L. lactis* and *E. coli* also revealed that both proteins are monodisperse.

Aside from the advantages of using *L. lactis* as an expression strain (as mentioned in Introduction section 1.13.1), NZ9000 pNZnQM could be argued to be a better system for the production of SiaQM than MC1061 pBADnQM because it has a slightly higher yield and there is no aggregate formation, unlike the *E. coli* system. However, either system produces adequate amounts of SiaQM for use in *in vitro* N-acetylneuraminic acid uptake assays and other biophysical analysis, such as crystallographic trials.

5.11 Chapter summary

This chapter documents the heterologous expression and purification of the integral membrane protein, SiaQM, from the sialic acid specific TRAP transporter from *H. influenzae*. SiaQM is a fusion of the normally separate large and small TRAP membrane components. Expression of SiaQM was first

observed using the pTTQ18 expression system in *E. coli*. Growth and induction conditions for the accumulation of SiaQM with a C-terminal hexa-histidine tag were optimised. The growth medium, time of harvesting after induction and the temperature of growth were varied simultaneously and optimal expression conditions were defined. SiaQM was found to be localised in the inner membrane of *E. coli*, as predicted and IMAC purification was used to purify SiaQM facilitated by the hexa-histidine tag. Unfortunately, expression from this system had a number of problems; the overall yield of SiaQM was low (0.15% of total inner membrane), there was low affinity between the IMAC resin and the hexahistidine tag and SiaQM was co-expressed with a histidine tag-containing contaminant.

To overcome these problems, accumulation of SiaQM was monitored when expressed from two different expression systems; the arabinose-inducible pBAD vector in *E. coli* or the nisin-inducible pNZ8048 vector in *L. lactis*. In both of these expression systems N- and C-terminal deca-histidine tagged variants were produced. The expression and induction conditions were optimised for both of these systems, which revealed that C-terminally tagged SiaQM accumulated to a much lower degree than the N-terminally tagged variant. IMAC purification of N-terminally tagged SiaQM revealed that the overall yield was between 2-3% of total membrane protein, far superior to the yield obtained from the pTTQ18 system. This yield equates to ~1.5 mg/L culture, which is high enough to facilitate the biochemical and biophysical characterisation of SiaQM.

Chapter 6

Reconstitution of SiaPQM and characterisation of *N*-acetylneuraminic acid transport mechanism

6.1 *In vivo* N-acetylneuraminic acid transport by SiaPQM

Although the overall goal of this project was to purify the components of SiaPQM and reconstitute the transporter for characterisation *in vitro*, there was also very little known about the transport properties *in vivo*. Roughly characterising the *in vivo* transport properties of SiaPQM would give insight into its transport activity and also define initial *in vitro* transport assay conditions to test for successful reconstitution of SiaPQM, i.e. to provide conditions that are known to energise Neu5Ac transport by SiaPQM.

The *in vivo* transport assays were performed using a previously established complementation system (developed by Dr. Emmanuele Severi) where SiaPQM was expressed from a low copy number, IPTG-inducible plasmid (pES7) in an *E. coli* mutant, BW25113 Δ *nanT*, which had the *E. coli* Neu5Ac transporter gene, *nanT*, deleted (Baba et al., 2006). The initial *in vivo* transport assay was designed to monitor transport of radiolabelled Neu5Ac in the presence of a proton motive force (PMF), a Na⁺ gradient and a Li⁺ gradient. The Li⁺ gradient was included as it has been found in other transport systems that Li⁺ and Na⁺ can be used interchangeably due to their similar size and charge (Pourcher et al., 1995).

The *in vivo* transport assays were performed using standard methods (Methods section 2.17). Briefly, 20 ml cultures of BW25113 Δ *nanT* and BW25113 Δ *nanT* pES7 were grown in LB (supplemented with 100 μ g/ml ampicillin) to mid-log phase and induced with 1 mM IPTG. Cultures were grown for a further 2 hours and then harvested. The PMF, Na⁺ gradient and Li⁺ gradient were established using standard methods (Methods section 2.17.2-2.17.4).

Monitoring the uptake of radiolabelled Neu5Ac in these different conditions revealed that transport activity is only observed in the strain BW25113 Δ *nanT* pES7 in the presence of a Na⁺ gradient (Fig. 6.1, closed triangles). The same strain in the presence of a PMF or Li⁺ gradient did not show any uptake of radiolabelled substrate into the cells (Fig. 6.1, closed circles and closed squares, respectively). The strain BW25113 Δ *nanT*, which was not complemented with pES7, showed no transport activity in all three of the energised transport conditions with a basal Neu5Ac accumulation of

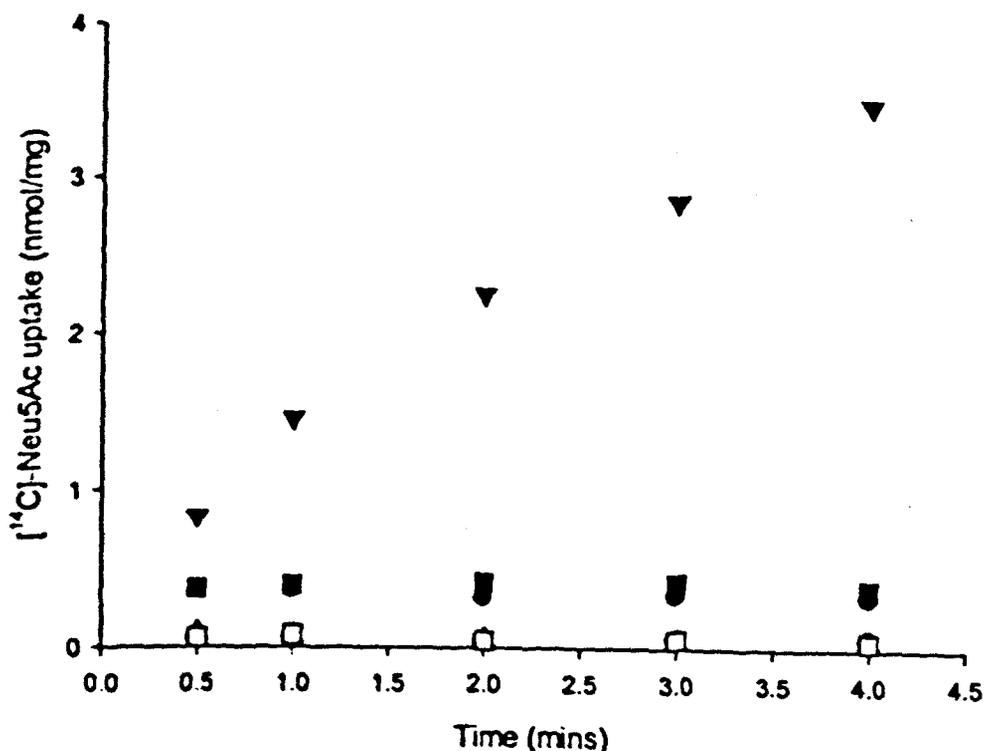


Figure 6.1 *In vivo* Na⁺-driven transport of [¹⁴C]-Neu5Ac by SiaPQM expressed in *E. coli* strain BW25113Δ*nanT*. Uptake of [¹⁴C]-Neu5Ac was monitored into *E. coli* strain BW25113Δ*nanT* and BW25113Δ*nanT* pES7 in the presence of the PMF (BW25113Δ*nanT* pES7-closed circles, BW25113Δ*nanT*-open circles), a sodium gradient (BW25113Δ*nanT* pES7-closed triangles, BW25113Δ*nanT*-open triangles) or a lithium gradient (BW25113Δ*nanT* pES7-closed squares, BW25113Δ*nanT*-open squares). 20 ml cultures of BW25113Δ*nanT* and BW25113Δ*nanT* pES7 were grown in LB (supplemented with 100 μg/ml ampicillin) to mid-log phase and induced with 1 mM IPTG. Cultures were grown for a further 2 hours and harvested. The PMF was created by diluting cells into buffer containing 50 mM KPi, pH 7, 2 mM MgSO₄ and 10 mM D-L-Hactate and incubated in continuous airflow for 2 minutes. A Na⁺ gradient was created by diluting cells into 50 mM KPi, pH 7, 2 mM MgSO₄ and 20 mM NaCl and a Li⁺ gradient was created by diluting cells into buffer containing 50 mM KPi, pH 7, 2 mM MgSO₄ and 20 mM LiCl. All transport assays were performed with 8 μM Neu5Ac at 30°C with constant stirring. Equal amount of cell suspension was added to each assay. A single data set is presented, but the experiment has been performed multiple times.

~0.05 nmol/mg (Fig. 6.1, open symbols). The basal Neu5Ac accumulation in BW25113 Δ *nanT* pES7 in the presence of PMF or a Li⁺ gradient where no transport was observed was ~0.3 nmol/mg. This is significantly higher than the BW25113 Δ *nanT* cultures and most likely due to SiaP in the periplasm of BW25113 Δ *nanT* pES7 binding radiolabelled Neu5Ac, but not able to accumulate it in the cell due to the presence of inappropriate electrochemical gradients.

Although previous transport assays with this *in vivo* system had shown that it complements for Neu5Ac transport (E. Severi, personal communication), individual electrochemical gradients had not been applied and the counter-ion for transport was unknown. This preliminary *in vivo* transport assay has shown the first strong evidence that Neu5Ac transport via SiaPQM is powered by a Na⁺ gradient. Although this finding in itself is noteworthy, the reason for doing this experiment was to choose conditions for the initial *in vitro* experiments to assess the success of reconstitution, which has been fulfilled.

6.2 Cloning, expression trials and expression optimisation of the *siaP* gene using the arabinose inducible pBAD vector in *E. coli*

To characterise functionally the *in vitro* transport properties of SiaPQM it was necessary to express and purify both components, SiaP and SiaQM, to an adequate level. The expression and purification of SiaQM has been demonstrated and optimised previously (Chapter 5). Wild-type SiaP has been purified, biochemically and biophysically characterised, resulting in the publication of the high resolution crystal structure (Muller et al., 2006). Unfortunately, wild-type SiaP from these studies could not be expressed or purified to the levels required for the *in vitro* characterisation of SiaPQM transport.

To improve the yield and ease of purification of SiaP, the gene *siaP* was cloned into the arabinose inducible pBAD vector. The region of the *H. influenzae* genome encoding *siaP* was amplified by PCR using primers SiaP_Nfor_minSP with SiaP_Nrev and SiaP_Cfor with SiaP_Crev and cloned into pBAD using LIC so that the gene sequence was in-frame with a 5'- or 3'-nucleotide sequence encoding a decahistidine tag (Methods section 2.6.5). Expression trials were

performed in *E. coli* strain MC1061 using standard methods (Method section 2.7.2) Accumulation of both C- and N-terminally tagged SiaP was observed (data not shown) and the expressing clones selected from the initial expression trial were named MC1061 pBADnP and MC1061 pBADcP. The optimisation of growth and induction conditions for the expression of SiaP from MC1061 pBADnP and MC1061 pBADcP was performed to maximise the overall accumulation of SiaP whilst minimising degradation. Using the observed accumulation levels in the initial expression trials and previous studies with these expression vectors (Geertsma, 2005) the temperature of growth, the concentration of inducer and the optical density (OD₆₆₀) of induction were varied.

SiaP accumulation was optimised in MC1061 pBADcP by growing the cultures as in the initial expression trials at 37°C, induced with 0.005% L-arabinose at an optical density (OD₆₆₀) of either 0.5 or 1. SiaP accumulation was optimised in MC1061 pBADnP by growing the cultures as in the initial expression trials at either 37°C or room temperature, induced with either 0.005% or 0.0005% L-arabinose at OD₆₆₀ of 0.5. All cultures were harvested 2 hours after induction and whole cell lysates were prepared using the protocol in Methods section 2.8.2, except that the whole cell lysates were obtained by collecting the samples after the 20000xg centrifugation.

Analysis of SiaP accumulation in the whole cell lysate prepared from cultures of MC1061 pBADcP and MC1061 pBADnP using Western blotting revealed that N-terminally tagged SiaP accumulated to higher levels than the C-terminally tagged variant (Fig. 6.2).

The OD₆₆₀ of induction was the only parameter to be varied for MC1061 pBADcP and revealed that there was no difference in accumulation levels when the culture was induced at either OD₆₆₀ of 0.5 or 1, however, there are two bands visible when the culture was induced at an OD₆₆₀ of 0.5 compared to a single band when induced at OD₆₆₀ of 1 (Fig. 6.2). The two bands correspond to the full length SiaP, including the signal peptide and probably localised to the

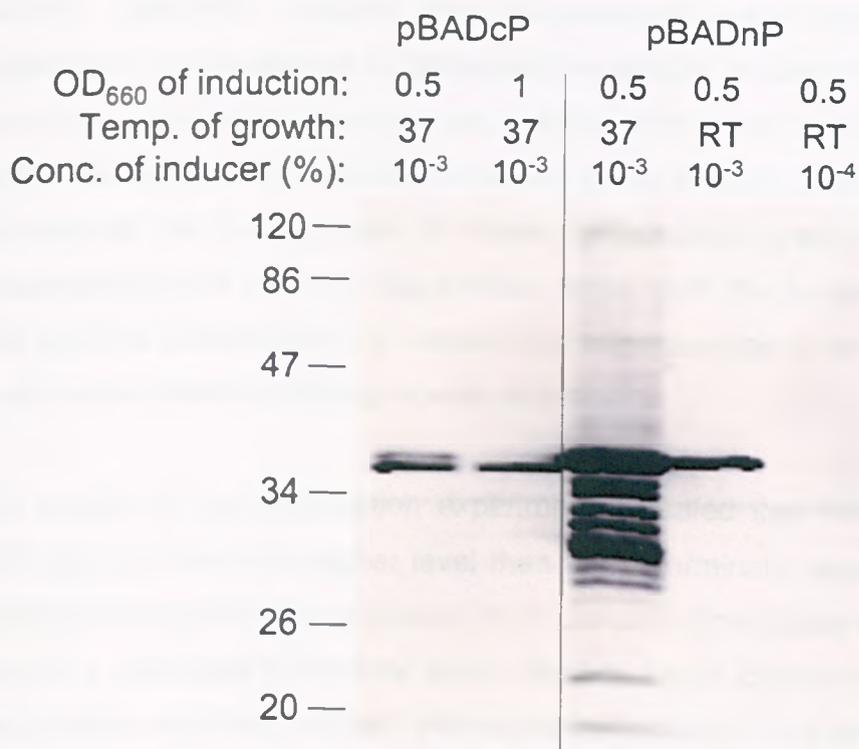


Figure 6.2 Optimisation of temperature of growth, inducer concentration and optical density (OD₆₆₀) of induction for the production of SiaP from MC1061 pBADcP and MC1061 pBADnP. Western blots with anti His-tag antibody of whole cell lysates from 5 ml MC1061 pBADcP and MC1061 pBADnP cultures grown at either room temperature or 37°C and induced at either OD₆₆₀ of 0.5 or 1 with either 0.005% (10⁻³) or 0.0005% (10⁻⁴) L-arabinose. All samples were harvested after a further 2 hours growth after induction. Equal amounts of protein were added for each sample. Molecular weight (MW) shown in kDa.

cytoplasm, and processed SiaP localised to the periplasm without a signal peptide.

The initial expression trial for the accumulation of N-terminally tagged SiaP from MC1061 pBADnP revealed that accumulation was very high, as was degradation. In an attempt to decrease the amount of degradation, the culture was grown at room temperature and induced with either 0.005% or 0.0005% L-arabinose. Analysis of SiaP accumulation using Western blotting revealed that decreasing the temperature to room temperature greatly decreased the accumulation and also the degradation, when both the temperature of growth and the final concentration of L-arabinose were decreased simultaneously only a very weak Western blot signal was visible.

The results of the optimisation experiment revealed that N-terminally tagged SiaP accumulates to a higher level than the C-terminally tagged variant. Most proteins destined for the periplasm in *E. coli* and other Gram negative bacteria contain a cleavable N-terminal signal peptide which directs the protein to the Sec translocon, which actively transports the protein to the periplasm at which point the signal peptide is cleaved (Martoglio & Dobberstein, 1998). C-terminally tagged SiaP was expressed with its cognate signal peptide and was localised to the periplasm, whereas N-terminally tagged SiaP did not have a signal peptide and was localised to the cytoplasm. Analysis of the whole cell lysate from MC1061 pBADcP grown at 37°C and induced with 0.005% L-arabinose at OD₆₆₀ of 0.5 using Western blotting revealed two similarly-sized immunoreactive species (Fig. 6.2). These species most likely correspond to the processed (signal peptide cleaved) SiaP from the periplasm and the unprocessed SiaP from the cytoplasm. This phenomenon was not observed in MC1061 pBADnP cultures as all of the SiaP was localised to the cytoplasm. The extra level of processing required for correct localisation of the C-terminally tagged variant and the higher capacity of the cytoplasm over the periplasm for recombinant proteins may have contributed to the difference in observed accumulation levels. It may, however, be caused by the position of the affinity tag as the same difference in expression levels was seen for N- or C-terminally tagged SiaQM (Chapter 5).

These observations indicate that MC1061 pBADnP should be used for the production of SiaP. Decreasing the growth temperature reduced degradation in MC1061 pBADnP, but also greatly reduced the extent of N-terminally tagged SiaP accumulation; therefore, the optimal conditions for SiaP accumulation were 37°C growth, induced at an OD₆₆₀ of 0.4-0.5 with 0.005% L-arabinose.

6.3 Purification of ligand-free N-terminally decahistidine tagged SiaP by immobilised metal affinity chromatography (IMAC) and anion exchange chromatography

LB medium contains casein which is a highly sialylated protein (Kim *et al.*, 2005), therefore sialic acids, including the substrate of SiaPQM, *N*-acetylneuraminic acid (Neu5Ac) are present in the medium. It has been observed that due to the high affinity interaction between SiaP and Neu5Ac, SiaP produced from cultures grown in LB result in a high percentage of SiaP being pre-bound to Neu5Ac (Severi *et al.*, 2005). The *in vitro* Neu5Ac transport activity of SiaQM is monitored using [¹⁴C]-labelled Neu5Ac, therefore it was necessary to produce a ligand-free version of SiaP, otherwise there would be a large concentration of unlabelled, Neu5Ac being transported which may affect the observed transport activity. *E. coli* can not produce Neu5Ac *de novo* (Vimr & Troy, 1985a); therefore, growing MC1061 pBADnP in a defined minimal medium devoid of Neu5Ac would mean that the SiaP produced would be 100% ligand-free.

A larger scale culture of MC1061 pBADnP, typically 2-5 L, was grown in enhanced M9 minimal medium and induced according to the optimised conditions previously established. SiaP was purified from whole cell lysate preparations, which was prepared by rupturing the cells with 2 passes through a French pressure cell (Thermo) at 10000 psi in the presence of 100 µg/ml deoxyribonuclease I (Sigma) and 1 mM MgCl₂. Cell lysate was incubated on ice for 5 minutes and 5 mM EDTA, pH7 was added. The sample was clarified by centrifugation at 18000 rpm for 20 minutes at 4°C (SS34 rotor in Sorvall Evolution centrifuge). SiaP was purified with a gravity flow column using the standard method (Methods section 2.14.3), except that none of the buffers contained the detergent, DDM. Anion exchange chromatography was

performed with DEAE sepharose using the standard protocol (Methods section 2.14.4).

Analysis of the fractions taken during the IMAC purification of SiaQM from *E. coli* whole cell lysate using Coomassie-stained SDS-PAGE revealed an intense band in elution fraction 2 that corresponded to the estimated molecular weight of SiaPH₁₀ (36.5 kDa, Fig. 6.3). A band was present in the whole cell lysate fraction, which was not present in the flowthrough from the column, indicating that this band was SiaPH₁₀ and was abundant enough to be visible in the whole cell lysate (Fig. 6.3). There was a second intense band in the elution fraction at ~25 kDa that was identified as YodA, which has been found to be a common contaminant of IMAC purification (Bolanos-Garcia & Davies, 2006). YodA is an *E. coli* metal-binding lipocalin, is important in response to cadmium and has been shown to have a high affinity for Ni²⁺ ions. This contaminant was not present when SiaPH₁₀ was produced from cells grown in LB (data not shown); therefore, its production is a consequence of growing the culture in minimal medium. YodA is known to be under the control of ferric uptake regulation (Fur) which is a repressor protein that employs ferrous iron as a co-repressor (Puskarova *et al.*, 2002). M9 minimal medium does not contain a source of ferrous iron and so the addition of FeSO₄ to the medium could repress the expression of YodA. This experiment was performed and in the presence of FeSO₄ there was no reduction in the YodA contamination (data not shown).

Anion exchange chromatography of the IMAC eluate was found to result in homogeneous SiaP with a high yield (Fig. 6.4). The absorbance (A₂₈₀) trace of the elution profile revealed 2 major peaks during the linear gradient from 0-500 mM NaCl (Fig. 6.4a). Analysis of the fractions spanning these peaks by SDS-PAGE revealed that the first peak consisted of the species at approximately the molecular weight expected for SiaPH₁₀ (36.5 kDa) and the second peak corresponded to YodA (Fig. 6.4b). Analysis of the protein content of each fraction revealed that there was a small overlap between the SiaPH₁₀ and YodA peaks, but there were a sufficient number of fractions that contained only

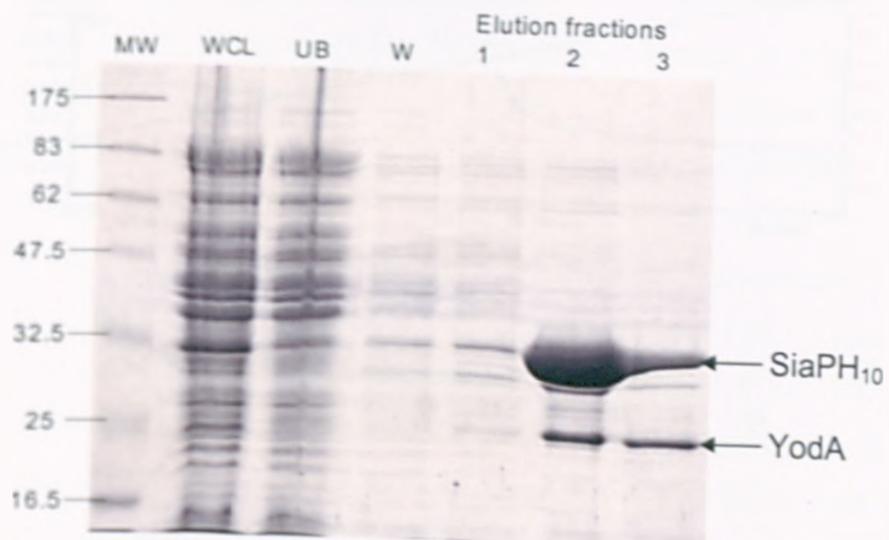


Figure 6.3 IMAC purification of SiaP from clarified whole cell lysate from MC1061 pBADnP. Coomassie-stained SDS-PAGE of SiaP purification with IMAC showing clarified whole cell lysate (WCL), the flowthrough (UB), the 40 mM imidazole washing step (W) and the elution fractions. Molecular weight (MW) shown in kDa.

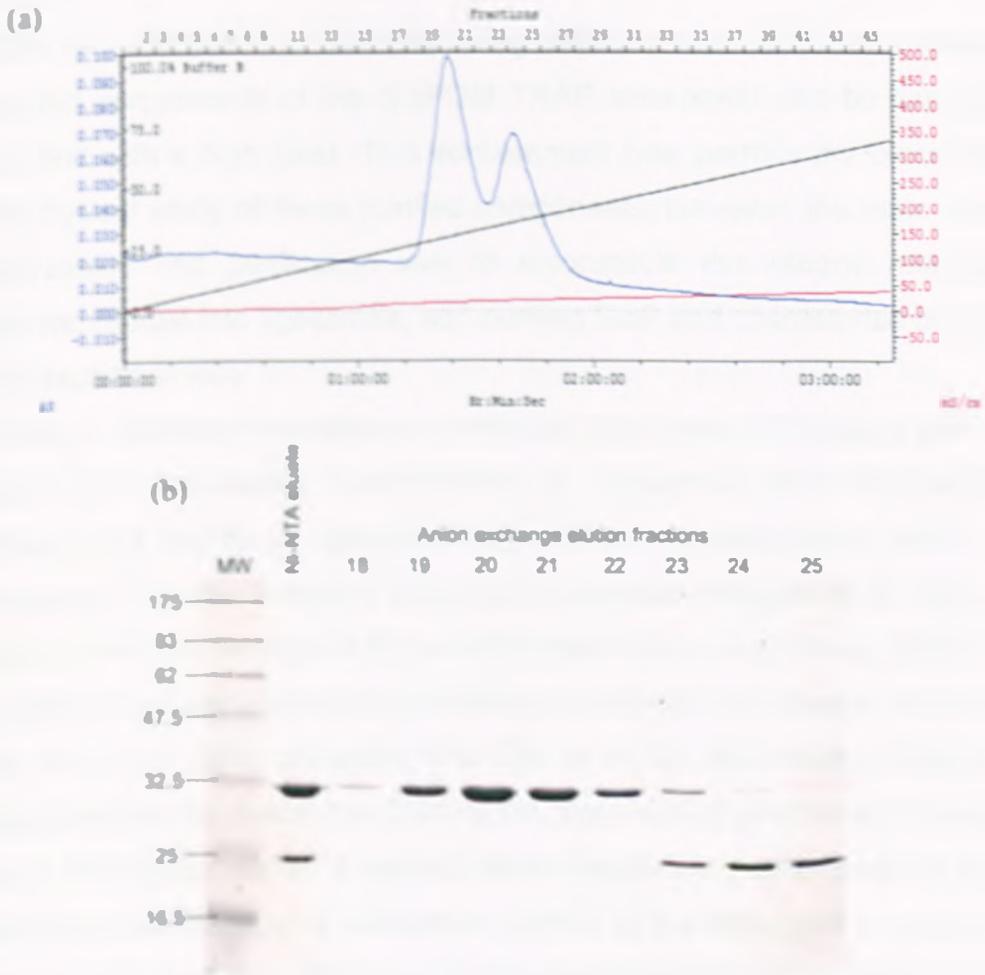


Figure 6.4 Anion exchange chromatography purification of SiaPH₁₀ from Ni-NTA eluate. (a) Elution profile from DEAE sepharose resin (Sigma) monitored by absorbance (A_{260}) of the Ni-NTA eluate (Fig. 6.3). A_{260} (blue line), conductance (red line) and percentage of 500 mM NaCl buffer (black line). (b) Coomassie-stained SDS-PAGE of SiaP purification using anion exchange chromatography. Molecular weight (MW) shown in kDa.

SiaPH₁₀. The total yield of homogeneous, ligand-free SiaPH₁₀ was 2.56 mg/L culture. SiaPH₁₀ is a robust protein that can be concentrated to high levels (up to 1.2 mM) and dialysed into dH₂O.

6.4 Initial reconstitution of SiaPQM using different reconstitution methods

The protein components of the SiaPQM TRAP transporter can be expressed and purified with a high yield. This achievement now permits the biochemical and biophysical study of these purified components, however, the main goal of the expression and purification was to reconstitute the integral membrane component, SiaQM into liposomes, add purified SiaP and characterise transport of [¹⁴C]-Neu5Ac *in vitro*.

A number of different reconstitution methods have been developed and it is apparent that successful reconstitution is subjective and reconstitution conditions have had to be optimised for particular systems (Knol, 1999). It is also apparent that the transport activity of a particular transporter *in vitro* may be only a small percentage of the activity observed *in vivo* (Knol, 1999). The most widely used reconstitution method is solubilising the integral membrane protein in a high CMC detergent like DM or β -OG and rapidly diluting the protein/detergent mix below the CMC in the presence of pre-formed liposomes (Baron & Thompson, 1975). A second reconstitution method that will be tested involves the solubilisation of membrane protein in the detergent of choice, for example, DDM, and then mixing with detergent-destabilised liposomes (using Triton X100, although other detergent have been used previously) followed by removal of the detergents by polystyrene beads. This method has also been subject to rigorous optimisation for other proteins (Knol et al., 1996).

The preliminary *in vivo* uptake assays strongly indicated Na⁺ is the counter ion for Neu5Ac transport, therefore in the preliminary *in vitro* uptake assays an inwardly directed Na⁺ gradient will be the only electrochemical gradient applied.

SiaQM was purified as detailed previously (Chapter 5) and reconstituted using either the detergent-destabilised liposome method (Methods section 2.15.3) with SiaQM solubilised in DDM or the rapid dilution method (Methods section 2.15.2) with SiaQM in the presence of DM or OG.

Standard *in vitro* Neu5Ac transport assays were performed (Methods section 2.161 and 2.16.2) in the presence of an inwardly directed Na⁺ gradient (Methods section 2.16.3.1). A control of empty liposomes, i.e. liposomes with no protein reconstituted, was included.

Analysis of the transport activity of the proteoliposomes prepared by three different methods revealed that, upon creation of an inwardly directed Na⁺ gradient, there was increased uptake of radiolabelled Neu5Ac into the proteoliposomes created using the rapid dilution method with SiaQM solubilised in DM (Fig. 6.5a, closed triangles). The substrate accumulation increased from 700 cpm to 1150 in 80 seconds at which point the uptake rate decreased and the cpm levelled off. Under the same conditions, the proteoliposomes created using the rapid dilution method with SiaQM in the presence of the β -OG (Fig. 6.5a, open triangles) and the proteoliposomes created using the detergent destabilised liposome method (Fig. 6.5a, open circles) did not show any significant increase in radiolabel influx. There was a general upward trend indicating very slow rate of influx; however, this was also observed for the empty liposome control (Fig. 6.5a, closed circles). This suggests that the observed radioactivity increase in these two samples is caused by the gradual diffusion of the [¹⁴C]-Neu5Ac across the liposome bilayer irrespective of the presence of SiaPQM.

Analysis of the protein content of the proteoliposomes prepared using the rapid dilution method with SiaQM in DM by SDS-PAGE revealed that SiaQM is the major species present (Fig. 6.5b).

As in the expression of membrane proteins, the reconstitution of integral membrane proteins is subjective. As in the case of protein expression, the reconstitution method that will work best can not be determined *a priori*, therefore each method must be tested, and possibly optimised, to determine its efficacy of reconstitution.

This preliminary *in vitro* transport assay was designed to determine which, if any, of the three reconstitution methods would result in successful reconstitution of SiaQM. Successful functional reconstitution was assessed by the ability of the reconstituted system to transport [¹⁴C]-Neu5Ac into the lumen of the

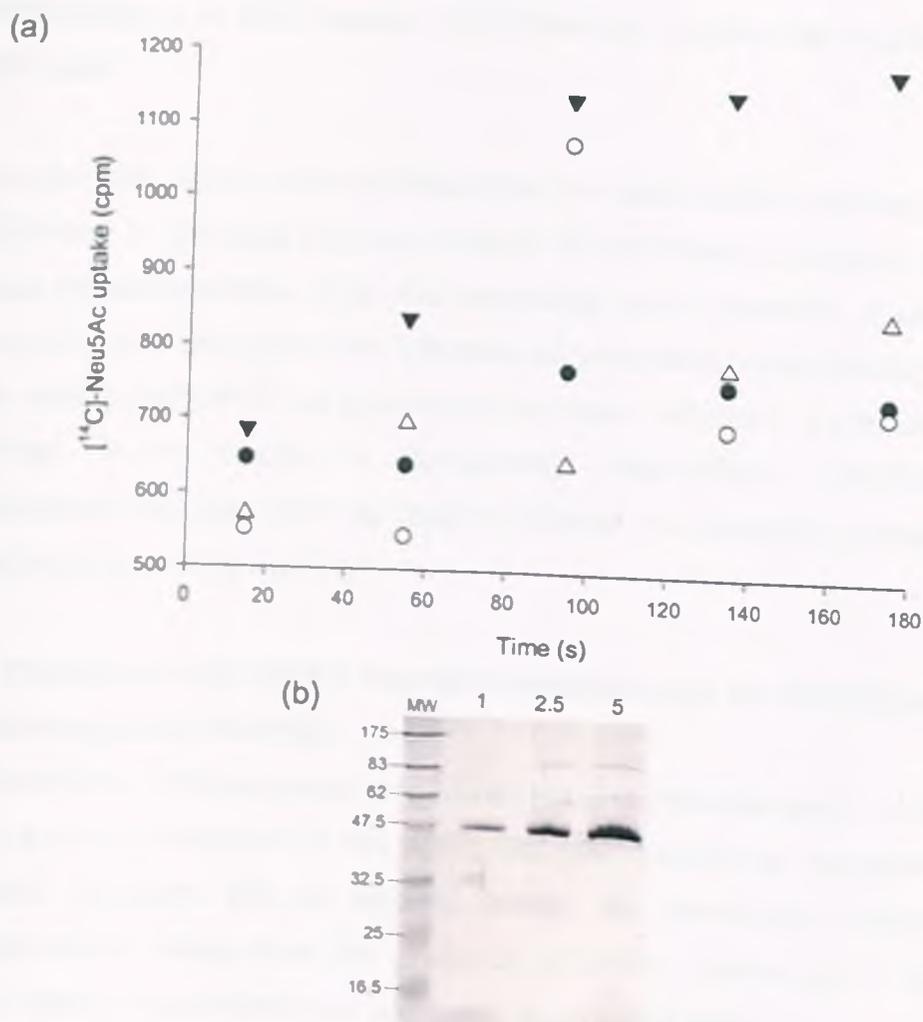


Figure 6.5 Na^+ -driven transport of $[^{14}\text{C}]\text{-Neu5Ac}$ by SiaPQM reconstituted by three different methods. **(a)** Uptake of $[^{14}\text{C}]\text{-Neu5Ac}$ into SiaQM-containing proteoliposomes in the presence of SiaP and an inwardly directed Na^+ gradient ($\Delta\mu\text{Na}^+$). SiaQM was reconstituted using the destabilised liposome method and SiaQM solubilised in DDM (open circles) and the rapid dilution method with SiaQM solubilised in either DM (closed triangles) or $\beta\text{-OG}$ (open triangles). Empty liposomes were included as a control (closed circles). SiaQM-containing proteoliposomes (and the empty liposomes) loaded with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, were diluted into buffer containing 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, to produce the sodium gradient. All reactions included 4 μM $[^{14}\text{C}]\text{-Neu5Ac}$, 4 μM Neu5Ac and 5 μM SiaP. **(b)** Coomassie-stained SDS-PAGE of 1 μl , 2.5 μl and 5 μl of proteoliposome suspension produced from the rapid dilution method with SiaQM solubilised in DM. Single experiments were performed for the purposes of experimental optimisation only. Molecular weight (MW) shown in kDa.

proteoliposome in the presence of an inwardly directed Na^+ gradient and the ESR, SiaP.

The transport assay result indicated that the rapid dilution method with SiaQM solubilised in DM was the only method of the three to produce functionally active proteoliposomes. Even the seemingly small alteration of using SiaQM solubilised in β -OG rather than DM resulted in no observable transport activity. The main objective of this experiment has been achieved: a method has been defined which results in functionally reconstituted SiaPQM. These proteoliposomes can now be used to dissect the energetic requirements of Neu5Ac transport by SiaPQM.

6.5 Characterisation of the energetic requirements for *N*-acetylneuraminic acid transport by SiaPQM

Reconstitution of transporters into liposomes and characterisation *in vitro* allows very specific dissection of the energetic requirements for transport because specific gradients can be applied across the membrane without fear of complications arising from the presence of other proteins (as in vesicles) or other residual electrochemical gradients (as in whole cells).

The initial *in vivo* transport assays revealed that only the creation of a Na^+ gradient across the membrane could power transport of Neu5Ac. The ability of a Na^+ gradient to power transport was also shown in the initial *in vitro* transport assays. This, however, does not preclude other electrochemical gradients, such as a H^+ (ΔpH) gradient or membrane potential ($\Delta\Psi$) from being able to power transport of Neu5Ac via SiaPQM. Transport assays with proteoliposomes can determine the effects of these other gradients when applied individually and in combination.

SiaQM was expressed and purified as detailed previously and reconstituted using the rapid dilution method with SiaQM solubilised in DM. Standard *in vitro* transport assays were performed (Methods section 2.16.1 and 2.16.2) with different combinations of electrochemical gradients created by diluting the proteoliposomes in specific buffers (detailed in Method sections 2.16.3.1-2.16.3.5). The following electrochemical gradients were created; Na^+ gradient

alone ($\Delta\mu\text{Na}^+$), Na^+ gradient and membrane potential combined ($\Delta\mu\text{Na}^+ + \Delta\Psi$), Na^+ gradient and pH gradient combined ($\Delta\mu\text{Na}^+ + \Delta\text{pH}$), membrane potential alone ($\Delta\Psi$) and a pH gradient alone (ΔpH). An empty liposome control was included which had a $\Delta\mu\text{Na}^+$ alone applied across the membrane. 5 μM SiaP and 5 μM [^{14}C]-Neu5Ac were used in each assay. Each assay was performed in triplicate (except for Na^+ gradient alone, which was performed in duplicate) and the average of the three data sets was plotted.

Analysis of the transport data revealed that there was no transport activity when a $\Delta\Psi$ or ΔpH are applied individually (Fig. 6.6, open triangles and closed squares, respectively). Linear uptake into the proteoliposomes is observed when the $\Delta\mu\text{Na}^+$, the $\Delta\mu\text{Na}^+ + \Delta\text{pH}$ and the $\Delta\mu\text{Na}^+ + \Delta\Psi$ were applied across the membrane, but all three have different transport rates (Fig. 6.6, closed circles, closed triangles and open circles, respectively). When the $\Delta\mu\text{Na}^+$ was applied, the rate of transport was 20 nmol/mg/min. The addition of ΔpH to make $\Delta\mu\text{Na}^+ + \Delta\text{pH}$ caused inhibition of the transport rate to ~8 nmol/mg/min. The inclusion of the $\Delta\Psi$ to make $\Delta\mu\text{Na}^+ + \Delta\Psi$ had the opposite effect and increased the rate to 28 nmol/mg/min.

These *in vitro* transport assays have shown that transport of Neu5Ac can not be powered by a pH gradient of membrane potential alone. These electrochemical gradients do, however, have an impact on the rate of transport when in the presence of a Na^+ gradient. Monitoring transport activity when a sodium gradient was applied in conjunction with the membrane potential revealed that the membrane potential has a synergistic effect with the Na^+ -gradient and increased the transport rate. A molecule of Neu5Ac has one negative charge and if it is assumed that only one Neu5Ac is transported at a time by SiaPQM, then the observation that the inclusion of the membrane potential has a synergistic effect on Na^+ driven transport implies that SiaPQM is an electrogenic transporter. If, in contrast, SiaPQM was an electroneutral transporter then the inclusion of a (negative inside) membrane potential would have no effect on the transport rate. This observed positive effect of the (negative inside) membrane potential on transport of Neu5Ac implied that there are more than one Na^+ ions involved in the transport of one Neu5Ac molecule. Further experiments will be

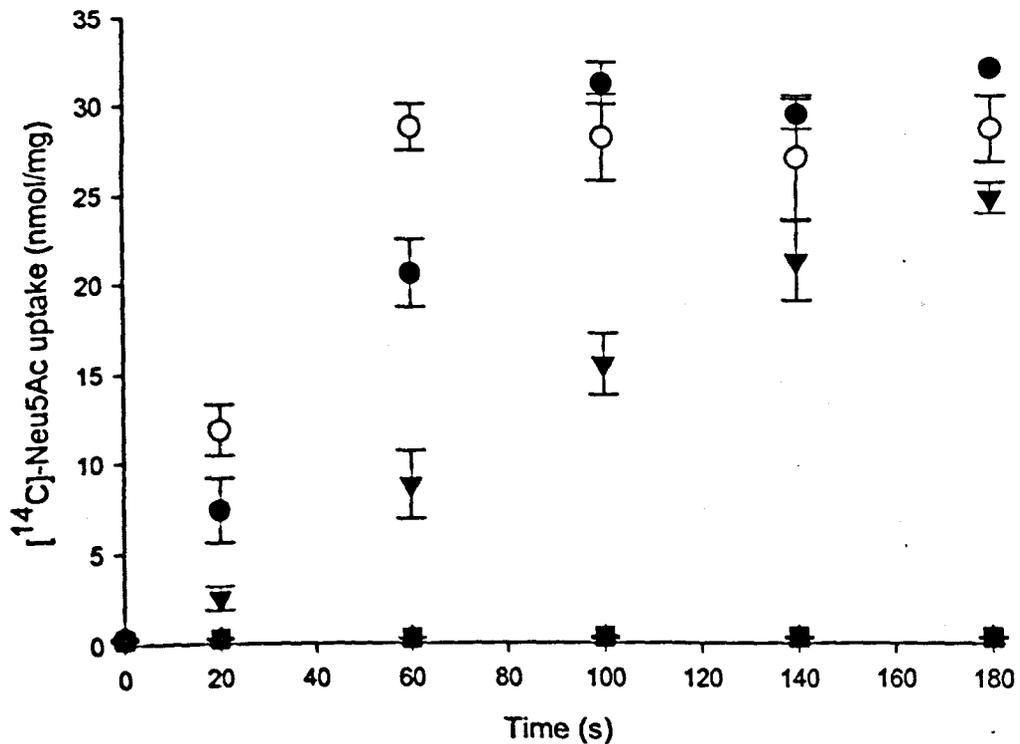


Figure 6.6 Energetic requirement for [^{14}C]-Neu5Ac by SiaQM containing proteoliposomes. Uptake of [^{14}C]-Neu5Ac into SiaQM-containing proteoliposomes in the presence of SiaP and different energetic conditions. SiaQM-containing proteoliposomes loaded with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, were diluted into the following buffers; 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, 2 μM valinomycin to produce a sodium gradient and membrane potential combined ($\Delta\mu\text{Na}^+ + \Delta\Psi$, open circles), 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, to produce a sodium gradient ($\Delta\mu\text{Na}^+$, closed circles), 120 mM sodium PIPES, pH 7, to produce a sodium gradient and pH gradient combined ($\Delta\mu\text{Na}^+ + \Delta\text{pH}$, closed triangles), 100 mM N-methyl glucamine (NMG) acetate, pH 7, 20 mM NMG phosphate, pH 7, to make a membrane potential ($\Delta\Psi$, open triangles), 120 mM potassium phosphate, pH 7, to make a pH gradient (ΔpH , closed squares). All reactions included 5 μM [^{14}C]-Neu5Ac and 5 μM SiaP. The average of triplicate data sets is shown, except for $\Delta\mu\text{Na}^+$, where the average of duplicate data sets is shown. Error bars shown are the standard deviations from the mean, except for $\Delta\mu\text{Na}^+$, in which the range is shown ($n=2$).

required to define the stoichiometry of counter-ion to substrate (see Discussion).

The reasons for the inhibition of Na^+ -driven transport by the addition of a pH gradient (alkaline inside) are unknown. It is unknown whether the pH gradient affected transport or whether the pH change within the lumen of the proteoliposome changed the properties of the protein.

6.6 Requirement of a Na^+ gradient over just the presence of Na^+ for the transport of *N*-acetylneuraminic acid via SiaPQM

It has been demonstrated that SiaPQM can transport Neu5Ac in the presence of a Na^+ gradient, however, it is not clear whether it is the Na^+ gradient or simply the presence of Na^+ that is important for transport. If, for example, Na^+ is only required for the interaction of SiaP with Neu5Ac (like in TakP (Gonin et al., 2007)) or the interaction between SiaP, Neu5Ac and SiaQM then the gradient would not be necessary. To test whether it is the Na^+ gradient that is important and not just presence of Na^+ , a transport assay was performed in which the transport activity of SiaQM-containing proteoliposomes were compared when a Na^+ gradient and membrane potential was applied or when equimolar concentrations of Na^+ were present on both the outside and inside of the proteoliposomes and a membrane potential was applied.

A standard *in vitro* transport assay (Methods section 2.16.1 and 2.16.2) was performed with the following modifications; SiaQM-containing proteoliposomes loaded with 60 mM potassium acetate, pH 7, and 60 mM sodium PIPES, pH 7, were diluted into buffer containing 60 mM *N*-methyl glucamine phosphate (NMG-Pi), pH 7, 60 mM sodium PIPES, pH 7, 2 μM valinomycin to create a situation where equimolar concentrations of Na^+ were present on both the inside and outside the proteoliposome. To produce a $\Delta\mu\text{Na}^+ + \Delta\Psi$, SiaQM-containing proteoliposomes loaded with 60 mM potassium acetate, pH 7, and 60 mM NMG-Pi, pH 7, were diluted into the same outside buffer as above. All reactions included 5 μM SiaP and 5 μM [^{14}C]-Neu5Ac.

Analysis of the transport activity of SiaQM-containing proteoliposomes in the presence of SiaP and a Na^+ gradient with a membrane potential (negative

inside) revealed rapid transport of the [^{14}C]-Neu5Ac into the lumen of the proteoliposomes reaching ~ 4.8 nmol/mg after 180 s (Fig. 6.7, open circles). The transport activity of proteoliposomes in the presence of a membrane potential and equimolar concentrations of Na^+ was negligible in comparison (Fig. 6.7, closed circles). There is a slight increase during the assay indicating slow transport into the proteoliposome. This is unlikely to be due to Na^+ -gradient independent transport and could be a consequence of a small Na^+ gradient being established due to the electrostatic attraction of Na^+ to the negatively charged lumen (relative to the outside) formed by the membrane potential. Monitoring the transport activity in these two different conditions revealed that it is indeed a Na^+ gradient that is required for transport and not the mere presence of Na^+ .

6.7 Investigation into *N*-acetylneuraminic acid transport by SiaQM and SiaP orthologue, VC1779

The process of transporting a substrate through an ESR-dependent transporter like SiaPQM consists of multiple steps. The ESR, SiaP, has to bind Neu5Ac, and then interact with the translocation machinery SiaQM. The binding of substrate to ESR has been studied in depth for ABC transporters as well as a number of TRAP ESRs (See Introduction section 1.7.3.1 and 1.8.1.2). The interaction of the ESR with ABC transporter integral membrane protein has also been studied using mutagenesis and more recently using co-crystallisation (Locher et al., 2002, Oldham et al., 2007). These studies have revealed that surface charge is the major driving force for the interaction between ESR and integral membrane protein. It is unknown, however, how the liganded (or unliganded) TRAP ESRs interact with their cognate integral membrane proteins. A possible method to determine which regions of SiaP confer specificity to SiaQM would be to determine whether a close orthologue of SiaP can form a functional transporter with SiaQM. VC1779 from *Vibrio cholerae* is known to bind Neu5Ac with a high affinity and has ~ 51 % identity to SiaP. If this ESR can interact functionally with SiaQM then conserved, solvent exposed regions of the two ESRs would be good targets for the region of interaction. Conversely, if VC1779 cannot interact functionally with SiaQM then solvent exposed regions where these very similar proteins differ could be investigated as the possible

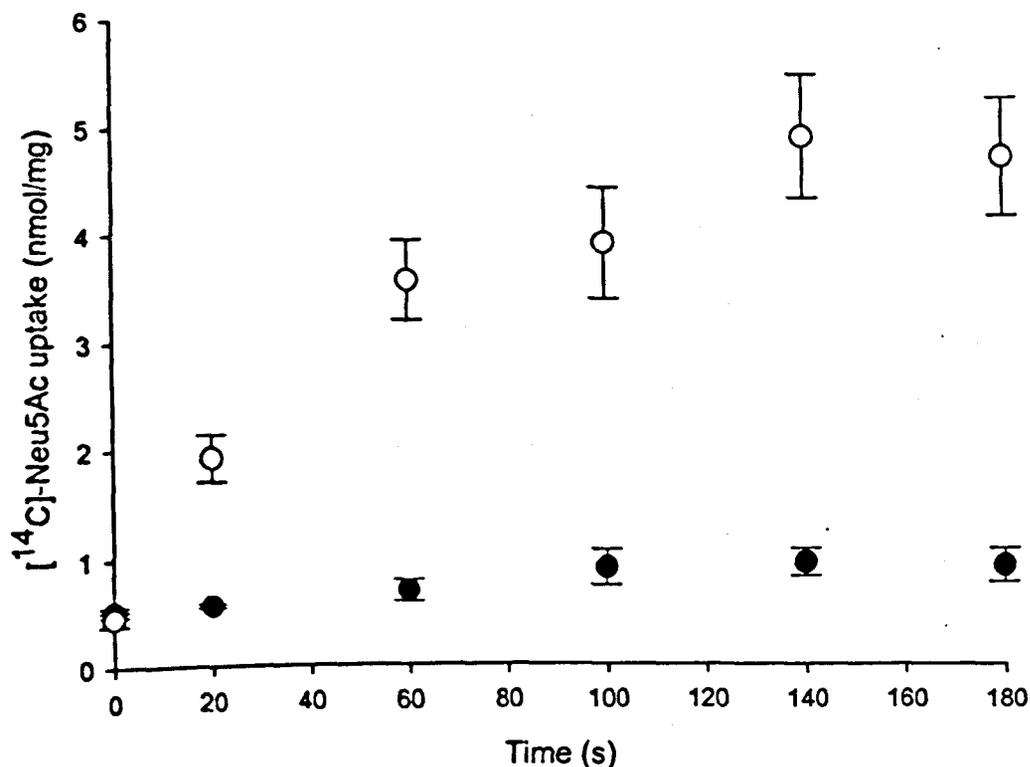


Figure 6.7 Transport of [¹⁴C]-Neu5Ac in the presence and absence of a Na⁺ gradient. Uptake of [¹⁴C]-Neu5Ac into SiaQM-containing proteoliposomes in the presence of SiaP and an inwardly-directed Na⁺ gradient and membrane potential ($\Delta\mu\text{Na}^+ + \Delta\Psi$, open circles) or in the absence of a Na⁺ gradient but with equimolar Na⁺ on both sides of the membrane and a membrane potential (closed circles). SiaQM-containing proteoliposomes loaded with 60 mM potassium acetate, pH 7, and 60 mM sodium PIPES, pH 7, were diluted into buffer containing 60 mM NMG-Pi, pH 7, 60 mM sodium PIPES, pH 7, 2 μM valinomycin to create a situation where equimolar concentrations of Na⁺ were present on both the inside and outside the proteoliposome. To produce the $\Delta\mu\text{Na}^+ + \Delta\Psi$, SiaQM-containing proteoliposomes loaded with 60 mM potassium acetate, pH 7, and 60 mM NMG-Pi, pH 7, were diluted into the same outside buffer as above. All reactions included 5 μM SiaP and 5 μM [¹⁴C]-Neu5Ac. The average of triplicate data is presented and error bars shown are the standard deviations from the mean.

region of interaction. To perform this experiment, VC1779 had to be purified and tested for its ability to bind Neu5Ac with a high affinity before it could be included in an *in vitro* uptake assay.

A more thorough investigation into the differences between SiaP and VC1779 is possible because the high resolution crystal structures are available for both ((Muller et al., 2006) and S. Ramaswamy, unpublished data).

The expression strain BL21 (DE3) plysS pRY11 was used to produce ligand-free VC1779. 6 L of double M9 minimal medium was inoculated to an OD₆₅₀ of 0.1 from an overnight culture of BL21 (DE3) pRY11. The culture was grown at 25°C to an OD₆₅₀ of 0.4-0.5 and expression was induced with 0.5 mM IPTG. The cells were harvested after overnight growth. Periplasmic fraction was prepared (Method section 2.8.5) and ligand-free VC1779 was purified using a HisTrap column (GE Healthcare) and standard methods (Methods section 2.14.4).

Tryptophan fluorescence spectroscopy was performed using a Fluoromax 3 with an excitation wavelength of 295 nm and an emission wavelength of 328 nm. A 25 nM solution of VC1779 diluted into 50 mM Tris, pH 8 was used. A time-based acquisition was performed in which increasing concentrations of Neu5Ac were added every 30 seconds up to a final concentration of 8 μM. The cumulative fluorescence change (ΔF) was plotted against concentration of Neu5Ac and the K_d was calculated by fitting the curve to a single rectangular hyperbola.

A standard uptake assay was performed (Methods section 2.16.1 and 2.16.2) with a $\Delta\mu_{Na^+} + \Delta\Psi$ applied across the membrane (Methods section 2.16.3.2) in presence of 5 μM SiaP or 5 μM VC1779.

Analysis of the SiaP (PDB code: 2CEY) and VC1779 (S. Ramaswamy, unpublished data) structures was performed using the program CCP4MG.

VC1779 is expressed in-frame with a C-terminal hexahistidine tag and it is expressed to a high degree in the expression strain BL21 (DE3) pRY11 grown in M9 minimal media, therefore, it is purified in its ligand-free form. The expression is from the pET20b plasmid (Novagen) and the protein is directed to the periplasm by fusion to the *E. coli* PelB signal peptide. Preparation of the

periplasm can be considered the first purification step as the contents of the cytoplasm are not released during the procedure. VC1779 is purified to ~99% homogeneity in one IMAC purification step (Fig. 6.8, inset). VC1779 was dialysed into 50 mM KPi, pH 7, as it has been found to be less stable than SiaP and it was unknown whether it would be soluble in dH₂O (E. Severi, personal communication).

Tryptophan fluorescence spectroscopy revealed that at an emission wavelength of 328 nm the binding of Neu5Ac causes a ~30% quench in the fluorescence signal (data not shown). The binding affinity (K_d) can be determined by titrating the extent of fluorescence quenching with increasing concentrations of Neu5Ac (Fig. 6.8). The affinity of Neu5Ac for VC1779 was calculated to be $0.113 \mu\text{M} \pm 0.002$ (an average of three experiments). This indicates that the VC1779 purified was active and was able to bind Neu5Ac with a high affinity. A second experiment was performed to determine the effects of the presence and absence of Na⁺ on the binding affinity of Neu5Ac to VC1779. Purified VC1779 was dialysed into Na⁺-free buffer and a fluorescence titration was performed as before, either in the absence of Na⁺ or in the presence of 120 mM NaCl. The affinity of Neu5Ac for VC1779 in the absence of Na⁺ was calculated to be $0.111 \mu\text{M} \pm 0.037$ (an average of four experiments, data not shown). In the presence of 120 mM NaCl the affinity was calculated to be $0.347 \mu\text{M} \pm 0.113$ (an average of four experiments, data not shown). This data suggests that Na⁺ is not integral to the binding mechanism of VC1779.

The *in vitro* transport assays revealed that there was no transport activity when VC1779 was included in the outside buffer (Fig. 6.9, closed circles), however, the same experiment performed with SiaP with the same proteoliposomes revealed uptake of radiolabelled Neu5Ac (Fig. 6.9, open circles). This indicated that, although VC1779 can bind Neu5Ac with a high affinity, it is not capable of interacting functionally with SiaQM to transport Neu5Ac.

High resolution crystal structures have been elucidated for ABC transporter integral membrane components in complex with their cognate ESRs for the maltose transporter (Oldham et al., 2007) and the molybdate transporter (Hollenstein et al., 2007b). The structures have shown that the ESR is positioned on the integral membrane component so that the mouth of the binding cleft is in-line with the opening of the translocation channel. These

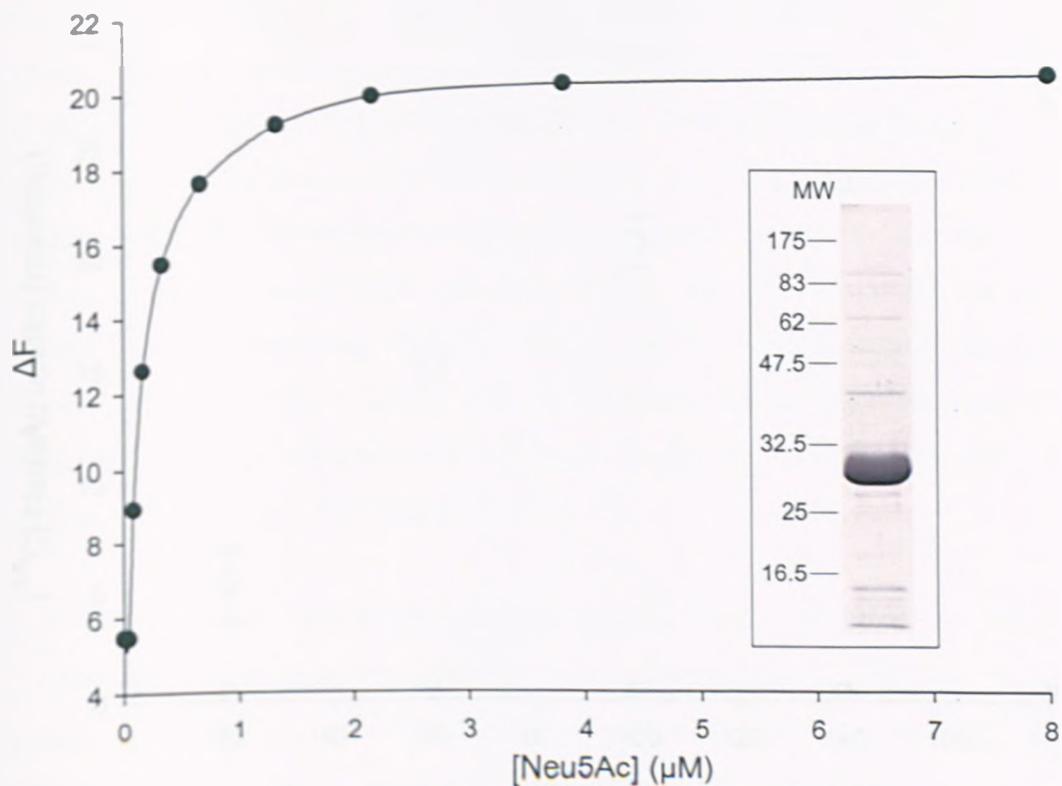


Figure 6.8 Neu5Ac concentration-dependent fluorescence change in VC1779. Fluorescence changes measured at 328 nm upon addition of increasing concentrations of Neu5Ac. Data were fitted to a single rectangular hyperbola. Inset: Coomassie-stained SDS-PAGE gel of VC1779 purified by Ni²⁺ affinity chromatography. Triplicate data sets were collected. Molecular weight (MW) shown in kDa.

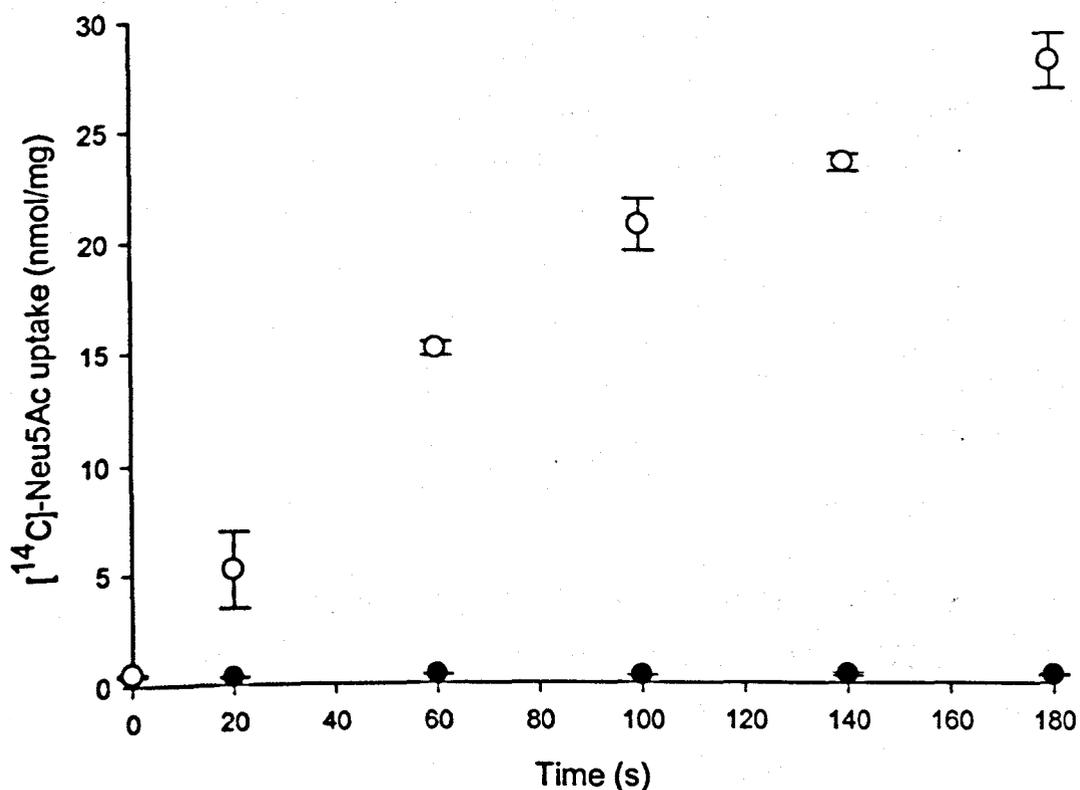


Figure 6.9 Uptake of [¹⁴C]-Neu5Ac by SiaQM in the presence of $\Delta\mu\text{Na}^+ + \Delta\Psi$ and either SiaP or VC1779. Uptake of [¹⁴C]-Neu5Ac into SiaQM-containing proteoliposomes in the presence of SiaP (open circles) or VC1779 (closed circles) and an inwardly directed Na^+ gradient plus negative-inside membrane potential ($\Delta\mu\text{Na}^+ + \Delta\Psi$). SiaQM-containing proteoliposomes loaded with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, were diluted into buffer containing 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, 2 μM valinomycin. All reactions included 5 μM [¹⁴C]-Neu5Ac. The average of duplicate data sets is presented. Error bars shown indicate the range between the two sets of data (n=2).

structures and other studies on ABC transporter ESRs (Prossnitz, 1991, Sebulsky *et al.*, 2003) have revealed that the interaction between the ESR and the integral membrane components is driven primarily by charged interactions between the proteins. It has been demonstrated that VC1779 cannot form a functional transporter with SiaQM and so the tertiary structure of SiaP and VC1779 were compared. This revealed that the backbone architecture was almost identical, which is remarkable considering the proteins are only 50 % identical at the amino acid level (Fig. 6.10a). Analysis of the surface exposed electrostatic properties of these two proteins on the protein face predicted to interact with the membrane components revealed a number of differences in charge distribution over this area (Fig. 6.10b and c). The conserved surface exposed region previously identified (Muller *et al.*, 2006) is essentially identical in both proteins (Fig. 6.10b and c). Unfortunately, without experimental evidence it would be extremely difficult to predict which residues on this interface differentiate SiaP from VC1779.

VC1779 is a close orthologue of SiaP and is part of a TRAP transporter in *V. cholerae*. *In vitro* transport assays have revealed that VC1779 can not form a functional transporter with SiaQM. The reason VC1779 can not function with SiaQM is likely due to variations in residues and electrostatic properties in regions of VC1779 that sterically or electrostatically hinder interaction with the reciprocal interaction patch on SiaQM.

6.8 Directionality of *N*-acetyl neuraminic acid transport by SiaPQM

One of the characteristics of secondary transporters is the ability to transport substrate reversibly, that is, depending on the direction of the substrate and counter-ion gradient, the substrate can be transported in either direction (See Introduction section 1.5.4). Although TRAP transporters are ESR-dependent, they are still secondary transporters and may have some residual secondary transporter characteristics. The ESR in the TRAP transporter confers a higher affinity to TRAP transporters than is found in secondary transporters, and ESRs also define the direction of transport, however, this does not mean it is not possible for the TRAP transporter to work reversibly. To assess this, two *in vitro*

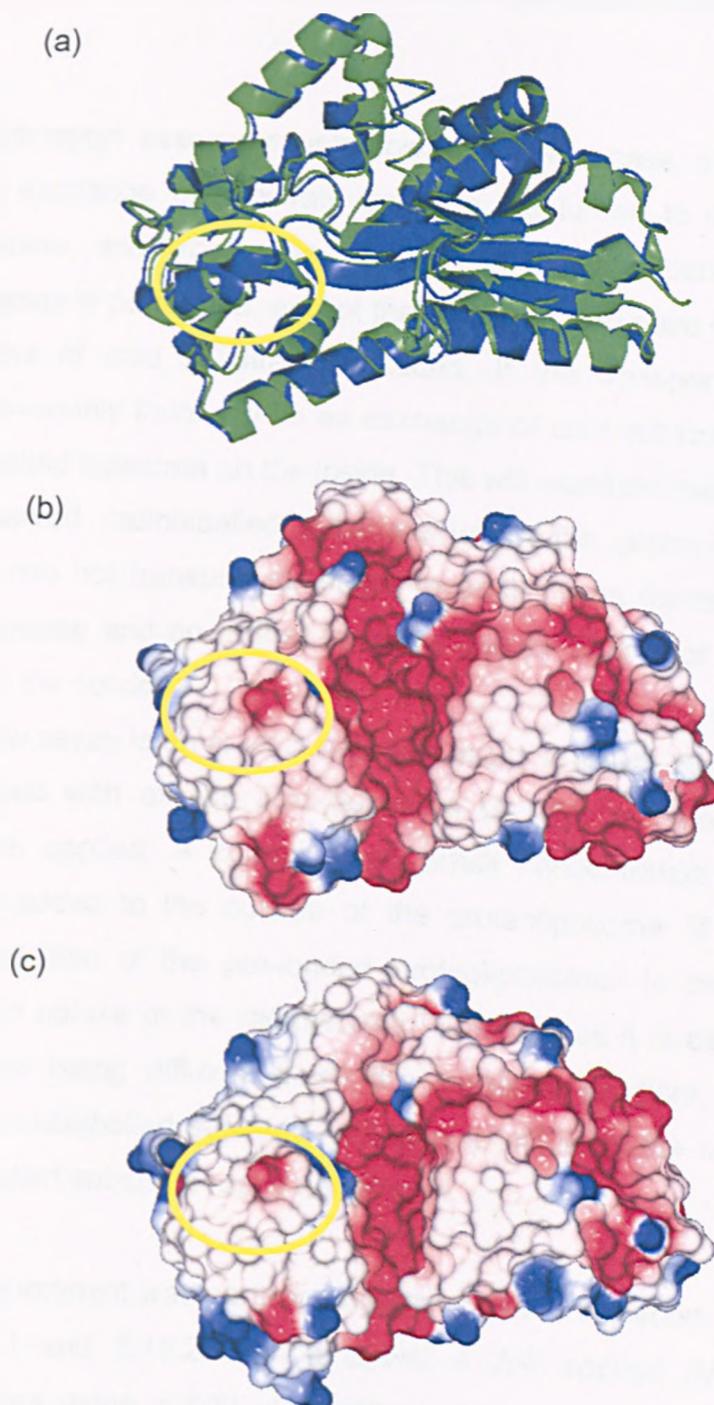


Figure 6.10 Comparison of the tertiary structure and surface exposed electrostatic characteristics of SiaP and VC1779. (a) Superimposition of carbon backbone of SiaP (green) and VC1779 (blue). Electrostatic characteristics of the face of (b) VC1779 and (c) SiaP predicted to interact with the integral membrane protein components. The conserved, surface exposed region predicted to be the interaction patch (Muller et al., 2006) is outlined in yellow.

transport assays were performed; a counterflow assay and a cold substrate chase.

The two transport assays; counterflow and cold chase are both a way of measuring exchange of substrate between the lumen to the outside of the proteoliposome, and *vice versa*. In the cold chase experiment, a standard transport assay is performed, except that at a specified point during the assay a large excess of cold substrate is added. If the transporter can transport substrate reversibly there will be an exchange of cold substrate on the outside for radiolabelled substrate on the inside. This will manifest itself with a decrease in the observed radiolabelled substrate within the proteoliposomes. If the transporter can not transport substrate reversibly then transport will level out with no decrease and no further increase due to dilution of the radiolabelled substrate on the outside.

A counterflow assay is an *in vitro* transport assay in which the proteoliposomes are pre-loaded with excess cold substrate (1 mM) and no electrochemical gradients are applied. A comparatively small concentration of radiolabelled substrate is added to the outside of the proteoliposome. If the transport is reversible, addition of the pre-loaded proteoliposomes to the outside buffer leads to rapid uptake of the radiolabelled substrate as it is exchanged for the cold substrate being effluxed down the substrate gradient. Eventually, the internalised radiolabelled substrate will start to efflux as the radiolabelled and non-radiolabelled substrate reach equilibrium.

The chase experiment was performed like a standard transport assay (Methods section 2.16.1 and 2.16.2) with a $\Delta\mu\text{Na}^+ + \Delta\Psi$ applied (Methods section 2.16.3.2) except using a 600 μl reaction volume and taking 50 μl samples. Either dH_2O or 1 mM cold Neu5Ac was added at 100 seconds (as indicated in the figure).

The counterflow experiment (Methods section 2.16.4) was performed in the presence and absence of 5 μM SiaP. Empty liposome controls were added for experiments in the presence and absence of SiaP.

There was no significant difference between the transport activities for assays

where a 1 mM Neu5Ac chase or dH₂O was added (Fig. 6.11). In both assays, transport of Neu5Ac increased quickly, and after 60 seconds the transport rate levelled off. After 150 seconds there was a slight decrease in the Neu5Ac uptake indicating efflux of substrate from the proteoliposomes. This phenomenon is observed for both the experiment where 1 mM Neu5Ac is added and the control experiment where dH₂O is added, therefore this is not a consequence of the 1 mM cold Neu5Ac chase. After this slight decrease in uptake, the rate levels off again and decreases no further (Fig. 6.11).

The counterflow assay was performed in the presence of Na⁺ as it has been shown that Na⁺ is required for transport, however, this was not a Na⁺ gradient as Counterflow assays are performed in the absence of electrochemical gradients. The experiments were also performed in the presence and absence of SiaP. To give a reference point for the transport activity data, the counterflow assays were plotted with the data from the cold chase control in which only dH₂O was added. The resultant graph revealed that there was no transport activity in the counterflow assay in the presence (Fig. 6.12a) or absence (Fig. 6.12b) of SiaP.

The results obtained from the counterflow and cold chase experiments strongly suggest that SiaPQM is unidirectional in the conditions used. A positive control for these experiments would have made this assertion more robust, for example, reconstituted NanT from *E. coli*, which is a classical secondary Neu5Ac transporter and would have shown bi-directional transport. Unfortunately, this positive control was not available; however, the counterflow assay and the cold chase experiments are well established and have been used on a number of occasions with other transporters *in vitro* and *in vivo*. Indeed, the cold chase experiment has been performed on SiaPQM *in vivo*, with NanT as a control, both expressed in *E. coli*. The results indicated that although bidirectional transport could be observed for NanT, it could not be observed for SiaPQM. (Dr. Emmanuele Severi, personal communication). This reflects the result from the *in vitro* experiment presented and adds weight to the assertion that SiaPQM can only transport Neu5Ac unidirectionally.

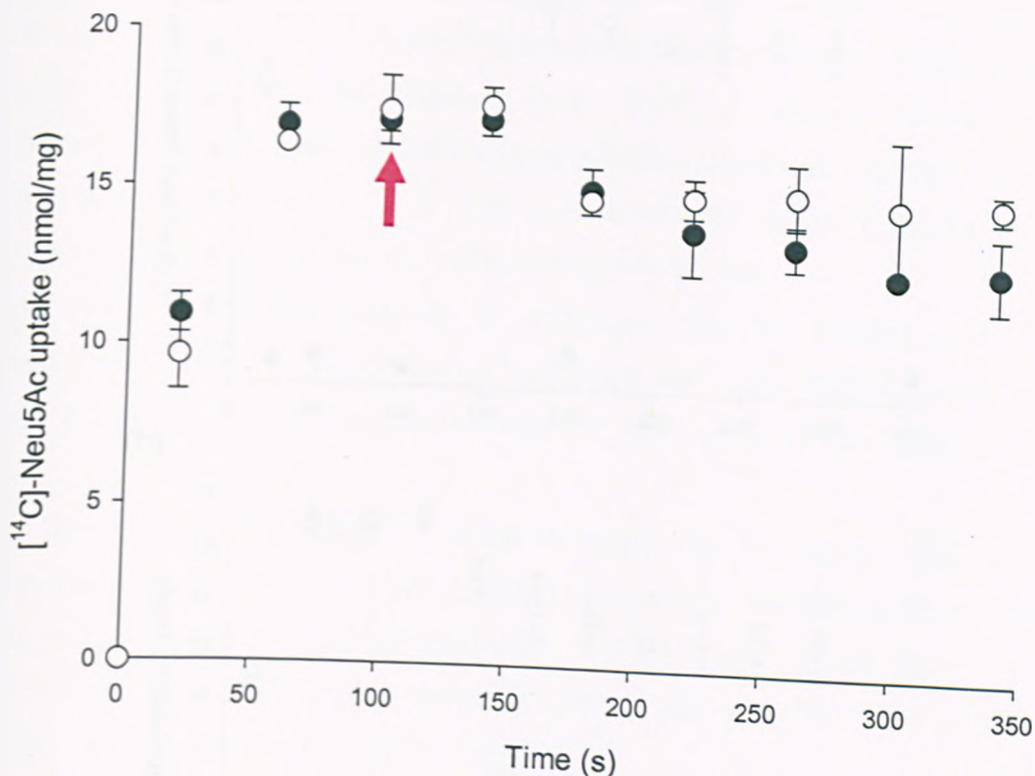


Figure 6.11 Effect on the accumulated levels of [¹⁴C]-Neu5Ac upon addition of excess cold Neu5Ac to assess the directionality of SiaQM. Uptake of [¹⁴C]-Neu5Ac into SiaQM-containing proteoliposomes in the presence of SiaP and an inwardly directed Na⁺ gradient plus (negative-inside) membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$) with the addition of either 1 mM Neu5Ac (open circles) or dH₂O (closed circles) at 100s. The $\Delta\mu_{\text{Na}^+} + \Delta\Psi$ was created by loading SiaQM-containing proteoliposomes with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, and diluting them into buffer containing 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, 2 μM valinomycin. All transport assays included 5 μM SiaP and 5 μM [¹⁴C]-Neu5Ac. Duplicate data sets were collected. Error bars shown indicate the range between the two sets of data ($n=2$). Red arrow indicates time at which excess cold substrate (or dH₂O) was added.

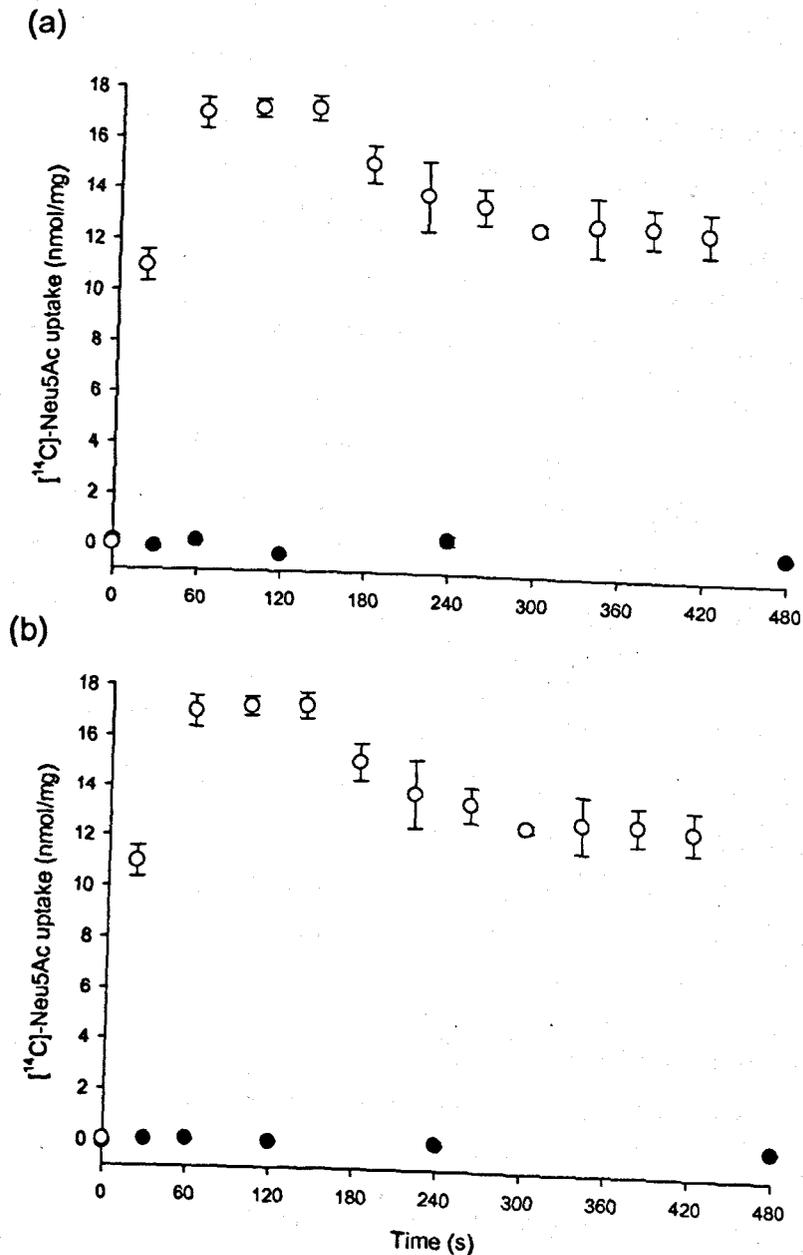


Figure 6.12 Counterflow assays in the presence and absence of SiaP. SiaQM-containing proteoliposomes loaded with 100 mM KAc, pH7, 20 mM KPi, pH7, 10 mM NaPIPES and 1 mM Neu5Ac were diluted into the same buffer devoid of Neu5Ac, but containing 5 μM [¹⁴C]-Neu5Ac and either (a) 5 μM SiaP or (b) no SiaP (closed circles). A control experiment was included in which SiaPQM was energised by $\Delta\mu\text{Na}^+ + \Delta\Psi$ (open circles). An empty liposome control was added and treated in the same way with and without SiaP. The values for the counterflow were calculated by subtracting the empty liposome control values. The average of duplicate data sets is presented. Error bars shown indicate the range between the two sets of data (n=2).

6.9 Effect of increasing concentrations of SiaP on *N*-acetylneuraminic acid transport by SiaPQM *in vitro*

SiaPQM is a secondary transporter, and the large subunit of the integral membrane components is a member of the IT (ion transport) superfamily and is thought to be related to classical secondary transporters that transport substrate with no need of an ESR (Prakash et al., 2003). It has been shown *in vivo* that deletion of either SiaP or SiaQM results in the inability to transport the Neu5Ac into the bacteria (Severi et al., 2005), however, the ability of SiaQM to transport Neu5Ac without SiaP has not been assessed *in vitro*.

An experiment was designed to determine whether SiaQM could transport Neu5Ac in the absence of SiaP. This experiment was then extended to assess the effects increasing the concentration of SiaP in the *in vitro* transport assay.

A standard *in vitro* transport assay (Methods section 2.16.1 and 2.16.2) was performed with a $\Delta\mu\text{Na}^+ + \Delta\Psi$ applied across the membrane (Methods section 2.16.3.1) except that the concentration of SiaP in the outside buffer was 0, 2.5, 5, 10 or 20 μM . An empty liposome control was also included which had the standard SiaP concentration of 5 μM .

Analysis of the transport activity in the presence of different concentrations of SiaP revealed that in the conditions tested, there was no transport activity in the absence of SiaP (Fig. 6.13, closed circles). In the presence of 5 μM SiaP, rapid uptake was observed up to ~ 11 nmol/mg, at which point the transport rate levels off (Fig. 6.13, closed triangles). In the presence of 2.5 μM SiaP, a slower transport rate was observed compared to transport in the presence of 5 μM SiaP and the value at the end of the assay was also lower at ~ 9 nmol/mg (Fig. 6.13, open circles). In the presence of 10 μM SiaP, the transport rate was lower and reached its peak at ~ 100 seconds at which point the uptake started to decrease, i.e. substrate was being effluxed from the proteoliposome (Fig. 6.13, open triangles). A similar transport activity was seen in the presence of 20 μM SiaP, except that the uptake peaked after only 60 seconds and there was markedly increased efflux from the proteoliposome resulting in a value of ~ 4 nmol/mg at the end of the assay (Fig. 6.13, closed squares).

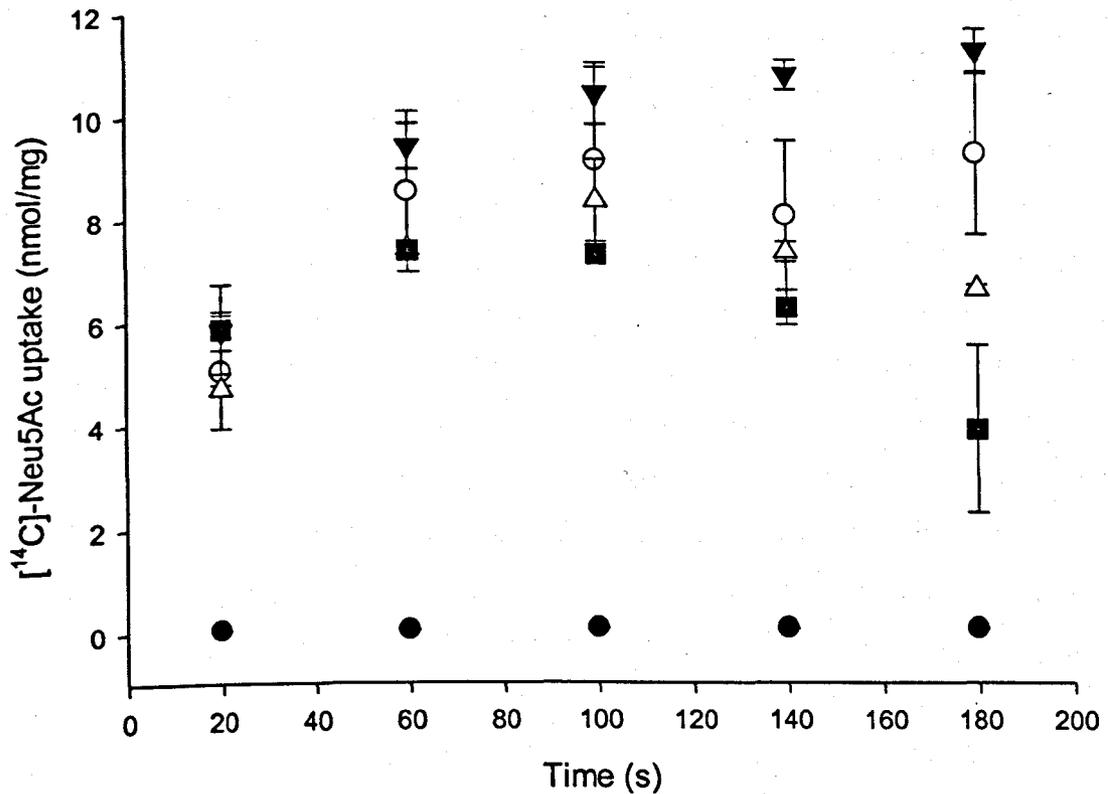


Figure 6.13 Transport activity of SiaQM in the presence of increasing concentrations of SiaP with a fixed [¹⁴C]-Neu5Ac concentration.

Uptake of [¹⁴C]-Neu5Ac into SiaQM-containing proteoliposomes in the presence of an inwardly directed Na⁺ gradient plus (negative-inside) membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$) and SiaP in concentrations of 0 μM (closed circles), 2.5 μM (open circles), 5 μM (closed triangles), 10 μM (open triangles) or 20 μM (closed squares). The $\Delta\mu_{\text{Na}^+} + \Delta\Psi$ was created by loading SiaQM-containing proteoliposomes with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, and diluting them into buffer containing 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, 2 μM valinomycin. All transport assays included 5 μM [¹⁴C]-Neu5Ac. Duplicate data sets were collected. Error bars shown indicate the range between the two sets of data (n=2).

The results of this experiment have shown that SiaQM cannot transport Neu5Ac in the absence of SiaP in the conditions used. It is possible that SiaQM is capable of transporting Neu5Ac in the absence of SiaP, but the affinity of SiaQM for Neu5Ac is very low. Perhaps if the concentration of Neu5Ac was increased then some transport activity would be observed.

Increasing the concentration of SiaP to 2.5, 5, 10 and 20 μM led to different transport activities. The results suggest that at 2.5 and 5 μM SiaP, Neu5Ac was transported into the proteoliposome and was stably retained within the lifetime of the assay. When the concentration of SiaP was increased to 10 and 20 μM , the uptake rate of Neu5Ac into the proteoliposomes was comparatively lower and a V_{max} was reached followed by efflux of Neu5Ac from the proteoliposome. This was SiaP concentration-dependent efflux as there was a more pronounced phenotype with the higher concentrations of SiaP. If this is indeed efflux of Neu5Ac from the proteoliposomes, then the obvious gate would be SiaQM as there is no evidence that Neu5Ac can permeate the lipid bilayer of the proteoliposome. This unexpected conclusion is contrary to the results of the counterflow assay and the cold chase experiment, which suggested that transport through SiaPQM was unidirectional.

6.10 Effects of excess SiaP chase on *N*-acetylneuraminic acid transport and accumulation *in vitro*

The previous experiment revealed that if increased amounts of SiaP were used in the *in vitro* transport assays a phenomenon occurred in which Neu5Ac was transported into the proteoliposomes, but then exited the proteoliposomes presumably via SiaQM. To determine whether this phenomenon is directly linked to increasing the concentration of SiaP, an experiment was designed in which a standard transport assay was performed with 2.5 μM SiaP and a $\Delta\mu\text{Na}^+$ + $\Delta\Psi$ applied across the membrane. At 180 seconds into the assay, excess SiaP (17.5 μM) was added and the effects on transport/efflux of Neu5Ac were monitored.

A standard *in vitro* transport assay (Methods section 2.16.1 and 2.16.2) was performed with a $\Delta\mu\text{Na}^+$ + $\Delta\Psi$ applied across the membrane (Methods section 2.16.3.1) except that the concentration of SiaP in the outside buffer was 2.5 μM .

Samples were taken at intervals and 17.5 μM SiaP was added at 180 seconds. A control was included where dH_2O was added instead of SiaP. Duplicate data sets were recorded for both experiments.

Upon application of the $\Delta\mu\text{Na}^+ + \Delta\Psi$ and addition of radiolabelled Neu5Ac, similar transport rates were observed for both transport assays, reaching between ~ 16 and 18 nmol/mg at 180 seconds (Fig. 6.14).

The Neu5Ac accumulation levels continue to increase and level off at ~ 21 nmol/mg when dH_2O was added, however, accumulated levels of Neu5Ac rapidly decreased upon addition of 17.5 μM SiaP, reaching ~ 5 nmol/mg by the end of the assay.

This experiment was designed to confirm whether it is the addition of excess SiaP that is causing the efflux of Neu5Ac from the proteoliposomes. The transport activity observed revealed that addition of excess SiaP led to efflux of Neu5Ac, whereas adding the same volume of dH_2O did not. This was an unexpected result considering the results of the counterflow assay and cold chase experiment, which demonstrated that transport through SiaPQM was unidirectional.

There are a number of possible reasons for this excess SiaP-dependent Neu5Ac efflux; for example, ligand-free SiaP could be binding to SiaQM and during the natural oscillations of the ligand-bound SiaQM translocation channel, the ligand binds to SiaP and is carried out of the proteoliposome. This is unlikely because there will be a large excess of ESR over translocation channel *in vivo* so if this was the reason for this phenomenon then Neu5Ac would never accumulate in the cell. A second scenario is that ligand-free SiaP is binding all Neu5Ac outside the proteoliposomes and forming a large pseudo-substrate gradient that is greater than the combined magnitude of the Na^+ gradient ($\Delta\mu\text{Na}^+$) and membrane potential ($\Delta\Psi$). It would be difficult to assess whether the first scenario is occurring, the second scenario, however could be examined with further experiments.

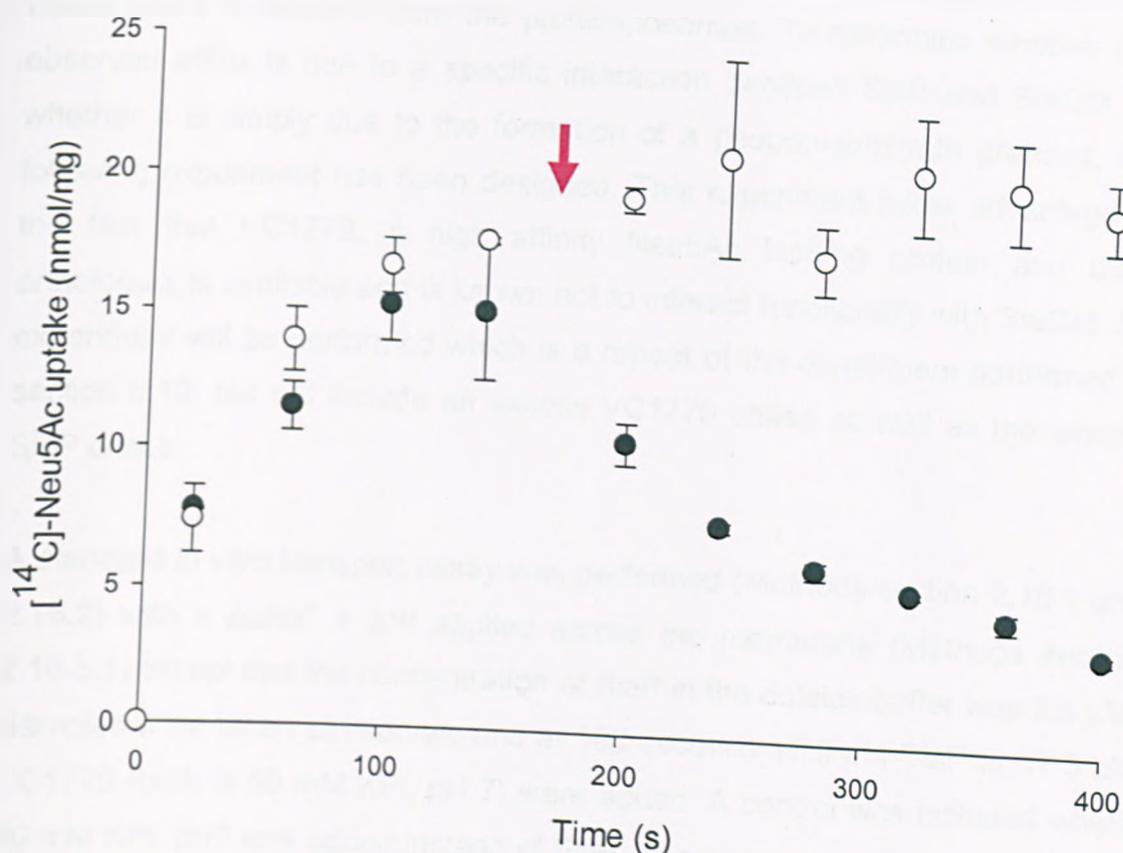


Figure 6.14 Effect of addition of excess SiaP on accumulated Neu5Ac. Accumulation of [¹⁴C]-Neu5Ac in SiaQM-containing proteoliposomes in the presence of 2.5 mM SiaP and an inwardly directed Na⁺ gradient plus (negative-inside) membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$) followed by the addition (red arrow) of 17.5 mM SiaP (closed circles) or dH₂O (open circles). The $\Delta\mu_{\text{Na}^+} + \Delta\Psi$ was created by loading SiaQM-containing proteoliposomes with 100 mM potassium acetate, pH 7, and 20 mM potassium phosphate, pH 7, and diluting them into buffer containing 100 mM sodium acetate, pH 7, 20 mM sodium PIPES, pH 7, 2 μM valinomycin. Both transport assays included 5 μM [¹⁴C]-Neu5Ac. The average of duplicate data sets are presented. Error bars shown indicate the range between the two sets of data ($n=2$). The red arrow indicates the time at which excess SiaP (or dH₂O) was added.

6.11 Comparison of effects of SiaP chase and VC1779 chase on *N*-acetylneuraminic acid transport and accumulation *in vitro*

The previous experiments have revealed that addition of excess SiaP can cause efflux of Neu5Ac from the proteoliposomes. To determine whether this observed efflux is due to a specific interaction between SiaP and SiaQM, or whether it is simply due to the formation of a pseudo-substrate gradient, the following experiment has been designed. This experiment takes advantage of the fact that VC1779, a high affinity Neu5Ac binding protein and SiaP orthologue, is available and is known not to interact functionally with SiaQM. An experiment will be performed which is a repeat of the experiment performed in section 6.10, but will include an excess VC1779 chase as well as the excess SiaP chase.

A standard *in vitro* transport assay was performed (Methods section 2.16.1 and 2.16.2) with a $\Delta\mu\text{Na}^+ + \Delta\Psi$ applied across the membrane (Methods section 2.16.3.1) except that the concentration of SiaP in the outside buffer was 2.5 μM . Samples were taken at intervals and at 100 seconds 17.5 μM SiaP or 17.5 μM VC1779 (both in 50 mM KPi, pH 7) were added. A control was included where 50 mM KPi, pH7 was added instead of ESR. The average of duplicate data sets for each experiment is presented.

Upon application of $\Delta\mu\text{Na}^+ + \Delta\Psi$ in the presence of 2.5 μM and 5 μM radiolabelled Neu5Ac, a high rate of uptake was observed reaching ~20 nmol/mg after 60 seconds (Fig. 6.15). Upon addition of 17.5 μM SiaP, a decrease in the accumulated levels of radiolabelled Neu5Ac reaching ~5.5 nmol/mg by the end of the assay (Fig. 6.15, open circles). However, upon addition of 17.5 μM VC1779 (Fig. 6.15, closed triangles) or 50 mM KPi, pH 7 (Fig. 6.15, closed circles), no such decrease in the accumulated levels of Neu5Ac was observed. A general downward trend is observed in the VC1779 chase data, but this is also observed in the control data where 50 mM KPi, pH 7, was added, indicating that this is not a consequence of adding VC1779.

This experiment and the previous experiment in section 6.10 have shown that efflux of radiolabelled Neu5Ac from the proteoliposomes after addition of 17.5

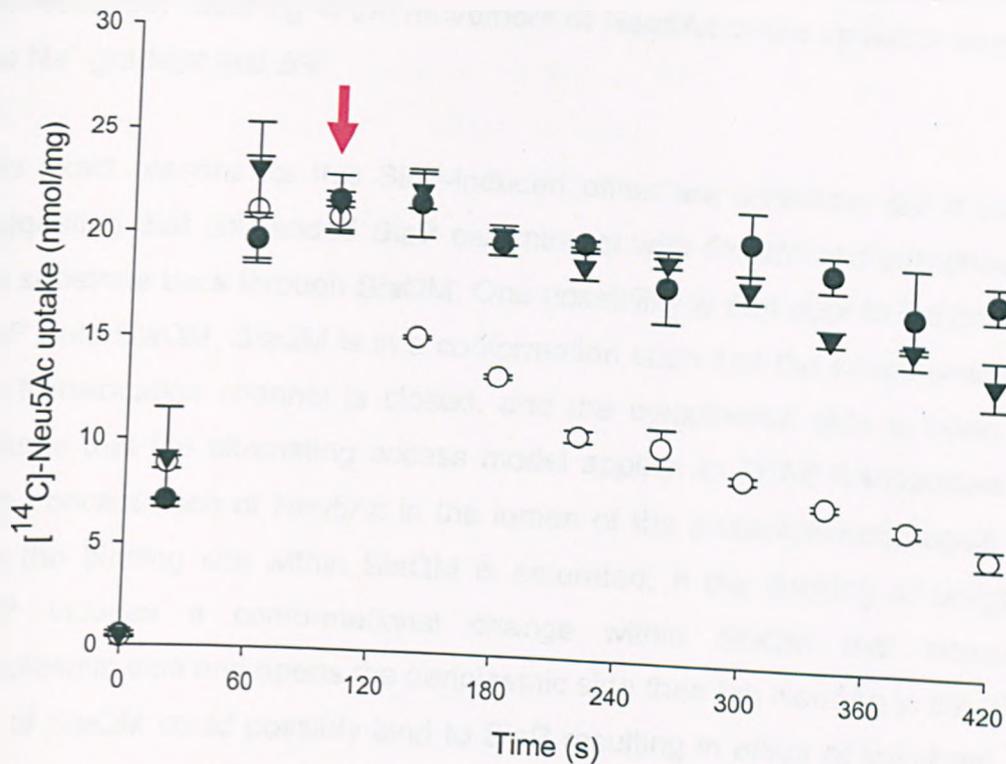


Figure 6.15 Effect of addition of excess SiaP and excess VC1779 on accumulated Neu5Ac. Accumulation of [^{14}C]-Neu5Ac in SiaQM-containing proteoliposomes in the presence of $2.5\ \mu\text{M}$ SiaP and an inwardly directed Na^+ gradient plus (negative-inside) membrane potential ($\Delta\mu_{\text{Na}^+} + \Delta\Psi$) followed by the addition of $17.5\ \mu\text{M}$ SiaP (open circles), $17.5\ \mu\text{M}$ VC1779 (closed triangles) or $50\ \text{mM}$ KPi, pH 7 (closed circles) at 100 seconds. The $\Delta\mu_{\text{Na}^+} + \Delta\Psi$ was created by loading SiaQM-containing proteoliposomes with $100\ \text{mM}$ potassium acetate, pH 7, and $20\ \text{mM}$ potassium phosphate, pH 7, and diluting them into buffer containing $100\ \text{mM}$ sodium acetate, pH 7, $20\ \text{mM}$ sodium PIPES, pH 7, $2\ \mu\text{M}$ valinomycin. $5\ \mu\text{M}$ [^{14}C]-Neu5Ac was used in each transport assay. The average of duplicate data sets are presented. Error bars shown indicate the range between the two sets of data ($n=2$). The red arrow indicates the time at which excess ESR (or $50\ \text{mM}$ KPi, pH 7) was added.

μM SiaP is a consequence of specific interaction between SiaP and SiaQM. Efflux was not induced by the addition of $17.5 \mu\text{M}$ VC1779, a SiaP orthologue that has been shown to be Neu5Ac with a similar affinity to SiaP. This is an unexpected result as it suggests that excess SiaP can make SiaQM work bidirectionally resulting in the movement of Neu5Ac in the opposite direction to the Na^+ gradient and $\Delta\Psi$.

The exact reasons for this SiaP-induced efflux are unknown, but it could be suggesting that unliganded SiaP can interact with SiaQM and somehow "pull" the substrate back through SiaQM. One possibility is that prior to the docking of SiaP onto SiaQM, SiaQM is in a conformation such that the periplasmic side of the translocation channel is closed, and the cytoplasmic side is open (if we assume that the alternating access model applies to TRAP transporters). The high concentration of Neu5Ac in the lumen of the proteoliposome could mean that the binding site within SiaQM is saturated. If the docking of unliganded SiaP induces a conformational change within SiaQM that closes the cytoplasmic side and opens the periplasmic side then the Neu5Ac in the binding site of SiaQM could possibly bind to SiaP resulting in efflux of substrate. This, however, is just conjecture and a number of experiments will have to be performed to determine the reasons for this observation.

6.12 Chapter summary

Using SiaQM that was purified as documented in Chapter 5, the whole Neu5Ac-specific TRAP transporter was reconstituted into liposomes. This chapter documents the first demonstration of TRAP transporter activity in an *in vitro* reconstituted system. Using this reconstituted system, the energetics of Neu5Ac transport by SiaPQM were probed. This has revealed that SiaPQM utilises a $\Delta\mu\text{Na}^+$ to power transport of Neu5Ac. This was corroborated by *in vivo* Neu5Ac transport assays in whole cells expressing SiaPQM.

Transport assays on the reconstituted SiaPQM system revealed that it can not use a ΔpH or the $\Delta\Psi$ to transport Neu5Ac. A $\Delta\mu\text{Na}^+$ in combination with a ΔpH was found to decrease the rate of transport, whereas the application of a $\Delta\mu\text{Na}^+$ in combination with a $\Delta\Psi$ increased the transport rate above that of $\Delta\mu\text{Na}^+$ alone. It was demonstrated that it was the Na^+ gradient and not simply the presence of Na^+ that drove transport of Neu5Ac. It was also demonstrated that

the interaction of the ESR, SiaP with SiaQM is specific as the close orthologue of SiaP, VC1779 could not form a functional transporter with SiaQM.

TRAP transporters are secondary transporters, however, it was shown that SiaPQM does not demonstrate characteristic secondary transporter behaviour. SiaPQM did not show any solute counterflow activity or exchange activity in a "cold chase" experiment. Contrary to these findings, the addition of excess unliganded SiaP resulted in efflux of Neu5Ac from the proteoliposomes, demonstrating that, under these non-physiological conditions, SiaQM could work reversibly. This was shown to be a result of a specific interaction between SiaP and SiaQM as VC1779 could not induce this reversible characteristic.

Chapter 7

Discussion

7.1 Distribution and diversity of TRAP transporters

As part of this body of work, the relational database TRAPdb was populated with data describing the characteristics of thousands of TRAP transporter genes. Collection of this data has clearly shown the extent of TRAP transporter distribution in prokaryotes, with 243 different organisms containing at least one component of a TRAP transporter. This data was correct as of February 2008 and has no doubt increased greatly since then. The last analysis of this kind was performed in 2001 (Kelly & Thomas, 2001) which identified approximately 10% of the total TRAP transporter genes that can now be found in TRAPdb.

During this data-mining exercise a number of novel organisations of TRAP genes were found that extend the already variable repertoire of TRAP transporter genetic organisation (Chapter 3). This analysis also identified a number of motifs, such as a serine/threonine kinase and a helix-turn-helix DNA binding domain, fused to TRAP transporter ESR genes, potentially rendering them able to perform non-transport related functions (discussed in Chapter 3). Although these findings clearly need to be experimentally verified, they extend the potential uses for TRAP transporter components

The most interesting finding from this data-mining was the high numbers of TRAP transporters found in bacteria whose environment contains high Na^+ concentrations (Chapter 3). A list of organisms with the highest numbers of TRAP transporters was compiled from the data collected. This analysis revealed that some organisms had up to 28 separate TRAP transporters encoded on their genomes, which in some cases accounted for ~10% of the total number of transporters. Further analysis revealed that a large proportion of organisms with a high number of TRAP transporters were marine-dwelling bacteria (Chapter 3). It has been shown that the bioenergetics of most marine bacteria depends on a Na^+ circuit as opposed to a H^+ circuit found in bacteria such as *E. coli* (Kogure, 1998). The sodium motive force (smf) is known to be used for a number of different processes such as transport and flagellar movement in a number of different bacteria (Hirota & Imae, 1983, Lanyi, 1979). This correlation alone is not sufficient to declare that all TRAP transporters utilise a Na^+ gradient, but it does add to a growing litany of research linking Na^+ to TRAP transporters. The first unequivocal demonstration that Na^+ is required for the transport mechanism by TRAP transporters is shown in this thesis by the transport of Neu5Ac by SiaPQM (Chapter 6).

The most powerful and useful application of the data collected for TRAPdb is in the prediction of substrate specificity. It is assumed that TRAP transporter ESRs, like ABC transporter ESRs, confer the substrate specificity to the transporter (Doeven *et al.*, 2004, Wilkinson, 2003). ESRs can be grouped by their sequence similarity and it has been demonstrated that groups of ESRs bind the same or similar compounds (Tam & Saier, 1993). This is powerful tool because if the substrate specificity of one protein in the group is known then the specificity of the rest of the group can be inferred (but would obviously need experimental verification). Another method that can be used to predict the substrate specificity of a TRAP transporter is to analyse its genome context. If the transporter is surrounded by genes that encode enzymes that catabolise a particular compound then it is likely that that particular compound is transported by the encoded TRAP system. This method has been demonstrated for a number of TRAP transporters including SiaPQM and YiaMNO (Severi *et al.*, 2005, Thomas *et al.*, 2006). Analysis of the data from TRAPdb revealed a TRAP transporter from *Sinorhizobium meliloti* whose large membrane component (DctM) is fused to a malonyl-CoA decarboxylase enzyme. Closer inspection of the genome context revealed that it is also in an operon with a gene encoding malonyl-CoA synthetase. Analysis of the genomes of closely related organisms revealed a similar linkage of a TRAP transporter to these two enzymes (although they were not fused), and one example of a TRAP transporter linked to a multisubunit malonyl-CoA decarboxylase, which is structurally different, but performs the same function. This strongly suggests that this TRAP transporter (and orthologues in related species) is specific for malonate. This was reinforced by a study that demonstrated malonate could induce the expression of the ESR of this system (Mauchline *et al.*, 2006). To confirm this specificity, the ESR could be purified and characterised or *in vivo* transport assays could be performed. It would be especially interesting to determine whether the TRAP transporter membrane component is actually fused to the malonyl-CoA decarboxylase enzyme.

There are now a number of predicted and experimentally defined substrates for TRAP transporters (see Introduction section 1.8.4). All of the experimentally defined substrates share a common feature; they all possess at least one carboxylate group. Crystal structures of TRAP transporter ESRs have shown

that a carboxylate group of the ligand is coordinated by an arginine residue (Arg147 in SiaP) in the binding site of the protein (Gonin *et al.*, 2007, Johnston *et al.*, 2008, Muller *et al.*, 2006, Rucktooa *et al.*, 2007). Phylogenetic analysis revealed that this arginine is conserved in 98% of TRAP ESRs (from 248 sequences), implying that this interaction is very important in the binding mechanism of TRAP transporters (Muller *et al.*, 2006). This also correlates well with the fact that all experimentally defined TRAP transporter substrates contain at least one carboxylate group. There are, however, a number of predicted substrates of TRAP transporters that do not possess a carboxylate group. For example, taurine does not possess a carboxylate group and is predicted to be a TRAP transporter substrate from genome context analysis (Bruggemann *et al.*, 2004, Denger *et al.*, 2006). Taurine contains a sulphonate group, which has chemical properties similar to carboxylate so it is likely that it could be a TRAP transporter substrate. Other predicted TRAP transporter substrates are not as straight forward. A number of sugars, including raffinose, lactose and arabinose, are predicted to be TRAP transporter substrates. These compounds were shown to induce the transcription of a TRAP transporter ESR from *S. meliloti* and are therefore predicted to be the substrates (Mauchline *et al.*, 2006). These sugars do not possess any acidic functional groups so it is unknown whether they are truly TRAP transporter substrates. ABC transporters are able to bind and transport a number of compounds with different chemical properties, including sugars and amino acids, so there should be no limitations *per se* in what TRAP transporters can transport. Experimental verification will be required to determine whether these are true TRAP transporters substrates.

7.2 Expression and purification of TRAP integral membrane proteins

In order to study the biochemical and energetic characteristics of TRAP transporters in more detail, the protein components need to be overexpressed and purified. Production of large quantities of ESR proteins is now fairly trivial and a number of different TRAP ESRs have been expressed and purified to homogeneity (Gonin *et al.*, 2007, Johnston *et al.*, 2008, Muller *et al.*, 2006, Rucktooa *et al.*, 2007, Thomas *et al.*, 2006). It is the production of the integral membrane components that has proved most challenging. During the course of this investigation, 3 integral membrane proteins have been expressed; YiaM (4 TMHs), YiaN (12 TMHs), both from *E. coli* and SiaQM (17

TMHs) from *H. influenzae*. Three different expression vectors have been used; pTTQ18 (C-terminal RGS_{H6} tag), two pBAD vectors (either N- or C-terminal H₁₀ tag) and two pNZ8048 derivatives (either N- or C-terminal H₁₀ tag). Four *E. coli* expression strains have been utilised (BL21 (DE3), C41 (DE3), C43 (DE3) and MC1061) and one *L. lactis* strain (NZ9000). Varying factors such as the expression vector and strain and optimisations of expression conditions has resulted in the overexpression of all three proteins. Unfortunately, the sample size and number of variables changed are not large enough to make general statements about the relative merits of each expression system in the overexpression of integral membrane proteins. However, a number of the parameters that contributed to the successful expression of these TRAP transporter membrane components have also been observed to contribute to the successful expression of other integral membrane proteins in the literature. As a preface to this section of the discussion, after reviewing the literature on the subject, it is apparent that expression of integral membrane proteins is overall very subjective.

All three proteins, YiaM, YiaN and SiaQM were expressed using the pTTQ18 vector, which has previously been used to overexpress a number of integral membrane proteins (Morrison *et al.*, 2003, Saidijam *et al.*, 2003, Saidijam *et al.*, 2005, Ward, 2000). Expression from this vector is tightly regulated and under the control of the moderately-strong *lac* promoter. A protein produced from this vector possesses a C-terminal RGS_{H6} tag. Expression of YiaM from this vector resulted in a yield of ~0.7 mg/L culture, whereas expression of SiaQM in the pTTQ18 vector resulted in a yield of ~0.12 mg/L culture. Unfortunately, YiaN was not purified to homogeneity, therefore a yield can not be calculated. A number of conditions were optimised for the expression of each protein including inducer concentration, temperature of growth, post-induction incubation time and the medium of growth (Chapter 5 and 6). The importance of optimising culture conditions is emphasised by studies showing that optimisation can lead to a 3-5 fold increase in protein production (Li *et al.*, 2001, Auer *et al.*, 2001).

In order to increase the overall yield of YiaN, the expression of YiaN was compared in three different strains of *E. coli*, BL21 (DE3), C41 (DE3) and C43 (DE3). C41 and C43 are derivatives of BL21 that were derived by selecting for lower susceptibility to the toxic effects of integral membrane protein expression

(Miroux & Walker, 1996). Comparing steady state levels of YiaN using Western blotting revealed that BL21 (DE3) accumulates the most YiaN after 2 hours, the signal then decreased in subsequent timepoints. YiaN accumulation in C43 (DE3) peaked after 1 hour, but was lower than that observed for BL21 (DE3). Accumulation in C41 (DE3) peaked after 1 hour, and attained similar accumulation levels to BL21 (DE3). There appeared to be higher stability of YiaN in C41 (DE3) as there was no decrease in signal during the timecourse, as seen for BL21 (DE3). A study was performed comparing expression levels, plasmid stability and toxicity from a number of different vectors in BL21, C41 and C43 (Dumon-Seignovert *et al.*, 2004). This study revealed that C41 and C43 had higher plasmid stability, lower toxicity and higher percentage of expression (percentages of expressing plasmids were 54%, 86% and 81% for BL21, C41 and C43, respectively), however, the relative levels of expression were not reported (Dumon-Seignovert *et al.*, 2004). Another study, this time working with only one protein, MexY, detected no accumulation of MexY in BL21 (DE3). However, upon switching to C43 (DE3) a yield of up to 5 mg/L culture was obtained (Mokhonova *et al.*, 2005). In a high throughput study of 49 different integral membrane proteins, comparative expression in BL21, C41 and C43 was performed (Eshaghi *et al.*, 2005). Out of the 49 targets, 51% were expressed from BL21, 49% from C41 and 45% from C43. There were very few differences in which targets were expressed by each strain, however, there were several targets where the relative expression levels were quite different for each strain. These studies and the findings from Chapter 4 of this thesis suggest that BL21, C41 and C43 will all express the same proteins, with some exceptions. However, the relative levels of expression may be quite different. Therefore, it would be beneficial to screen all 3 strains for expression of the target membrane protein.

SiaQM was the major target for study and this *H. influenzae* protein was heterologously expressed from pTTQ18 and pBAD vectors in *E. coli*. The protein produced from the pTTQ18 vector had a C-terminal RGSH₆ tag, whereas SiaQM was produced with either an N- or C-terminal H₁₀ tag from the pBAD vectors. The variable that made the biggest difference in expression levels was the location of the affinity tag. SiaQM expressed from pTTQ18 had an overall yield of 0.12mg/L, whereas, N-terminally tagged SiaQM from the pBAD vector had an overall yield of 1.5 mg/L. The C-terminally tagged version

of SiaQM expressed from a pBAD vector was never purified so an overall yield was not calculated. However, expression trials and optimisation suggest that it was significantly lower than the yield of N-terminally tagged SiaQM expressed from pBAD and was more in the range of the yield of SiaQM from pTTQ18 (Chapter 6). A study on how accumulation of aquaporin was affected by histidine tag length and location has been performed that corroborates this finding (Mohanty & Wiener, 2004). Aquaporin from *E. coli* was expressed by a pET vector (T7 promoter) with either a hexa- or deca-histidine tag at the N- or C- terminus. This study revealed that for this protein, the longer the affinity tag the lower the accumulation, and having the affinity tag at the N-terminus results in higher accumulation. A much wider study on the expression optimisation of 36 P-type ATPases from 11 different bacteria also revealed that positioning of the affinity tag at the N-terminus resulted in higher accumulation of protein (Lewinson *et al.*, 2008). This study was performed in both pET vectors and pBAD vectors with the same overall result in each. As mentioned previously, membrane protein expression is very subjective and this is reflected in the fact that approximately 1/3 of the P-type ATPases in this study accumulated to a higher degree with a C-terminal affinity tag. This higher accumulation of the N-terminally tagged variant compared to the C-terminally tagged variant was also observed for SiaQM expressed in *L. lactis* and for SiaP expressed from the pBAD vectors in *E. coli* (Chapter 6). The reason for the different levels of accumulation of integral membrane proteins according to the position of the affinity tag is not known, but these studies clearly show that it is a variable that requires significant consideration.

A study that directly compared the heterologous expression of 37 targets from 3 different bacteria from pTTQ18 and pBAD in *E. coli* showed that pTTQ18 was able to express the highest number of proteins (68% compared to 65% from pBAD), however, there was higher accumulation of protein from the pBAD vector (Surade *et al.*, 2006). The study concludes that there is no single expression vector that is best for every protein, therefore using multiple vectors (even just two) will undoubtedly increase success rate of membrane protein expression.

As well as expression in *E. coli*, SiaQM was also expressed in *L. lactis* with a C- or N-terminal decahistidine tag (Chapter 6). Expression of SiaQM from the nisin

inducible pNZ8048 vector in *L. lactis* resulted in a higher overall yield of protein than expression from pBAD in *E. coli*. This difference in accumulation was quite small with SiaQM making up 2-3% of the total membrane protein in each strain. The previously mentioned study that compared the accumulation of 37 targets from 3 different bacteria expressed from different vectors also compared the accumulation of these targets in *E. coli* from the pTTQ18 and pBAD vectors with expression in *L. lactis* from the pNZ8048 expression vector (Surade et al., 2006). The study showed that 84% of targets were successfully expressed in *E. coli* (from at least one of three vectors) compared to just 40% in *L. lactis* (utilising only one expression vector). It was shown that each individual *E. coli* expression vector used in this study (pTTQ18, pBAD and pQE) outperformed the *L. lactis* expression vector (pNZ8048). It was also revealed that there were no targets expressed by *L. lactis* that were not expressed by at least one of the *E. coli* vectors with approximately the same level of accumulation (Surade et al., 2006). This may suggest that expression in *L. lactis* is not worthwhile; however, as is apparent from all these studies, expression of membrane proteins is very subjective so varying the expression species is an excellent idea if expression in *E. coli* is not ideal. As an example of this, *L. lactis* has been shown to be effective in the overexpression of a number of eukaryotic membrane proteins (Kunji et al., 2003, Monne et al., 2005, Kunji et al., 2005).

7.3 Basic structural requirements of TRAP transporters

7.3.1 ESR component

It has been demonstrated unequivocally in two separate TRAP systems, DctPQM and SiaPQM, that the membrane component (whether fused or encoded by two separate genes) and the ESR component are essential for transport activity *in vivo* (Forward et al., 1997, Severi et al., 2005). This dependence on the presence of SiaP has also been demonstrated *in vitro* in the purified, reconstituted system (Chapter 6). In the presence of SiaP, SiaQM-containing proteoliposomes and a $\Delta\mu\text{Na}^+ + \Delta\Psi$, transport of Neu5Ac into the lumen of the proteoliposomes is observed. However, under the same conditions, with the same batch of proteoliposomes, but in the absence of SiaP no transport activity was observed (Chapter 6). This phenotype is also observed in ABC transporters. ABC transporters were first differentiated from secondary

transporters by their osmotic shock sensitivity, a process in which the periplasm, and all periplasmic components, is removed (Neu & Heppel, 1965). ABC transporters require an ESR for transport and are therefore osmotic shock sensitive. Secondary transporters that require no periplasmic components are osmotic shock insensitive. TRAP transporters, however, are not ABC transporters, and it has been postulated that the likely translocation domain, DctM was once a standalone secondary transporter that recruited the accessory proteins, the small integral membrane component, DctQ and an ESR (Rabus *et al.*, 1999). It is therefore not beyond the realms of possibility that the integral membrane component can still work independently of the ESR component, but with a very low affinity for substrate. In the *in vitro* transport assay where SiaP was not present, 5 μ M Neu5Ac was added to the transport mixture, and it is possible that this concentration may be too low to observe any transport. A simple experiment to determine whether this were true would be to increase the concentration of Neu5Ac and monitor transport activity in the absence of ESR.

A potential function for the small membrane component in TRAP transporters is the ability to recruit the ESR domain. It may also function as a clamp to prevent backflow through the large membrane component, which is why the counterflow and the substrate exchange phenomena are not observed (Chapter 6). If the DctM component was purified in the absence of the DctQ component, perhaps bidirectional and ESR-independent transport could be observed.

Due to the scarcity of structural information on TRAP transporters, derived either biochemically or biophysically, a number of vital characteristics are unknown. One of these characteristics is the position at which the ESR interacts with the membrane components. In the high resolution structure of SiaP, a potential site of interaction is observed that is conserved in members of the sialic acid clade, but not outside of it (Muller *et al.*, 2006). It was proposed that this surface exposed conserved patch interacted with the integral membrane components and would allow promiscuity of the ESRs from these transporters between the transporters in the sialic acid-specific grouping. *In vitro* transport assays in the presence of SiaQM-containing proteoliposomes and VC1779, the SiaP orthologue from *V. cholerae*, showed no transport activity, thus disproving this theory (Chapter 6).

The positions at which SiaP and SiaQM interact with each other remain unknown. This interaction has been shown to be specific, therefore it is likely that the sites of interaction on each protein have evolved synchronously. The question of where SiaP and SiaQM interact could be answered using a number of different experiments. The complex of the integral membrane component and the ESR could be crystallised. This is not a trivial feat as integral membrane proteins are particularly difficult to crystallise and a number of reviews have been published on the subject (Wiener, 2004). One of the major problems is the highly dynamic nature of transporters, especially secondary transporters. The elucidation of the high resolution crystal structure of lactose permease was only possible at first by using a conformationally constrained mutant (Abramson *et al.*, 2003). However, recent advances have allowed the wild type structure to be crystallised (Guan *et al.*, 2007). The third major limitation in the elucidation of a high resolution crystal structure of a complete TRAP transporter complex is stably trapping the ESR onto the membrane components. This was only possible for the maltose ABC transporter structure by using a mutation in the ATPase domain that could bind ATP, but could not hydrolyse it (Oldham *et al.*, 2007). Unfortunately, no inhibitors that can trap a transition state in the transport cycle of TRAP transporters are currently available. One route to circumvent this would be to use chemical crosslinking to covalently attach the membrane component to the ESR. However, one would have to question the physiological relevance of a complex trapped in this way.

The interaction between the ESR and membrane components could be mapped using specific chemical crosslinking. A cysteine-free SiaQM could be constructed and monitored for activity (SiaP is naturally cysteine-free). A number of single cysteine mutations in each protein component could be introduced onto the surface of SiaP proposed to interaction with SiaQM and the periplasmic loops of SiaQM predicted from hydropathy analysis. This would allow the mapping of the site of interaction as the cysteine residues would have to be very close to form a disulphide bond. A more accurate topology model for SiaQM would be required first, which could be produced using cysteine scanning mutagenesis and hydrophilic thiol-reactive agents, as seen for lactose permease (Frillingos *et al.*, 1998), or PhoA/ β -lactamase fusions, as used with DctQ and MalFG (Boyd *et al.*, 1993, Ehrmann *et al.*, 1990, Wyborn *et al.*, 2001).

An issue that has not been addressed for TRAP transporters is whether the unliganded ESR interacts with the integral membrane components or whether it is only after ligand binding that the ESR and membrane components interact. This is a contentious issue in ABC transporters as it has been demonstrated that it is only the ligand bound form that interacts with the integral membrane components in at least one well characterised ABC transporter, OppABCDF (Doeven *et al.*, 2008). In this case, interaction was only observed between liganded ESR OppA and the membrane components using fluorescence cross-correlation spectroscopy (FCCS) Ligand-free OppA was shown not to interact (Doeven *et al.*, 2008). *In vitro* transport studies on the vitamin B₁₂ ABC transporter suggested that in the absence of ATP radiolabelled substrate was trapped in between the ESR and membrane components. This trapped radiolabelled substrate could be slowly displaced with unlabelled substrate leading to the proposition that the ESR is in constant contact with the membrane components throughout the transport cycle (Borths *et al.*, 2005). It has also been shown that the ESR and membrane components of the BtuCDF system form a very stable complex *in vitro* (Borths *et al.*, 2002). There is also evidence that the unliganded maltose-specific ESR MalE from *E. coli* and histidine specific ESR HisJ from *S. typhimurium* interact with their cognate membrane components with a similar affinity to when they have ligand bound (Ames *et al.*, 1996, Austermuhle *et al.*, 2004). This was demonstrated for the maltose system using site directed spin labelling followed by electron paramagnetic resonance (EPR) and for the histidine ABC system using chemical crosslinking.

It is unknown at this time which of these mechanisms TRAP transporters employ. However, this could be addressed using a number of different methods such as those mentioned above for the maltose and histidine ABC systems as well as ITC or surface plasmon resonance (SPR). Both of the latter methods can be used to detect protein:protein interactions and could be used to determine whether unliganded or liganded ESR can interact with detergent solubilised membrane components.

7.3.2 DctQ and DctM components

Due to the aforementioned problems with integral membrane protein expression and purification, the DctQ and DctM (and fusions of the two) are by far the least

studied components of TRAP transporters.

There are many fundamental questions that remain to be answered regarding these proteins and now with the high yield production of a number of these proteins, we are able to study them in more detail. One aspect that remains to be answered is the topology of the large integral membrane subunit, DctM, and fused integral membrane proteins, DctQM. Another fundamental question that has not been addressed is the subunit stoichiometry required to form a functional transporter.

Expression and purification of the hexahistidine-tagged DctQ subunit, YiaM, from the *E. coli* TRAP transporter revealed a ladder of immunoreactive species. The species with apparent molecular weights corresponding to a monomer and dimer of YiaM were confirmed to be YiaM using peptide mass fingerprinting (Chapter 4). This ladder of immunoreactive species is composed of discrete bands that have apparent molecular weights that approximately correlate with the consecutive addition of YiaM protomers. The highest oligomeric structure observed by Western blotting corresponds approximately with a complex of five YiaM subunits. There is reason to believe that all of these species are YiaM complexes because the monomer and dimer were confirmed by peptide mass fingerprinting and the higher oligomeric species are sensitive to boiling and treatment with the reducing agent, β -mercaptoethanol. The actual oligomeric state of YiaM in solution is not known and it is also unknown whether any of these apparent complexes are physiologically relevant.

There are now known to be many examples of a fusion between the DctQ and DctM subunits. Unless there are unforeseen structural and mechanistic differences between TRAP systems composed of two separate membrane components or one a single membrane component, this would suggest that the DctQ and DctM subunits in TRAP transporters are in a 1:1 stoichiometry with each other. This implies that if the YiaM (DctQ subunit) is forming oligomeric structures then its cognate DctM subunit, YiaN is forming equivalent oligomeric structures. However, when hexahistidine-tagged YiaN was co-expressed with untagged YiaM, two immunoreactive species were observed; one at the size predicted for a monomer and one at the size approximately predicted for a dimer. Unfortunately, the purification was not very efficacious and the elution fraction contained a number of contaminants so it could not be determined whether the potential dimeric YiaN species was actually YiaN. A protein

corresponding to YiaM was not identified in this elution fraction, however, it may have been masked by the high abundance of contaminants. It would be interesting to discover whether YiaM is still accumulated when co-expressed with YiaN and whether YiaM and YiaN form a complex in the membrane. Answering these fundamental questions would help to determine the function of YiaM.

The stoichiometry of the integral membrane components in a functional transporter is still unknown. Classical secondary transporter integral membrane components (MFS members) are invariably composed of ~12-14 TMHs (Pao *et al.*, 1998) and ABC transporter membrane components have on average ~12 TMHs (a dimer of proteins with 6 TMHs each (Davidson *et al.*, 2008)). The average TRAP DctM subunit (or DctM domain in fused membrane components), which is considered to form the translocation channel, consists of a predicted ~12 TMHs (Rabus *et al.*, 1999). Unless there are huge differences between TRAP transporters and these other families of transporters then one would predict that 12 TMHs in the DctM domain is sufficient for a single transporter. This reasoning would suggest that the stoichiometry of the DctM component in a single TRAP transporter is 1. If a model had to be made from the current body of evidence, one DctQ and one DctM would be predicted for a single TRAP transporter. However, further work is required to confirm this.

There are several potential functions suggested for the DctQ component in the literature. These include; a role in mediating the interaction between the DctP and DctM components, as a chaperone to maintain the stability of DctM or in the assembly of DctM in the membrane. It is also speculated that DctQ may be involved in electrochemical energy coupling to substrate transport (Kelly & Thomas, 2001). The first of these possibilities was initially corroborated by the identification of a DctP-DctQ fusion. However, this is the only example of this type of fusion, despite the comprehensive data mining required for the compilation of TRAPdb and is therefore considered to be a sequencing error. The possibility of DctQ acting as an assembly factor or as a chaperone is reinforced slightly by the observations that YiaN from *E. coli* and DctM from *R. capsulatus* are accumulated to a much higher degree in the presence of their cognate DctQ subunits, YiaM and DctQ, respectively (Chapter 4 and (Kelly & Thomas, 2001)). However, this apparent instability of the DctM component in

the absence of the DctQ subunit may simply be due to DctQ interacting with DctM and burying protease cleavage sites or ordering disordered polypeptide that is prone to degradation, and not that it is directly involved in insertion or assembly of DctM.

There is currently no evidence to suggest that DctQ is involved in coupling the electrochemical energy to solute transport. There is strong evidence in the maltose ABC transporter system that the ESR, MalE, interacting with the periplasmic side of the integral membrane proteins, MalFG, induces a conformational change that results in the stimulation of ATPase activity (Davidson *et al.*, 1992). The model for maltose transport by this ABC transporter suggests that interaction of the liganded MalE with MalFG induces a conformational change which opens the periplasmic side of the transmembrane channel, opening the ESR and releasing the ligand into the membrane component (Oldham *et al.*, 2007). It is possible that in TRAP transporters the DctQ subunit is responsible for inducing this conformational change upon interaction with liganded DctP subunit. Whilst phylogenetic analysis indicates the DctM component was likely a standalone secondary transporter in its evolutionary past. DctQ homologues are not found associated with any other protein and may have been recruited to facilitate the use of an ESR component with the translocation machinery.

Much work will be required to define the role of DctQ in the transport cycle. A number of approaches for this would be possible. Structural studies are an obvious avenue of investigation and the major bottleneck with membrane protein crystallography is producing enough protein for the screening process. This is not a problem for YiaM or SiaQM, as expression systems have been developed produce a high yield of protein (Chapter 4 and 5). Another approach is to determine how much of the DctQ subunit is required for transport by truncating the protein. It would also be interesting to determine whether the DctM component alone is sufficient for transport, which could be assessed using *in vitro* transport assays.

The stoichiometry of the integral membrane components could be assessed using chemical crosslinking. This is not an ideal solution as there is no way to prove whether the oligomeric state defined is relevant in solution. A more accurate method would be to use analytical ultracentrifugation (AUC). However, this method is not without its problems. AUC can be used to calculate the

molecular weight and overall size of the protein. The problem with integral membrane proteins is that the detergent micelle is required to keep the protein soluble. The micelle will increase the size of the particle massively and unless the contribution of the detergent can be determined there is no way to measure the size of the protein alone. To determine the detergent contribution, radiolabelled detergent can be used as for AmtB from *E. coli* (Blakey *et al.*, 2002). One of the most commonly used detergents, DDM (and DM for that matter) have a maltoside headgroup and there are colorimetric assays available to determine the concentration of detergent in a given sample (Lau & Bowie, 1997). This information can be used to determine the contribution of detergent and thus subtracting it from the total mass, therefore gleaning what the actual mass of the protein is. The ideal solution to this problem is to use neutrally buoyant detergents such as C₈E₅ (Fleming *et al.*, 1997). These detergents do not contribute to the sedimentation during AUC and are essentially ignored. The major problem with this method is solubilising enough of the integral membrane protein in the detergent to get a good signal in the AUC.

7.4 Energetic requirements and characteristics of Neu5Ac transport via SiaPQM

The first TRAP transporter to be reconstituted and examined *in vitro* has been reported (Chapter 6). Assessment of the energetic requirements for transporter by reconstituted SiaPQM revealed that transport of Neu5Ac into the lumen of the proteoliposomes was dependent on the presence of an inwardly-directed $\Delta\mu\text{Na}^+$. Application of specific gradients across the membrane revealed that a ΔpH or a $\Delta\Psi$ alone could not support transport of Neu5Ac. This dependence of Na^+ is supported by data from a number of different sources. It has been shown that there is a strong correlation between high numbers of TRAP transporters and environments with high Na^+ concentrations (Chapter 3). Marine-dwelling bacteria appear to have a propensity for TRAP transporter above bacteria found in other environments. A more direct line of evidence that TRAP transporters are Na^+ -dependent transporters is that whole cell transport assays revealed an ESR-dependent transporter in *R. sphaeroides*, assumed to be a TRAP transporter, which was found to be Na^+ -dependent (Jacobs *et al.*, 1996). Another example similar to this is that Na^+ -dependent glutamate transport in

Synecocystis was found to be mediated by an atypical TRAP transporter that utilised an ABC transporter ESR (homologous to GlnH (Quintero *et al.*, 2001)). The high resolution structure of a TRAP ESR from *R. sphaeroides* has been solved in complex with a pyruvate molecule and a Na⁺ ion in the binding site. This sodium ion interacts with both the protein and the ligand indicating that it is integral to the binding process. This does not appear to be the case for all TRAP transporter ESRs as it has been shown that the binding affinity of Neu5Ac to VC1779 in Na⁺-free buffer is not significantly different to when 120 mM NaCl is added (Chapter 6). The same observation was made for the ESR from the presumed glutamate-specific TRAP transporter from *R. sphaeroides* (Jacobs *et al.*, 1996). This indicates in these two cases that Na⁺ is not required for high affinity binding of ligand.

In the case of SiaPQM, Na⁺ does not appear play a role in facilitating binding of the ligand to the ESR. It has been observed that a gradient of Na⁺ is required for transport and not simply the presence of Na⁺ in the buffer (Chapter 6). Unfortunately, it was not possible to assess the effects of Na⁺ ions on the binding affinity of Neu5Ac for SiaP using tryptophan fluorescence spectroscopy. SiaP has a very low fluorescence signal due to the lack of tryptophan residues, which makes assessing the importance of Na⁺ in ligand binding very difficult and potentially unreliable. A method to circumvent this problem would be to isothermal titration calorimetry (ITC).

Although the effects of Na⁺ on Neu5Ac binding to SiaP have not been assessed directly, the evidence would suggest that the contribution of Na⁺ to the transport mechanism is not in the coordination of the ligand in the ESR component. However, this characteristic does not necessarily extend itself to all TRAP transporters.

The application of different combinations of electrochemical gradients across the membrane was performed on SiaQM-containing proteoliposomes (Chapter 6). Aside from the separate application of a $\Delta\mu\text{Na}^+$, a ΔpH and a $\Delta\Psi$ across the membrane, combinations of $\Delta\mu\text{Na}^+ + \Delta\text{pH}$ and $\Delta\mu\text{Na}^+ + \Delta\Psi$ were also applied. Transport measured under these conditions revealed that upon application of $\Delta\mu\text{Na}^+ + \Delta\text{pH}$ there was a decrease in the transport rate when compared to $\Delta\mu\text{Na}^+$ alone. There are a number of ways the internal pH, external pH and the ΔpH across the membrane can affect transport by secondary transporters

(reviewed in (Poolman *et al.*, 1987)). The pH can affect transport by ABC transporters, secondary transporters and group translocators (Poolman, 1994). It is easy to envisage how the change in pH may affect secondary transporters that use ΔpH as a driving force as this could interfere with discreet (de)protonation steps integral to the transport mechanism (Poolman *et al.*, 1987). However, it is less obvious how the ΔpH could affect a $\Delta\mu\text{Na}^+$ driven transporter like SiaPQM. There is of course the effect of pH on the ionic state of the substrate or its effects on the transporter itself, which may affect critical residues that would result in decreased transport. The magnitude of the internal pH has also been implicated in allosteric effects for certain transporters (Poolman *et al.*, 1987). However, the reason for the observed decrease in transport rate via SiaPQM when a $\Delta\mu\text{Na}^+$ + ΔpH are applied is not known and further investigations would be required to draw any conclusions.

Application of a combination of $\Delta\mu\text{Na}^+$ + $\Delta\Psi$ led to an increase in the transport rate compared to application of $\Delta\mu\text{Na}^+$ alone. Neu5Ac contains a carboxylate group and is negatively charged at physiological pH ($\text{pK}_a = 2.6$). The $\Delta\Psi$ that is applied produces a charge separation across the membrane with the inside negative compared to outside. The fact that Neu5Ac is negatively charged and the addition of a $\Delta\Psi$ increases transport rate despite it making the inside of the proteoliposome negative compared to the outside suggests that more than one Na^+ ion is transported per molecule of Neu5Ac. The exact stoichiometry of Neu5Ac to Na^+ is not known, however, it is an important detail to establish with regard to the mechanism of transport. A number of methods are available to determine the stoichiometry of H^+ and substrate for other transporters, most of which can also be applied to determining the $\text{Na}^+:\text{Neu5Ac}$ stoichiometry in SiaPQM. One method is to include pH probe in the transport assay and determine the change in pH as transport proceeds. Symport of substrate with H^+ will lead to acidification of the proteoliposome lumen and an increase in pH value in the external buffer, which can be converted into the number of H^+ transported (Zilberstein *et al.*, 1979). If a parallel transport assay is performed in which transport rate of substrate is measured then the two values can be combined and a stoichiometry established. Na^+ -specific probes are available, therefore this method can be directly applied to SiaPQM. Another method is to load the lumen of the vesicle/proteoliposome with a dye such as 2',7'-bis-(2-carboxyethyl)-5(6)-carboxyfluorescein (BCECF) that alters its emission profile

upon acidification of the lumen (Graves *et al.*, 2008). A dye that reacts to the presence of Na^+ is also available and could be used to determine the rate of Na^+ transport into the proteoliposome (Lo *et al.*, 2006).

As detailed in section 1.5.4 of the introduction, classical secondary transporters share common characteristics under certain experimental conditions. Solute counterflow and exchange ("cold chase") reactions are characteristic of secondary transporters and are only possible because of the bidirectional nature of transport by secondary transporters. TRAP transporters are secondary active transporters, but utilise an ESR component, which makes them inherently unidirectional. To assess whether SiaQM could transport Neu5Ac bidirectionally, solute counterflow and exchange reactions were performed on SiaQM-containing proteoliposomes. These experiments revealed that SiaQM (in the presence and absence of 5 μM SiaP) could not catalyse bidirectional movement of substrate (Chapter 6). In an ideal situation, these experiments would have been accompanied by a positive control for counterflow and exchange. The perfect control for these experiments would have been the classical secondary transporter NanT from *E. coli* and the perfect negative control would have been the ABC transporter SatABCD from *H. ducreyi*. Both of these systems are specific for Neu5Ac, but neither system had been reconstituted into liposomes for use in these types of experiments.

An experiment was performed to assess the effects of varying the concentration of SiaP whilst maintaining the same concentration of SiaQM and Neu5Ac (Chapter 6). The Neu5Ac concentration was maintained at 5 μM and SiaP was varied between 0 μM , 2.5 μM , 5 μM , 10 μM and 20 μM . With no SiaP present, there was no transport indicating that transport of Neu5Ac by SiaQM is dependent on the presence of SiaP. The transport rates in the presence of 2.5 μM or 5 μM SiaP were almost indistinguishable from each other. The presence of 10 μM SiaP facilitated uptake of Neu5Ac, but unexpectedly eventually led to efflux of the substrate. This was also observed in the presence of 20 μM , but Neu5Ac was effluxed to higher degree, suggesting that this efflux was facilitated by SiaP.

To clarify whether this efflux was a direct effect of excess SiaP, a transport assay was performed in the presence of 5 μM Neu5Ac and 2.5 μM SiaP. At 100

seconds, a SiaP "chase" was added making the final concentration of SiaP 20 μM . This resulted in rapid efflux of Neu5Ac. An experiment was performed to determine whether this efflux was caused by a specific interaction between SiaP and SiaQM or whether it was a consequence of the excess SiaP binding all of the free Neu5Ac in the external buffer and forming a pseudo-substrate gradient that was higher in magnitude than the $\Delta\mu\text{Na}^+$ and $\Delta\Psi$ combined. The experiment performed to ascertain which of these possibilities was correct took advantage of the availability of VC1779, a high affinity Neu5Ac ESR that was known not to functionally interact with SiaQM. If a VC1779 "chase" experiment produced the same efflux as the SiaP chase then it would suggest that efflux is not due to a specific interaction between SiaP and SiaQM. This would indicate that there is a different reason for the efflux, possibly the formation of a pseudo-substrate gradient as mentioned above. The results showed that efflux was not facilitated by a VC1779 chase as it was for a SiaP chase indicating that a specific interaction between SiaP and SiaQM caused this efflux.

Whilst this is an interesting observation, this situation is unlikely to occur *in vivo* as the transporter and catabolic genes are co-regulated. Upon transport of Neu5Ac into the cytoplasm, the substrate would be catabolised making efflux of Neu5Ac impossible. By using an artificial system, it has been demonstrated that under certain non-physiological conditions SiaPQM is able to perform bidirectional transport.

7.5 Proposed model for TRAP transporters

Using what is now known for TRAP transporters and what has been defined for classical secondary transporters and ABC transporters, a mechanism of transport can be proposed for TRAP transporters. The model will be a general TRAP transporter model so the different protein subunits will be referred to as DctP, DctQ and DctM. The transport mechanism begins with binding of ligand to the ESR domain, DctP (Fig. 7.1a). The ESR docks with the integral membrane components DctQM in their assumed resting state. This interaction has been shown to be specific as VC1779 cannot form a functional transporter with SiaQM *in vitro* (Chapter 6). It is currently unknown whether DctP interacts with DctQ or DctM or both and it is also unknown whether unliganded DctP can interact with the membrane components. The resting state of DctM, which is the component predicted to form the translocation channel, is assumed to be such

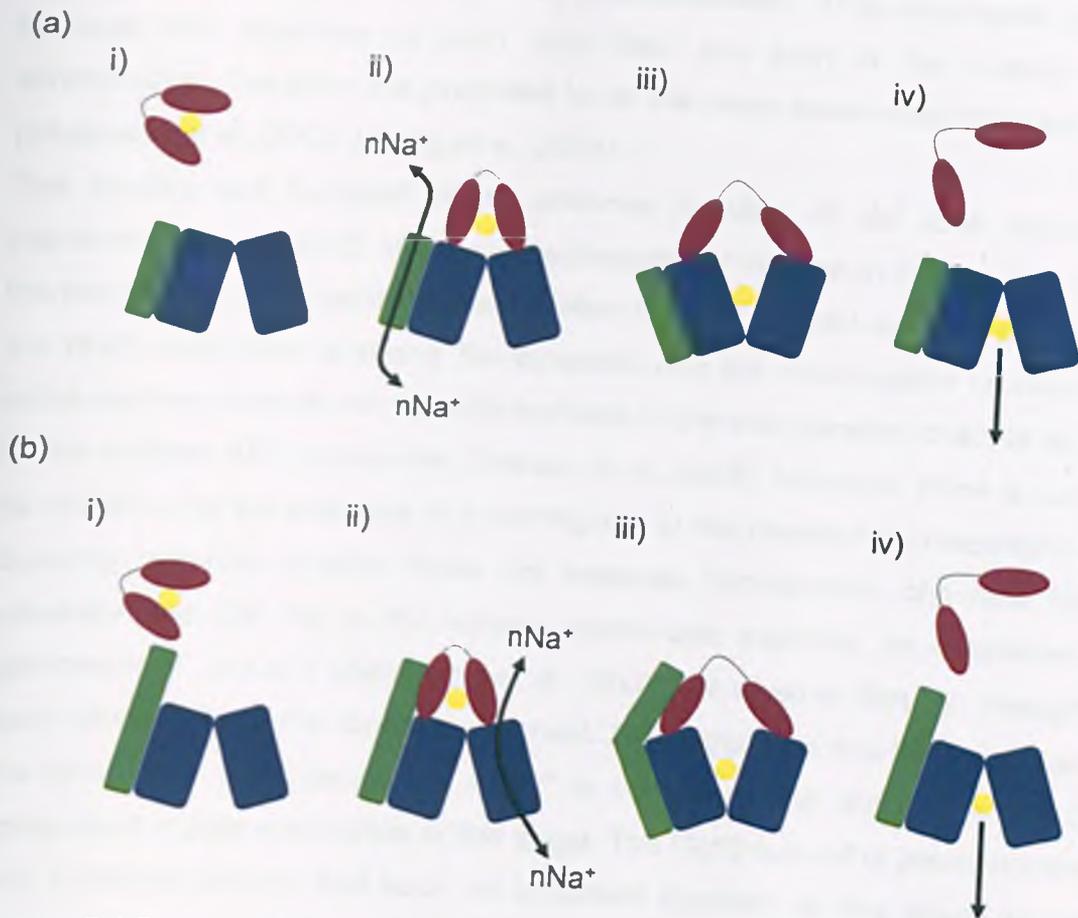


Figure 7.1 Proposed models for the transport mechanism TRAP transporters. Proposed model (a) begins with i) substrate (yellow circle) binding to the ESR (red) in the periplasm. ii) Liganded ESR docks with DctM (blue rectangles) in the outside-closed conformation. Transport of $n\text{Na}^+$ ions through DctQ (green rectangle) induces the outside-open conformation, opening of the ESR and releasing the substrate into DctM, as depicted in iii). iv) DctM relaxes back to outside-closed conformation, opening the cytoplasmic side of DctM and allowing release of substrate into the cytoplasm. Unliganded ESR is released back into the periplasm. Proposed model (b) begins i) substrate binding to the ESR in the periplasm. ii) Liganded ESR docks with DctQ with DctM in the outside-closed conformation. Transport of $n\text{Na}^+$ ions through DctM induces the outside-open conformation, opening of the ESR and releasing the substrate into DctM, as depicted in iii). iv) DctM relaxes back to outside-closed conformation, opening the cytoplasmic side of DctM and allowing release of substrate into the cytoplasm. Unliganded ESR is released back into the periplasm.

that the periplasmic opening is closed (outside-closed). This assumption is made because the structures of LacY and GlpT are both in the outside-closed conformation, therefore it is predicted to be the most stable conformational state (Abramson et al., 2003, Huang *et al.*, 2003).

The binding and transport of an unknown number of Na⁺ ions across the membrane through DctQ induces a conformational change in DctM which opens the periplasmic entrance to the translocation channel in DctM and also prises open the DctP component releasing the substrate into the translocation channel. The substrate may interact with specific residues in the translocation channel as seen for the maltose ABC transporter (Oldham et al., 2007). However, there is currently no evidence for the presence of a binding site in the membrane components. It is currently unknown whether there are separate translocation channels for the substrate and Na⁺ ion in the integral membrane subunits, as determined for galactoside:H⁺ in LacY (Abramson et al., 2003), or whether they go through the same channel in the membrane component. The suggestion that DctQ functions as the component that couples the $\Delta\mu\text{Na}^+$ to conformational changes in the DctM component is pure speculation at this stage. The DctQ subunit is poorly conserved and it seems unlikely that such an important function as this would not have conserved residues. Upon release of substrate into the translocation channel the DctM collapses back to its more favourable resting state (outside-closed), which releases the DctP component and opens the cytoplasmic entrance of the DctM component allowing release of the substrate into the cytoplasm.

Another model could be proposed with a different role for DctQ (Fig. 7.1b). In this alternate model, liganded DctP interacts with DctQ. DctM alone forms the translocation channel for the substrate *and* Na⁺ ion. The argument in favour of this is that DctQ is not conserved enough for such an important function as forming the Na⁺ channel. Furthermore, standalone secondary transporters are able to couple transport of substrate and counter ion without the need for auxiliary proteins like DctQ, so there is no reason why DctM could not. DctP and DctQ have a lot more sequence divergence than DctM so it possible they have diverged together thus conferring specificity of ESR on each transporter.

The models so far suggested do not take into account the apparent efflux of substrate observed upon addition of excess unliganded ESR (Chapter 6). Although this result is difficult to explain without further experiments, the results obtained in Chapter 6 allow for some speculation. The following explanation of the results from Chapter 6 will be assumed to be true for the benefit of this model, however, the validity of this explanation is still to be confirmed. Prior to any interaction with the DctP component, Dct(Q)M is in its resting state, which is assumed to be an outside-closed conformation, which is also assumed for LacY and GlpT (Abramson et al., 2003, Huang et al., 2003). Due to the high concentration of substrate in the lumen of the proteoliposome, the binding site within DctM is saturated. Upon interaction with the excess unliganded DctP component, the conformation of DctM changes to outside-open. At this point, DctP either detaches and substrate diffuses out or DctP binds substrate and detaches in a liganded form. The latter option is possible as it has been shown that Neu5Ac makes the majority of contacts with only one globular domain of SiaP (Muller et al., 2006). The model that could be suggested from this is that liganded DctP docks with DctM and induces a conformational change in DctM. This is analogous to the model suggested for the maltose ABC transporter (Oldham et al., 2007). This conformational change opens DctP releasing the substrate, which enters the translocation channel of DctM, possibly interacting with specific residues within the channel. The binding and transport of Na^+ could then induce the outside-closed conformation, breaking the interaction with DctP and releasing substrate into the cytoplasm. There are clearly several problems with this model, such as, what is stopping the unliganded DctP from rebinding to DctM in its resting state? Also, if the outside-closed conformation is the resting state of DctM, there is presumably no energetic requirement (from $\Delta\mu\text{Na}^+$) to return to that conformation. Would it not be more likely that the opening of DctP and the induction of the outside-open conformation is the energy – dependent step in the transport cycle?

Clearly, there are a number of potential models that could be proposed. This large number of possibilities is due to the lack of experimental evidence regarding the structure and mechanism of TRAP transporters.

An important question to consider is: why does a secondary transporter require an ESR component? The recruitment of an ESR by a secondary transporter is probably a way to increase the affinity of the transporter to enable scavenging of low concentrations of substrate from the environment. In organisms containing TRAP transporters it is presumably less energetically expensive to use a $\Delta\mu\text{Na}^+$ to transport a particular substrate than to synthesise ATP to drive transport. This correlation is clearly observed in the propensity for marine-dwelling bacteria to have high numbers of TRAP transporters (Chapter 3). However, these organisms still require the presence of high affinity transport via ABC transporters. The reason why an organism requires both ABC and TRAP transporters may be due to the substrate range of TRAP transporters being confined to organic anions, whereas ABC transporters can transport a plethora of different substrates. The reason for this limitation of substrate range is not immediately obvious as TRAP transporters are assumed to be an ancient family of transporters and will have had time to diversify, like ABC transporters (Rabus et al., 1999). However, the very nature of the substrate may have profound implications on the mechanism of transport by TRAP transporters that we do not know about yet.

Other questions that still remain unanswered, but are pivotal for the understanding of TRAP transporters are; what is the function of the DctQ component? How does the ESR component interact with the membrane components? Can the DctM subunit function on its own as a standalone secondary transporter? Are all TRAP transporters limited to the transport of organic anionic substrates? Do all TRAP transporters utilise a $\Delta\mu\text{Na}^+$ for driving transport? More experiments, using these proposed models as a basis, will lead to a deeper understanding of the functions and interactions of TRAP transporter subunits which will shed more light on the mechanism of transport.

References

- Abramson, J., I. Smirnova, V. Kasho, G. Verner, H. R. Kaback & S. Iwata, (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**: 610-615.
- Allen, S., A. Zaleski, J. W. Johnston, B. W. Gibson & M. A. Apicella, (2005) Novel sialic acid transporter of *Haemophilus influenzae*. *Infect Immun* **73**: 5291-5300.
- Ames, G. F., C. E. Liu, A. K. Joshi & K. Nikaido, (1996) Liganded and unliganded receptors interact with equal affinity with the membrane complex of periplasmic permeases, a subfamily of traffic ATPases. *J Biol Chem* **271**: 14264-14270.
- Ames, G. F. & K. Nikaido, (1978) Identification of a membrane protein as a histidine transport component in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* **75**: 5447-5451.
- Ames, G. F., K. Nikaido, J. Groarke & J. Petithory, (1989) Reconstitution of periplasmic transport in inside-out membrane vesicles. Energization by ATP. *J Biol Chem* **264**: 3998-4002.
- An, J. H. & Y. S. Kim, (1998) A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in *Rhizobium trifolii*--cloning, sequencing, and expression of the enzymes in *Escherichia coli*. *Eur J Biochem* **257**: 395-402.
- An, J. H., H. Y. Lee, K. N. Ko, E. S. Kim & Y. S. Kim, (2002) Symbiotic effects of *deltamatB Rhizobium leguminosarum* bv. *trifolii* mutant on clovers. *Mol Cells* **14**: 261-266.
- Angata, T. & A. Varki, (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev* **102**: 439-469.
- Antoine, R., F. Jacob-Dubuisson, H. Drobecq, E. Willery, S. Lesjean & C. Locht, (2003) Overrepresentation of a gene family encoding extracytoplasmic solute receptors in *Bordetella*. *J Bacteriol* **185**: 1470-1474.
- Asai, K., S. H. Baik, Y. Kasahara, S. Moriya & N. Ogasawara, (2000) Regulation of the transport system for C4-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* **146 (Pt 2)**: 263-271.

- Auer, M., M. J. Kim, M. J. Lemieux, A. Villa, J. Song, X. D. Li & D. N. Wang, (2001) High-yield expression and functional analysis of Escherichia coli glycerol-3-phosphate transporter. *Biochemistry* **40**: 6628-6635.
- Austermuhle, M. I., J. A. Hall, C. S. Klug & A. L. Davidson, (2004) Maltose-binding protein is open in the catalytic transition state for ATP hydrolysis during maltose transport. *J Biol Chem* **279**: 28243-28250.
- Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner & H. Mori, (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006 0008.
- Baron, C. & T. E. Thompson, (1975) Solubilization of bacterial membrane proteins using alkyl glucosides and dioctanoyl phosphatidylcholine. *Biochim Biophys Acta* **382**: 276-285.
- Blakey, D., A. Leech, G. H. Thomas, G. Coutts, K. Findlay & M. Merrick, (2002) Purification of the Escherichia coli ammonium transporter AmtB reveals a trimeric stoichiometry. *Biochem J* **364**: 527-535.
- Bluschke, B., V. Eckey, B. Kunert, S. Berendt, H. Landmesser, M. Portwich, R. Volkmer & E. Schneider, (2007) Mapping putative contact sites between subunits in a bacterial ATP-binding cassette (ABC) transporter by synthetic peptide libraries. *J Mol Biol* **369**: 386-399.
- Bolanos-Garcia, V. M. & O. R. Davies, (2006) Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta* **1760**: 1304-1313.
- Borths, E. L., K. P. Locher, A. T. Lee & D. C. Rees, (2002) The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter. *Proc Natl Acad Sci U S A* **99**: 16642-16647.
- Borths, E. L., B. Poolman, R. N. Hvorup, K. P. Locher & D. C. Rees, (2005) In Vitro Functional Characterization of BtuCD-F, the Escherichia coli ABC Transporter for Vitamin B(12) Uptake. *Biochemistry* **44**: 16301-16309.
- Bos, M. P., V. Robert & J. Tommassen, (2007) Biogenesis of the gram-negative bacterial outer membrane. *Annu Rev Microbiol* **61**: 191-214.
- Bouchet, V., D. W. Hood, J. Li, J. R. Brisson, G. A. Randle, A. Martin, Z. Li, R. Goldstein, E. K. Schweda, S. I. Pelton, J. C. Richards & E. R. Moxon, (2003)

- Host-derived sialic acid is incorporated into Haemophilus influenzae lipopolysaccharide and is a major virulence factor in experimental otitis media. *Proc Natl Acad Sci U S A* **100**: 8898-8903.
- Boyd, D., B. Traxler & J. Beckwith, (1993) Analysis of the topology of a membrane protein by using a minimum number of alkaline phosphatase fusions. *J Bacteriol* **175**: 553-556.
- Bruggemann, C., K. Denger, A. M. Cook & J. Ruff, (2004) Enzymes and genes of taurine and isethionate dissimilation in Paracoccus denitrificans. *Microbiology* **150**: 805-816.
- Buchel, D. E., B. Gronenborn & B. Muller-Hill, (1980) Sequence of the lactose permease gene. *Nature* **283**: 541-545.
- Casadaban, M. J. & S. N. Cohen, (1980) Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *J Mol Biol* **138**: 179-207.
- Chae, J. C., Y. Kim, Y. C. Kim, G. J. Zylstra & C. K. Kim, (2000) Genetic structure and functional implication of the fcb gene cluster for hydrolytic dechlorination of 4-chlorobenzoate from Pseudomonas sp. DJ-12. *Gene* **258**: 109-116.
- Chen, J., G. Lu, J. Lin, A. L. Davidson & F. A. Quijcho, (2003) A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol Cell* **12**: 651-661.
- Chen, J., S. Sharma, F. A. Quijcho & A. L. Davidson, (2001) Trapping the transition state of an ATP-binding cassette transporter: evidence for a concerted mechanism of maltose transport. *Proc Natl Acad Sci U S A* **98**: 1525-1530.
- Chohnan, S., Y. Kurusu, H. Nishihara & Y. Takamura, (1999) Cloning and characterization of mdc genes encoding malonate decarboxylase from Pseudomonas putida. *FEMS Microbiol Lett* **174**: 311-319.
- Cole, S. P., G. Bhardwaj, J. H. Gerlach, J. E. Mackie, C. E. Grant, K. C. Almquist, A. J. Stewart, E. U. Kurz, A. M. Duncan & R. G. Deeley, (1992) Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* **258**: 1650-1654.
- Corfield, T., (1992) Bacterial sialidases--roles in pathogenicity and nutrition. *Glycobiology* **2**: 509-521.

- Covitz, K. M., C. H. Panagiotidis, L. I. Hor, M. Reyes, N. A. Treptow & H. A. Shuman, (1994) Mutations that alter the transmembrane signalling pathway in an ATP binding cassette (ABC) transporter. *EMBO J* **13**: 1752-1759.
- Curnow, P., M. Lorch, K. Charalambous & P. J. Booth, (2004) The reconstitution and activity of the small multidrug transporter EmrE is modulated by non-bilayer lipid composition. *J Mol Biol* **343**: 213-222.
- Dassa, E. & M. Hofnung, (1985) Sequence of gene malG in E. coli K12: homologies between integral membrane components from binding protein-dependent transport systems. *EMBO J* **4**: 2287-2293.
- Daus, M. L., S. Berendt, S. Wuttge & E. Schneider, (2007) Maltose binding protein (MalE) interacts with periplasmic loops P2 and P1 respectively of the MalFG subunits of the maltose ATP binding cassette transporter (MalFGK(2)) from Escherichia coli/Salmonella during the transport cycle. *Mol Microbiol* **66**: 1107-1122.
- Davidson, A. L., E. Dassa, C. Orelle & J. Chen, (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev* **72**: 317-364, table of contents.
- Davidson, A. L., S. S. Laghaeian & D. E. Mannering, (1996) The maltose transport system of Escherichia coli displays positive cooperativity in ATP hydrolysis. *J Biol Chem* **271**: 4858-4863.
- Davidson, A. L. & H. Nikaido, (1991) Purification and characterization of the membrane-associated components of the maltose transport system from Escherichia coli. *J Biol Chem* **266**: 8946-8951.
- Davidson, A. L., H. A. Shuman & H. Nikaido, (1992) Mechanism of maltose transport in Escherichia coli: transmembrane signaling by periplasmic binding proteins. *Proc Natl Acad Sci U S A* **89**: 2360-2364.
- Davies, S. J., P. Golby, D. Omrani, S. A. Broad, V. L. Harrington, J. R. Guest, D. J. Kelly & S. C. Andrews, (1999) Inactivation and regulation of the aerobic C(4)-dicarboxylate transport (dctA) gene of Escherichia coli. *J Bacteriol* **181**: 5624-5635.
- Dawson, R. J., K. Hollenstein & K. P. Locher, (2007) Uptake or extrusion: crystal structures of full ABC transporters suggest a common mechanism. *Mol Microbiol* **65**: 250-257.

- Dawson, R. J. & K. P. Locher, (2006) Structure of a bacterial multidrug ABC transporter. *Nature* **443**: 180-185.
- Dean, D. A., A. L. Davidson & H. Nikaido, (1989) Maltose transport in membrane vesicles of *Escherichia coli* is linked to ATP hydrolysis. *Proc Natl Acad Sci U S A* **86**: 9134-9138.
- Denger, K., T. H. Smits & A. M. Cook, (2006) Genome-enabled analysis of the utilization of taurine as sole source of carbon or of nitrogen by *Rhodobacter sphaeroides* 2.4.1. *Microbiology* **152**: 3197-3206.
- Dey, S. & B. P. Rosen, (1995) Dual mode of energy coupling by the oxyanion-translocating ArsB protein. *J Bacteriol* **177**: 385-389.
- Diederichs, K., J. Diez, G. Grellner, C. Muller, J. Breed, C. Schnell, C. Vornrhein, W. Boos & W. Welte, (2000) Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*. *EMBO J* **19**: 5951-5961.
- Dietzel, I., V. Kolb & W. Boos, (1978) Pole cap formation in *Escherichia coli* following induction of the maltose-binding protein. *Arch Microbiol* **118**: 207-218.
- Diez, J., K. Diederichs, G. Grellner, R. Horlacher, W. Boos & W. Welte, (2001) The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic Archaeon *Thermococcus litoralis* at 1.85 Å. *J Mol Biol* **305**: 905-915.
- Dinh, T., I. T. Paulsen & M. H. Saier, Jr., (1994) A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria. *J Bacteriol* **176**: 3825-3831.
- Dippel, R. & W. Boos, (2005) The maltodextrin system of *Escherichia coli*: metabolism and transport. *J Bacteriol* **187**: 8322-8331.
- Doeven, M. K., R. Abele, R. Tampe & B. Poolman, (2004) The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *J Biol Chem* **279**: 32301-32307.
- Doeven, M. K., G. van den Bogaart, V. Krasnikov & B. Poolman, (2008) Probing receptor-translocator interactions in the oligopeptide ABC transporter by fluorescence correlation spectroscopy. *Biophys J* **94**: 3956-3965.

- Driessen, A. J., C. van Leeuwen & W. N. Konings, (1989) Transport of basic amino acids by membrane vesicles of *Lactococcus lactis*. *J Bacteriol* **171**: 1453-1458.
- Dumon-Seignover, L., G. Cariot & L. Vuillard, (2004) The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr Purif* **37**: 203-206.
- Ehrle, R., C. Pick, R. Ulrich, E. Hofmann & M. Ehrmann, (1996) Characterization of transmembrane domains 6, 7, and 8 of MalF by mutational analysis. *J Bacteriol* **178**: 2255-2262.
- Ehrmann, M., D. Boyd & J. Beckwith, (1990) Genetic analysis of membrane protein topology by a sandwich gene fusion approach. *Proc Natl Acad Sci U S A* **87**: 7574-7578.
- Eshaghi, S., M. Hedren, M. I. Nasser, T. Hammarberg, A. Thornell & P. Nordlund, (2005) An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. *Protein Sci* **14**: 676-683.
- Ferenci, T., W. Boos, M. Schwartz & S. Szmelcman, (1977) Energy-coupling of the transport system of *Escherichia coli* dependent on maltose-binding protein. *Eur J Biochem* **75**: 187-193.
- Figueira, M. A., S. Ram, R. Goldstein, D. W. Hood, E. R. Moxon & S. I. Pelton, (2007) Role of complement in defense of the middle ear revealed by restoring the virulence of nontypeable *Haemophilus influenzae* siaB mutants. *Infect Immun* **75**: 325-333.
- Fleming, K. G., A. L. Ackerman & D. M. Engelman, (1997) The effect of point mutations on the free energy of transmembrane alpha-helix dimerization. *J Mol Biol* **272**: 266-275.
- Forward, J. A., M. C. Behrendt, N. R. Wyborn, R. Cross & D. J. Kelly, (1997) TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *J Bacteriol* **179**: 5482-5493.
- Foster, D. L., M. Boublik & H. R. Kaback, (1983) Structure of the lac carrier protein of *Escherichia coli*. *J Biol Chem* **258**: 31-34.
- Fox, K. L., A. D. Cox, M. Gilbert, W. W. Wakarchuk, J. Li, K. Makepeace, J. C. Richards, E. R. Moxon & D. W. Hood, (2006) Identification of a bifunctional

- lipopolysaccharide sialyltransferase in *Haemophilus influenzae*: incorporation of disialic acid. *J Biol Chem* **281**: 40024-40032.
- Frillingos, S., M. Sahin-Toth, J. Wu & H. R. Kaback, (1998) Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins. *FASEB J* **12**: 1281-1299.
- Fu, D., A. Libson, L. J. Miercke, C. Weitzman, P. Nollert, J. Krucinski & R. M. Stroud, (2000) Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**: 481-486.
- Fukami-Kobayashi, K., Y. Tateno & K. Nishikawa, (1999) Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J Mol Biol* **286**: 279-290.
- Fukami-Kobayashi, K., Y. Tateno & K. Nishikawa, (2003) Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. *Mol Biol Evol* **20**: 267-277.
- Gaudet, R. & D. C. Wiley, (2001) Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing. *EMBO J* **20**: 4964-4972.
- Geertsma, E. R., (2005) What lies between: Functional interfaces in a dimeric transporter. In: Groningen: University of Groningen, pp. 99.
- Geertsma, E. R. & B. Poolman, (2007) High-throughput cloning and expression in recalcitrant bacteria. *Nat Methods* **4**: 705-707.
- Gilson, E., G. Alloing, T. Schmidt, J. P. Claverys, R. Dudler & M. Hofnung, (1988) Evidence for high affinity binding-protein dependent transport systems in gram-positive bacteria and in *Mycoplasma*. *EMBO J* **7**: 3971-3974.
- Golby, P., D. J. Kelly, J. R. Guest & S. C. Andrews, (1998) Topological analysis of DcuA, an anaerobic C4-dicarboxylate transporter of *Escherichia coli*. *J Bacteriol* **180**: 4821-4827.
- Gonin, S., P. Arnoux, B. Pierru, J. Lavergne, B. Alonso, M. Sabaty & D. Pignol, (2007) Crystal structures of an Extracytoplasmic Solute Receptor from a TRAP transporter in its open and closed forms reveal a helix-swapped dimer requiring a cation for alpha-keto acid binding. *BMC Struct Biol* **7**: 11.
- Grammann, K., A. Volke & H. J. Kunte, (2002) New type of osmoregulated solute transporter identified in halophilic members of the bacteria domain: TRAP

- transporter TeaABC mediates uptake of ectoine and hydroxyectoine in *Halomonas elongata* DSM 2581(T). *J Bacteriol* **184**: 3078-3085.
- Graves, A. R., P. K. Curran, C. L. Smith & J. A. Mindell, (2008) The Cl⁻/H⁺ antiporter CIC-7 is the primary chloride permeation pathway in lysosomes. *Nature* **453**: 788-792.
- Guan, L., O. Mirza, G. Verner, S. Iwata & H. R. Kaback, (2007) Structural determination of wild-type lactose permease. *Proc Natl Acad Sci U S A* **104**: 15294-15298.
- Guan, L., F. D. Murphy & H. R. Kaback, (2002) Surface-exposed positions in the transmembrane helices of the lactose permease of *Escherichia coli* determined by intermolecular thiol cross-linking. *Proc Natl Acad Sci U S A* **99**: 3475-3480.
- Guzman, L. M., D. Belin, M. J. Carson & J. Beckwith, (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Hamblin, M. J., J. G. Shaw, J. P. Curson & D. J. Kelly, (1990) Mutagenesis, cloning and complementation analysis of C₄-dicarboxylate transport genes from *Rhodobacter capsulatus*. *Mol Microbiol* **4**: 1567-1574.
- Harold, F. M., (1986) *The vital force: a study of bioenergetics*. W.H. Freeman and company.
- Heyne, R. I., W. de Vrij, W. Crielaard & W. N. Konings, (1991) Sodium ion-dependent amino acid transport in membrane vesicles of *Bacillus stearothermophilus*. *J Bacteriol* **173**: 791-800.
- Higgins, C. F. a. L., K.J., (2003) ABC transporter: an introduction and overview. In: ABC transporter. From bacteria to man. I. B. Holland, Cole, S.P.C., Kuchler, K. and Higgins, C.F. (ed). Academic Press, pp.
- Hirai, T., J. A. Heymann, P. C. Maloney & S. Subramaniam, (2003) Structural model for 12-helix transporters belonging to the major facilitator superfamily. *J Bacteriol* **185**: 1712-1718.
- Hirota, N. & Y. Imae, (1983) Na⁺-driven flagellar motors of an alkalophilic *Bacillus* strain YN-1. *J Biol Chem* **258**: 10577-10581.
- Hollenstein, K., R. J. Dawson & K. P. Locher, (2007a) Structure and mechanism of ABC transporter proteins. *Curr Opin Struct Biol* **17**: 412-418.

- Hollenstein, K., D. C. Frei & K. P. Locher, (2007b) Structure of an ABC transporter in complex with its binding protein. *Nature* **446**: 213-216.
- Hong, J. S., A. G. Hunt, P. S. Masters & M. A. Lieberman, (1979) Requirements of acetyl phosphate for the binding protein-dependent transport systems in *Escherichia coli*. *Proc Natl Acad Sci U S A* **76**: 1213-1217.
- Hood, D. W., A. D. Cox, M. Gilbert, K. Makepeace, S. Walsh, M. E. Deadman, A. Cody, A. Martin, M. Mansson, E. K. Schweda, J. R. Brisson, J. C. Richards, E. R. Moxon & W. W. Wakarchuk, (2001) Identification of a lipopolysaccharide alpha-2,3-sialyltransferase from *Haemophilus influenzae*. *Mol Microbiol* **39**: 341-350.
- Hood, D. W., K. Makepeace, M. E. Deadman, R. F. Rest, P. Thibault, A. Martin, J. C. Richards & E. R. Moxon, (1999) Sialic acid in the lipopolysaccharide of *Haemophilus influenzae*: strain distribution, influence on serum resistance and structural characterization. *Mol Microbiol* **33**: 679-692.
- Huang, Y., M. J. Lemieux, J. Song, M. Auer & D. N. Wang, (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* **301**: 616-620.
- Hung, L. W., I. X. Wang, K. Nikaido, P. Q. Liu, G. F. Ames & S. H. Kim, (1998) Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature* **396**: 703-707.
- Hunke, S., M. Mourez, M. Jehanno, E. Dassa & E. Schneider, (2000) ATP modulates subunit-subunit interactions in an ATP-binding cassette transporter (MalFGK2) determined by site-directed chemical cross-linking. *J Biol Chem* **275**: 15526-15534.
- Huvent, I., H. Belrhali, R. Antoine, C. Bompard, C. Locht, F. Jacob-Dubuisson & V. Villeret, (2006a) Crystal structure of *Bordetella pertussis* BugD solute receptor unveils the basis of ligand binding in a new family of periplasmic binding proteins. *J Mol Biol* **356**: 1014-1026.
- Huvent, I., H. Belrhali, R. Antoine, C. Bompard, C. Locht, F. Jacob-Dubuisson & V. Villeret, (2006b) Structural analysis of *Bordetella pertussis* BugE solute receptor in a bound conformation. *Acta Crystallogr D Biol Crystallogr* **62**: 1375-1381.

- Hvorup, R. N., B. A. Goetz, M. Niederer, K. Hollenstein, E. Perozo & K. P. Locher, (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* **317**: 1387-1390.
- Hyde, S. C., P. Emsley, M. J. Hartshorn, M. M. Mimmack, U. Gileadi, S. R. Pearce, M. P. Gallagher, D. R. Gill, R. E. Hubbard & C. F. Higgins, (1990) Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* **346**: 362-365.
- Jacob, F. & J. Monod, (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-356.
- Jacobs, M. H., T. van der Heide, A. J. Driessen & W. N. Konings, (1996) Glutamate transport in *Rhodobacter sphaeroides* is mediated by a novel binding protein-dependent secondary transport system. *Proc Natl Acad Sci U S A* **93**: 12786-12790.
- Janausch, I. G., E. Zientz, Q. H. Tran, A. Kroger & G. Unden, (2002) C₄-dicarboxylate carriers and sensors in bacteria. *Biochim Biophys Acta* **1553**: 39-56.
- Johnston, J. W., N. P. Coussens, S. Allen, J. C. Houtman, K. H. Turner, A. Zaleski, S. Ramaswamy, B. W. Gibson & M. A. Apicella, (2008) Characterization of the N-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable *Haemophilus influenzae* strain 2019. *J Biol Chem* **283**: 855-865.
- Joshi, A. K., S. Ahmed & G. Ferro-Luzzi Ames, (1989) Energy coupling in bacterial periplasmic transport systems. Studies in intact *Escherichia coli* cells. *J Biol Chem* **264**: 2126-2133.
- Kaback, H. R., M. Sahin-Toth & A. B. Weinglass, (2001) The kamikaze approach to membrane transport. *Nat Rev Mol Cell Biol* **2**: 610-620.
- Kaneko, M., A. Yamaguchi & T. Sawai, (1985) Energetics of tetracycline efflux system encoded by Tn10 in *Escherichia coli*. *FEBS Lett* **193**: 194-198.
- Karlsson, K. A., (1998) Meaning and therapeutic potential of microbial recognition of host glycoconjugates. *Mol Microbiol* **29**: 1-11.
- Karpowich, N., O. Martsinkevich, L. Millen, Y. R. Yuan, P. L. Dai, K. MacVey, P. J. Thomas & J. F. Hunt, (2001) Crystal structures of the MJ1267 ATP binding

- cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter. *Structure* **9**: 571-586.
- Kellermann, O. & S. Szmelcman, (1974) Active transport of maltose in *Escherichia coli* K12. Involvement of a "periplasmic" maltose binding protein. *Eur J Biochem* **47**: 139-149.
- Kelly, D. J. & G. H. Thomas, (2001) The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol Rev* **25**: 405-424.
- Kennedy, K. A. & B. Traxler, (1999) MalK forms a dimer independent of its assembly into the MalFGK2 ATP-binding cassette transporter of *Escherichia coli*. *J Biol Chem* **274**: 6259-6264.
- Kim, Y. J., S. Park, Y. K. Oh, W. Kang, H. S. Kim & E. Y. Lee, (2005) Purification and characterization of human caseinomacropeptide produced by a recombinant *Saccharomyces cerevisiae*. *Protein Expr Purif* **41**: 441-446.
- Knol, J., (1999) Membrane reconstitution and functional analysis of a sugar transport system. In. Groningen: University of Groningen, pp. 104.
- Knol, J., L. Veenhoff, W. J. Liang, P. J. Henderson, G. Leblanc & B. Poolman, (1996) Unidirectional reconstitution into detergent-destabilized liposomes of the purified lactose transport system of *Streptococcus thermophilus*. *J Biol Chem* **271**: 15358-15366.
- Kogure, K., (1998) Bioenergetics of marine bacteria. *Curr Opin Biotechnol* **9**: 278-282.
- Kolker, E., K. S. Makarova, S. Shabalina, A. F. Picone, S. Purvine, T. Holzman, T. Cherny, D. Armbruster, R. S. Munson, Jr., G. Kolesov, D. Frishman & M. Y. Galperin, (2004) Identification and functional analysis of 'hypothetical' genes expressed in *Haemophilus influenzae*. *Nucleic Acids Res* **32**: 2353-2361.
- Koman, A., S. Harayama & G. L. Hazelbauer, (1979) Relation of chemotactic response to the amount of receptor: evidence for different efficiencies of signal transduction. *J Bacteriol* **138**: 739-747.
- Kuhlbrandt, W., (2004) Biology, structure and mechanism of P-type ATPases. *Nat Rev Mol Cell Biol* **5**: 282-295.
- Kuipers, O. P., M. M. Beerthuyzen, R. J. Siezen & W. M. De Vos, (1993) Characterization of the nisin gene cluster nisABTCIPR of *Lactococcus lactis*.

- Requirement of expression of the *nisA* and *nisl* genes for development of immunity. *Eur J Biochem* **216**: 281-291.
- Kunji, E. R., K. W. Chan, D. J. Slotboom, S. Floyd, R. O'Connor & M. Monne, (2005) Eukaryotic membrane protein overproduction in *Lactococcus lactis*. *Curr Opin Biotechnol* **16**: 546-551.
- Kunji, E. R., D. J. Slotboom & B. Poolman, (2003) *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochim Biophys Acta* **1610**: 97-108.
- Kvint, K., L. Nachin, A. Diez & T. Nystrom, (2003) The bacterial universal stress protein: function and regulation. *Curr Opin Microbiol* **6**: 140-145.
- Lanyi, J. K., (1979) The role of Na⁺ in transport processes of bacterial membranes. *Biochim Biophys Acta* **559**: 377-397.
- Lanyi, J. K., (1993) Proton translocation mechanism and energetics in the light-driven pump bacteriorhodopsin. *Biochim Biophys Acta* **1183**: 241-261.
- Lau, F. W. & J. U. Bowie, (1997) A method for assessing the stability of a membrane protein. *Biochemistry* **36**: 5884-5892.
- Lewinson, O., A. T. Lee & D. C. Rees, (2008) The funnel approach to the precrystallization production of membrane proteins. *J Mol Biol* **377**: 62-73.
- Li, X. D., A. Villa, C. Gownley, M. J. Kim, J. Song, M. Auer & D. N. Wang, (2001) Monomeric state and ligand binding of recombinant GABA transporter from *Escherichia coli*. *FEBS Lett* **494**: 165-169.
- Liang, W. J., K. J. Wilson, H. Xie, J. Knol, S. Suzuki, N. G. Rutherford, P. J. Henderson & R. A. Jefferson, (2005) The *gusBC* genes of *Escherichia coli* encode a glucuronide transport system. *J Bacteriol* **187**: 2377-2385.
- Lilley, G. G., J. A. Barbosa & L. A. Pearce, (1998) Expression in *Escherichia coli* of the putative N-acetylneuraminase lyase gene (*nanA*) from *Haemophilus influenzae*: overproduction, purification, and crystallization. *Protein Expr Purif* **12**: 295-304.
- Lo, C. J., M. C. Leake & R. M. Berry, (2006) Fluorescence measurement of intracellular sodium concentration in single *Escherichia coli* cells. *Biophys J* **90**: 357-365.

- Locher, K. P., A. T. Lee & D. C. Rees, (2002) The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* **296**: 1091-1098.
- Londer, Y. Y., P. R. Pokkuluri, V. Orshonsky, L. Orshonsky & M. Schiffer, (2006) Heterologous expression of dodecaheme "nanowire" cytochromes c from *Geobacter sulfurreducens*. *Protein Expr Purif* **47**: 241-248.
- Lu, G., J. M. Westbrooks, A. L. Davidson & J. Chen, (2005) ATP hydrolysis is required to reset the ATP-binding cassette dimer into the resting-state conformation. *Proc Natl Acad Sci U S A* **102**: 17969-17974.
- Madigan, M. T. & H. Gest, (1979) Growth of the photosynthetic bacterium *Rhodospseudomonas capsulata* chemoautotrophically in darkness with H₂ as the energy source. *J Bacteriol* **137**: 524-530.
- Manson, M. D., W. Boos, P. J. Bassford, Jr. & B. A. Rasmussen, (1985) Dependence of maltose transport and chemotaxis on the amount of maltose-binding protein. *J Biol Chem* **260**: 9727-9733.
- Martoglio, B. & B. Dobberstein, (1998) Signal sequences: more than just greasy peptides. *Trends Cell Biol* **8**: 410-415.
- Marty-Teyssset, C., J. S. Lolkema, P. Schmitt, C. Divies & W. N. Konings, (1996) The citrate metabolic pathway in *Leuconostoc mesenteroides*: expression, amino acid synthesis, and alpha-ketocarboxylate transport. *J Bacteriol* **178**: 6209-6215.
- Mauchline, T. H., J. E. Fowler, A. K. East, A. L. Sartor, R. Zaheer, A. H. Hosie, P. S. Poole & T. M. Finan, (2006) Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc Natl Acad Sci U S A* **103**: 17933-17938.
- Mayfield, J. E., B. J. Bricker, H. Godfrey, R. M. Crosby, D. J. Knight, S. M. Halling, D. Balinsky & L. B. Tabatabai, (1988) The cloning, expression, and nucleotide sequence of a gene coding for an immunogenic *Brucella abortus* protein. *Gene* **63**: 1-9.
- Miller, D. M., 3rd, J. S. Olson, J. W. Pflugrath & F. A. Quiocho, (1983) Rates of ligand binding to periplasmic proteins involved in bacterial transport and chemotaxis. *J Biol Chem* **258**: 13665-13672.

- Miller, D. M., 3rd, J. S. Olson & F. A. Quiocho, (1980) The mechanism of sugar binding to the periplasmic receptor for galactose chemotaxis and transport in *Escherichia coli*. *J Biol Chem* **255**: 2465-2471.
- Miroux, B. & J. E. Walker, (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**: 289-298.
- Mohanty, A. K. & M. C. Wiener, (2004) Membrane protein expression and production: effects of polyhistidine tag length and position. *Protein Expr Purif* **33**: 311-325.
- Mokhonova, E. I., V. V. Mokhonov, H. Akama & T. Nakae, (2005) Forceful large-scale expression of "problematic" membrane proteins. *Biochem Biophys Res Commun* **327**: 650-655.
- Monne, M., K. W. Chan, D. J. Slotboom & E. R. Kunji, (2005) Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. *Protein Sci* **14**: 3048-3056.
- Morbach, S., S. Tebbe & E. Schneider, (1993) The ATP-binding cassette (ABC) transporter for maltose/maltodextrins of *Salmonella typhimurium*. Characterization of the ATPase activity associated with the purified MalK subunit. *J Biol Chem* **268**: 18617-18621.
- Morrison, S., A. Ward, C. J. Hoyle & P. J. Henderson, (2003) Cloning, expression, purification and properties of a putative multidrug resistance efflux protein from *Helicobacter pylori*. *Int J Antimicrob Agents* **22**: 242-249.
- Mourez, M., M. Hofnung & E. Dassa, (1997) Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits. *EMBO J* **16**: 3066-3077.
- Moxon, E. R. a. M., T.F., (2000) *Haemophilus influenzae*. In: Mandell, Douglas and Bennett's Principles and Practice of Infectious Disease. G. K. Mandell, Bennett, J.E., and Dolin, R. (ed). Philadelphia: Churchill Livingstone, pp. 2369-2378.
- Muller, A., E. Severi, C. Mulligan, A. G. Watts, D. J. Kelly, K. S. Wilson, A. J. Wilkinson & G. H. Thomas, (2006) Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a

- sialic acid-binding virulence factor from haemophilus influenzae. *J Biol Chem*.
- Murakami, S., R. Nakashima, E. Yamashita & A. Yamaguchi, (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **419**: 587-593.
- Neu, H. C. & L. A. Heppel, (1965) The release of enzymes from Escherichia coli by osmotic shock and during the formation of spheroplasts. *J Biol Chem* **240**: 3685-3692.
- Newman, M. J., D. L. Foster, T. H. Wilson & H. R. Kaback, (1981) Purification and reconstitution of functional lactose carrier from Escherichia coli. *J Biol Chem* **256**: 11804-11808.
- Nicholls, G. D. a. F., S.J., (2002) *Bioenergetics* 3, p. 297. Academic Press.
- Nikaido, H., (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**: 593-656.
- Nikaido, H. & E. Y. Rosenberg, (1983) Porin channels in Escherichia coli: studies with liposomes reconstituted from purified proteins. *J Bacteriol* **153**: 241-252.
- Olah, G. A., S. Trakhanov, J. Trehwella & F. A. Quioco, (1993) Leucine/isoleucine/valine-binding protein contracts upon binding of ligand. *J Biol Chem* **268**: 16241-16247.
- Oldham, M. L., D. Khare, F. A. Quioco, A. L. Davidson & J. Chen, (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**: 515-521.
- Ovchinnikov, Y. A., (1979) Physico-chemical basis of ion transport through biological membranes: ionophores and ion channels. *Eur J Biochem* **94**: 321-336.
- Padan, E. & S. Schuldiner, (1994) Molecular physiology of Na⁺/H⁺ antiporters, key transporters in circulation of Na⁺ and H⁺ in cells. *Biochim Biophys Acta* **1185**: 129-151.
- Pao, S. S., I. T. Paulsen & M. H. Saier, Jr., (1998) Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**: 1-34.
- Paulsen, I. T., L. Nguyen, M. K. Sliwinski, R. Rabus & M. H. Saier, Jr., (2000) Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. *J Mol Biol* **301**: 75-100.

- Paulsen, I. T., M. K. Sliwinski & M. H. Saier, Jr., (1998) Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *J Mol Biol* **277**: 573-592.
- Perkins, J. B., S. K. Guterman, C. L. Howitt, V. E. Williams, 2nd & J. Pero, (1990) Streptomyces genes involved in biosynthesis of the peptide antibiotic valinomycin. *J Bacteriol* **172**: 3108-3116.
- Pinkett, H. W., A. T. Lee, P. Lum, K. P. Locher & D. C. Rees, (2007) An inward-facing conformation of a putative metal-chelate-type ABC transporter. *Science* **315**: 373-377.
- Plantinga, T. H., C. Van Der Does, J. Badia, J. Aguilar, W. N. Konings & A. J. Driessen, (2004) Functional characterization of the Escherichia coli K-12 yiaMNO transport protein genes. *Mol Membr Biol* **21**: 51-57.
- Plate, C. A., (1979) Requirement for membrane potential in active transport of glutamine by Escherichia coli. *J Bacteriol* **137**: 221-225.
- Plumbridge, J. & E. Vimr, (1999) Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by Escherichia coli. *J Bacteriol* **181**: 47-54.
- Poolman, B., A. J. Driessen & W. N. Konings, (1987) Regulation of solute transport in streptococci by external and internal pH values. *Microbiol Rev* **51**: 498-508.
- Poolman, B. & W. N. Konings, (1993) Secondary solute transport in bacteria. *Biochim Biophys Acta* **1183**: 5-39.
- Poolman, B., D. Molenaar, E. J. Smid, T. Ubbink, T. Abee, P. P. Renault & W. N. Konings, (1991) Malolactic fermentation: electrogenic malate uptake and malate/lactate antiport generate metabolic energy. *J Bacteriol* **173**: 6030-6037.
- Poolman, B., Molenaar, D. and Konings, W.N., (1994) Diversity of transport mechanisms in bacteria. In: Biomembranes. pp.
- Pos, K. M., P. Dimroth & M. Bott, (1998) The Escherichia coli citrate carrier CitT: a member of a novel eubacterial transporter family related to the 2-oxoglutarate/malate translocator from spinach chloroplasts. *J Bacteriol* **180**: 4160-4165.

- Post, D. M., R. Mungur, B. W. Gibson & R. S. Munson, Jr., (2005) Identification of a novel sialic acid transporter in *Haemophilus ducreyi*. *Infect Immun* **73**: 6727-6735.
- Pourcher, T., S. Leclercq, G. Brandolin & G. Leblanc, (1995) Melibiose permease of *Escherichia coli*: large scale purification and evidence that H⁺, Na⁺, and Li⁺ sugar symport is catalyzed by a single polypeptide. *Biochemistry* **34**: 4412-4420.
- Prakash, S., G. Cooper, S. Singhi & M. H. Saier, Jr., (2003) The ion transporter superfamily. *Biochim Biophys Acta* **1618**: 79-92.
- Prossnitz, E., (1991) Determination of a region of the HisJ binding protein involved in the recognition of the membrane complex of the histidine transport system of *Salmonella typhimurium*. *J Biol Chem* **266**: 9673-9677.
- Puskarova, A., P. Ferienc, J. Kormanec, D. Homerova, A. Farewell & T. Nystrom, (2002) Regulation of yodA encoding a novel cadmium-induced protein in *Escherichia coli*. *Microbiology* **148**: 3801-3811.
- Quintero, M. J., M. L. Montesinos, A. Herrero & E. Flores, (2001) Identification of genes encoding amino acid permeases by inactivation of selected ORFs from the *Synechocystis* genomic sequence. *Genome Res* **11**: 2034-2040.
- Rabus, R., D. L. Jack, D. J. Kelly & M. H. Saier, Jr., (1999) TRAP transporters: an ancient family of extracytoplasmic solute-receptor-dependent secondary active transporters. *Microbiology* **145** (Pt 12): 3431-3445.
- Richarme, G., A. el Yaagoubi & M. Kohiyama, (1993) The MglA component of the binding protein-dependent galactose transport system of *Salmonella typhimurium* is a galactose-stimulated ATPase. *J Biol Chem* **268**: 9473-9477.
- Ringenberg, M. A., S. M. Steenbergen & E. R. Vimr, (2003) The first committed step in the biosynthesis of sialic acid by *Escherichia coli* K1 does not involve a phosphorylated N-acetylmannosamine intermediate. *Mol Microbiol* **50**: 961-975.
- Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J. L. Chou & et al., (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245**: 1066-1073.

- Rosenberg, T. & W. Wilbrandt, (1957) Uphill transport induced by counterflow. *J Gen Physiol* **41**: 289-296.
- Rucktooa, P., R. Antoine, J. Herrou, I. Huvent, C. Loch, F. Jacob-Dubuisson, V. Villeret & C. Bompard, (2007) Crystal structures of two Bordetella pertussis periplasmic receptors contribute to defining a novel pyroglutamic acid binding DctP subfamily. *J Mol Biol* **370**: 93-106.
- Sahin-Toth, M., R. L. Dunten, A. Gonzalez & H. R. Kaback, (1992) Functional interactions between putative intramembrane charged residues in the lactose permease of Escherichia coli. *Proc Natl Acad Sci U S A* **89**: 10547-10551.
- Sahin-Toth, M., J. le Coutre, D. Kharabi, G. le Maire, J. C. Lee & H. R. Kaback, (1999) Characterization of Glu126 and Arg144, two residues that are indispensable for substrate binding in the lactose permease of Escherichia coli. *Biochemistry* **38**: 813-819.
- Saidijam, M., K. E. Bettaney, G. Szakonyi, G. Psakis, K. Shibayama, S. Suzuki, J. L. Clough, V. Blessie, A. Abu-Bakr, S. Baumberg, J. Mueller, C. K. Hoyle, S. L. Palmer, P. Butaye, K. Walravens, S. G. Patching, J. O'Reilly, N. G. Rutherford, R. M. Bill, D. I. Roper, M. K. Phillips-Jones & P. J. Henderson, (2005) Active membrane transport and receptor proteins from bacteria. *Biochem Soc Trans* **33**: 867-872.
- Saidijam, M., G. Psakis, J. L. Clough, J. Mueller, S. Suzuki, C. J. Hoyle, S. L. Palmer, S. M. Morrison, M. K. Pos, R. C. Essenberg, M. C. Maiden, A. Abu-bakr, S. G. Baumberg, A. A. Neyfakh, J. K. Griffith, M. J. Stark, A. Ward, J. O'Reilly, N. G. Rutherford, M. K. Phillips-Jones & P. J. Henderson, (2003) Collection and characterisation of bacterial membrane proteins. *FEBS Lett* **555**: 170-175.
- Saier, M. H., R. N. Hvorup & R. D. Barabote, (2005) Evolution of the bacterial phosphotransferase system: from carriers and enzymes to group translocators. *Biochem Soc Trans* **33**: 220-224.
- Saier, M. H., Jr., (2000) Families of transmembrane sugar transport proteins. *Mol Microbiol* **35**: 699-710.
- Schaffner, W. & C. Weissmann, (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal Biochem* **56**: 502-514.

- Schirmer, T., T. A. Keller, Y. F. Wang & J. P. Rosenbusch, (1995) Structural basis for sugar translocation through maltoporin channels at 3.1 Å resolution. *Science* **267**: 512-514.
- Schuurman-Wolters, G. K. & B. Poolman, (2005) Substrate specificity and ionic regulation of GlnPQ from *Lactococcus lactis*. An ATP-binding cassette transporter with four extracytoplasmic substrate-binding domains. *J Biol Chem* **280**: 23785-23790.
- Screpanti, E., E. Padan, A. Rimón, H. Michel & C. Hunte, (2006) Crucial steps in the structure determination of the Na⁺/H⁺ antiporter NhaA in its native conformation. *J Mol Biol* **362**: 192-202.
- Sebulsky, M. T., B. H. Shilton, C. D. Speziali & D. E. Heinrichs, (2003) The role of FhuD2 in iron(III)-hydroxamate transport in *Staphylococcus aureus*. Demonstration that FhuD2 binds iron(III)-hydroxamates but with minimal conformational change and implication of mutations on transport. *J Biol Chem* **278**: 49890-49900.
- Severi, E., D. W. Hood & G. H. Thomas, (2007) Sialic acid utilization by bacterial pathogens. *Microbiology* **153**: 2817-2822.
- Severi, E., A. Muller, J. R. Potts, A. Leech, D. Williamson, K. S. Wilson & G. H. Thomas, (2008) Sialic acid mutarotation is catalyzed by the *Escherichia coli* beta-propeller protein YjhT. *J Biol Chem* **283**: 4841-4849.
- Severi, E., G. Randle, P. Kivlin, K. Whitfield, R. Young, R. Moxon, D. Kelly, D. Hood & G. H. Thomas, (2005) Sialic acid transport in *Haemophilus influenzae* is essential for lipopolysaccharide sialylation and serum resistance and is dependent on a novel tripartite ATP-independent periplasmic transporter. *Mol Microbiol* **58**: 1173-1185.
- Shakhnovich, E. A., S. J. King & J. N. Weiser, (2002) Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infect Immun* **70**: 7161-7164.
- Sharff, A. J., L. E. Rodseth, J. C. Spurlino & F. A. Quiocho, (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two

- domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **31**: 10657-10663.
- Shaw, J. G., M. J. Hamblin & D. J. Kelly, (1991) Purification, characterization and nucleotide sequence of the periplasmic C4-dicarboxylate-binding protein (DctP) from *Rhodobacter capsulatus*. *Mol Microbiol* **5**: 3055-3062.
- Shaw, J. G. a. K., D. J., (1991) Binding protein dependent transport of C4-dicarboxylates in *Rhodobacter capsulatus*. *Arch Microbiol* **155**: 6.
- Shuman, H. A., (1982) Active transport of maltose in *Escherichia coli* K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. *J Biol Chem* **257**: 5455-5461.
- Shuman, H. A. & T. J. Silhavy, (1981) Identification of the malK gene product. A peripheral membrane component of the *Escherichia coli* maltose transport system. *J Biol Chem* **256**: 560-562.
- Siebers, A., R. Kollmann, G. Dirkes & K. Altendorf, (1992) Rapid, high yield purification and characterization of the K(+)-translocating Kdp-ATPase from *Escherichia coli*. *J Biol Chem* **267**: 12717-12721.
- Sistrom, W. R., (1960) A requirement for sodium in the growth of *Rhodopseudomonas spheroides*. *J Gen Microbiol* **22**: 778-785.
- Smirnova, I., V. Kasho, J. Y. Choe, C. Altenbach, W. L. Hubbell & H. R. Kaback, (2007) Sugar binding induces an outward facing conformation of LacY. *Proc Natl Acad Sci U S A* **104**: 16504-16509.
- Smirnova, I. N. & H. R. Kaback, (2003) A mutation in the lactose permease of *Escherichia coli* that decreases conformational flexibility and increases protein stability. *Biochemistry* **42**: 3025-3031.
- Sohanpal, B. K., S. El-Labany, M. Lahooti, J. A. Plumbridge & I. C. Blomfield, (2004) Integrated regulatory responses of fimB to N-acetylneuraminic (sialic) acid and GlcNAc in *Escherichia coli* K-12. *Proc Natl Acad Sci U S A* **101**: 16322-16327.
- Sohanpal, B. K., S. Friar, J. Roobol, J. A. Plumbridge & I. C. Blomfield, (2007) Multiple co-regulatory elements and IHF are necessary for the control of fimB expression in response to sialic acid and N-acetylglucosamine in *Escherichia coli* K-12. *Mol Microbiol* **63**: 1223-1236.

- Sousa, M. C. & D. B. McKay, (2001) Structure of the universal stress protein of *Haemophilus influenzae*. *Structure (Camb)* **9**: 1135-1141.
- Springer, M. S., M. F. Goy & J. Adler, (1977) Sensory transduction in *Escherichia coli*: two complementary pathways of information processing that involve methylated proteins. *Proc Natl Acad Sci U S A* **74**: 3312-3316.
- Spurlino, J. C., G. Y. Lu & F. A. Quiocho, (1991) The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem* **266**: 5202-5219.
- 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.
- Steenbergen, S. M., C. A. Lichtensteiger, R. Caughlan, J. Garfinkle, T. E. Fuller & E. R. Vimr, (2005) Sialic Acid metabolism and systemic pasteurellosis. *Infect Immun* **73**: 1284-1294.
- Steinke, A., S. Grau, A. Davidson, E. Hofmann & M. Ehrmann, (2001) Characterization of transmembrane segments 3, 4, and 5 of MalF by mutational analysis. *J Bacteriol* **183**: 375-381.
- Studier, F. W., A. H. Rosenberg, J. J. Dunn & J. W. Dubendorff, (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**: 60-89.
- Surade, S., M. Klein, P. C. Stolt-Bergner, C. Muenke, A. Roy & H. Michel, (2006) Comparative analysis and "expression space" coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci* **15**: 2178-2189.
- Sweet, G., C. Gandor, R. Voegelé, N. Wittekindt, J. Beuerle, V. Truniger, E. C. Lin & W. Boos, (1990) Glycerol facilitator of *Escherichia coli*: cloning of glpF and identification of the glpF product. *J Bacteriol* **172**: 424-430.
- Sweet, G. D., J. M. Somers & W. W. Kay, (1979) Purification and properties of a citrate-binding transport component, the C protein of *Salmonella typhimurium*. *Can J Biochem* **57**: 710-715.
- Takahashi, H., E. Inagaki, C. Kuroishi & T. H. Tahirov, (2004) Structure of the *Thermus thermophilus* putative periplasmic glutamate/glutamine-binding protein. *Acta Crystallogr D Biol Crystallogr* **60**: 1846-1854.

- Takami, H., K. Nakasone, Y. Takaki, G. Maeno, R. Sasaki, N. Masui, F. Fuji, C. Hiram, Y. Nakamura, N. Ogasawara, S. Kuhara & K. Horikoshi, (2000) Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. *Nucleic Acids Res* **28**: 4317-4331.
- Tam, R. & M. H. Saier, Jr., (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* **57**: 320-346.
- Tame, J. R., G. N. Murshudov, E. J. Dodson, T. K. Neil, G. G. Dodson, C. F. Higgins & A. J. Wilkinson, (1994) The structural basis of sequence-independent peptide binding by OppA protein. *Science* **264**: 1578-1581.
- Teather, R. M., J. Bramhall, I. Riede, J. K. Wright, M. Furst, G. Aichele, U. Wilhelm & P. Overath, (1980) Lactose carrier protein of *Escherichia coli*. Structure and expression of plasmids carrying the Y gene of the lac operon. *Eur J Biochem* **108**: 223-231.
- Thomas, G. H., S. T., L.-K. M.R., A. Leech & D. Kelly, (2006) Novel ligands for the extracellular solute receptors of two bacterial TRAP transporters. *Microbiology* **152**: 187-198.
- Tolner, B., T. Ubbink-Kok, B. Poolman & W. N. Konings, (1995a) Cation-selectivity of the L-glutamate transporters of *Escherichia coli*, *Bacillus stearothermophilus* and *Bacillus caldotenax*: dependence on the environment in which the proteins are expressed. *Mol Microbiol* **18**: 123-133.
- Tolner, B., T. Ubbink-Kok, B. Poolman & W. N. Konings, (1995b) Characterization of the proton/glutamate symport protein of *Bacillus subtilis* and its functional expression in *Escherichia coli*. *J Bacteriol* **177**: 2863-2869.
- Trakhanov, S., N. K. Vyas, H. Luecke, D. M. Kristensen, J. Ma & F. A. Quiocho, (2005) Ligand-free and -bound structures of the binding protein (LivJ) of the *Escherichia coli* ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation and ligand specificity. *Biochemistry* **44**: 6597-6608.
- Treptow, N. A. & H. A. Shuman, (1985) Genetic evidence for substrate and periplasmic-binding-protein recognition by the MalF and MalG proteins,

- cytoplasmic membrane components of the *Escherichia coli* maltose transport system. *J Bacteriol* **163**: 654-660.
- Tsuchiya, T. & S. Shinoda, (1985) Respiration-driven Na⁺ pump and Na⁺ circulation in *Vibrio parahaemolyticus*. *J Bacteriol* **162**: 794-798.
- Tynkkynen, S., G. Buist, E. Kunji, J. Kok, B. Poolman, G. Venema & A. Haandrikman, (1993) Genetic and biochemical characterization of the oligopeptide transport system of *Lactococcus lactis*. *J Bacteriol* **175**: 7523-7532.
- Ullmann, R., R. Gross, J. Simon, G. Uden & A. Kroger, (2000) Transport of C(4)-dicarboxylates in *Wolinella succinogenes*. *J Bacteriol* **182**: 5757-5764.
- van der Rest, M. E., D. Molenaar & W. N. Konings, (1992) Mechanism of Na⁽⁺⁾-dependent citrate transport in *Klebsiella pneumoniae*. *J Bacteriol* **174**: 4893-4898.
- Varki, A., (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**: 97-130.
- Varki, A., (1997) Sialic acids as ligands in recognition phenomena. *FASEB J* **11**: 248-255.
- Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C.

- Xiao, C. Yan, et al., (2001) The sequence of the human genome. *Science* **291**: 1304-1351.
- Viitanen, P., M. L. Garcia & H. R. Kaback, (1984) Purified reconstituted lac carrier protein from *Escherichia coli* is fully functional. *Proc Natl Acad Sci U S A* **81**: 1629-1633.
- Vimr, E., C. Lichtensteiger & S. Steenbergen, (2000) Sialic acid metabolism's dual function in *Haemophilus influenzae*. *Mol Microbiol* **36**: 1113-1123.
- Vimr, E. R., (1994) Microbial sialidases: does bigger always mean better? *Trends Microbiol* **2**: 271-277.
- Vimr, E. R., K. A. Kalivoda, E. L. Deszo & S. M. Steenbergen, (2004) Diversity of microbial sialic acid metabolism. *Microbiol Mol Biol Rev* **68**: 132-153.
- Vimr, E. R. & F. A. Troy, (1985a) Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli*. *J Bacteriol* **164**: 845-853.
- Vimr, E. R. & F. A. Troy, (1985b) Regulation of sialic acid metabolism in *Escherichia coli*: role of N-acylneuraminate pyruvate-lyase. *J Bacteriol* **164**: 854-860.
- Vyas, N. K., M. N. Vyas & F. A. Quioco, (1991) Comparison of the periplasmic receptors for L-arabinose, D-glucose/D-galactose, and D-ribose. Structural and Functional Similarity. *J Biol Chem* **266**: 5226-5237.
- Wagner, S., L. Baars, A. J. Ytterberg, A. Klussmeier, C. S. Wagner, O. Nord, P. A. Nygren, K. J. van Wijk & J. W. de Gier, (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol Cell Proteomics* **6**: 1527-1550.
- Walker, J. E., M. Saraste, M. J. Runswick & N. J. Gay, (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**: 945-951.
- Walmsley, A. R., J. G. Shaw & D. J. Kelly, (1992) The mechanism of ligand binding to the periplasmic C4-dicarboxylate binding protein (DctP) from *Rhodobacter capsulatus*. *J Biol Chem* **267**: 8064-8072.
- Walter, C., K. Honer zu Bentrup & E. Schneider, (1992) Large scale purification, nucleotide binding properties, and ATPase activity of the MalK subunit of

- Salmonella typhimurium maltose transport complex. *J Biol Chem* **267**: 8863-8869.
- Wandersman, C., M. Schwartz & T. Ferenci, (1979) Escherichia coli mutants impaired in maltodextrin transport. *J Bacteriol* **140**: 1-13.
- Ward, A. e. a., (2000) *Membrane transport: A practical approach*. Oxford University Press, Oxford.
- Weinglass, A. B., J. P. Whitelegge, Y. Hu, G. E. Verner, K. F. Faull & H. R. Kaback, (2003) Elucidation of substrate binding interactions in a membrane transport protein by mass spectrometry. *EMBO J* **22**: 1467-1477.
- Widdas, W. F., (1952) Inability of diffusion to account for placental glucose transfer in the sheep and consideration of the kinetics of a possible carrier transfer. *J Physiol* **118**: 23-39.
- Widenhorn, K. A., W. Boos, J. M. Somers & W. W. Kay, (1988) Cloning and properties of the Salmonella typhimurium tricarboxylate transport operon in Escherichia coli. *J Bacteriol* **170**: 883-888.
- Wiener, M. C., (2004) A pedestrian guide to membrane protein crystallization. *Methods* **34**: 364-372.
- Wilkinson, A. J. a. V., K.H.G., (2003) Crystal structures of periplasmic solute-binding proteins in ABC transport complexes illuminate their function. In: ABC transporters. From bacteria to man. I. B. Holland, Cole, S.P.C., Kuchler, K. and Higgins, C.F. (ed). Academic Press, pp.
- Winnen, B., R. N. Hvorup & M. H. Saier, Jr., (2003) The tripartite tricarboxylate transporter (TTT) family. *Res Microbiol* **154**: 457-465.
- Witholt, B. & M. Boekhout, (1978) The effect of osmotic shock on the accessibility of the murein layer of exponentially growing Escherichia coli to lysozyme. *Biochim Biophys Acta* **508**: 296-305.
- Wong, P. T. & T. H. Wilson, (1970) Counterflow of galactosides in Escherichia coli. *Biochim Biophys Acta* **196**: 336-350.
- Wyborn, N. R., J. Alderson, S. C. Andrews & D. J. Kelly, (2001) Topological analysis of DctQ, the small integral membrane protein of the C4-dicarboxylate TRAP transporter of Rhodobacter capsulatus. *FEMS Microbiol Lett* **194**: 13-17.

- Yao, N., S. Trakhanov & F. A. Quioco, (1994) Refined 1.89-Å structure of the histidine-binding protein complexed with histidine and its relationship with many other active transport/chemosensory proteins. *Biochemistry* **33**: 4769-4779.
- Yew, W. S. & J. A. Gerlt, (2002) Utilization of L-ascorbate by *Escherichia coli* K-12: assignments of functions to products of the *yjf-sga* and *yia-sgb* operons. *J Bacteriol* **184**: 302-306.
- Yuan, Y. R., S. Blecker, O. Martsinkevich, L. Millen, P. J. Thomas & J. F. Hunt, (2001) The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter. *J Biol Chem* **276**: 32313-32321.
- Zhou, Z., K. A. White, A. Polissi, C. Georgopoulos & C. R. Raetz, (1998) Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. *J Biol Chem* **273**: 12466-12475.
- Zientz, E., S. Six & G. Uden, (1996) Identification of a third secondary carrier (DcuC) for anaerobic C₄-dicarboxylate transport in *Escherichia coli*: roles of the three Dcu carriers in uptake and exchange. *J Bacteriol* **178**: 7241-7247.
- Zilberstein, D., S. Schuldiner & E. Padan, (1979) Proton electrochemical gradient in *Escherichia coli* cells and its relation to active transport of lactose. *Biochemistry* **18**: 669-673.