

Molecular and biochemical characterisation of  
SiaP as a sialic acid binding protein component  
of a TRAP transporter for sialic acid

Adam Paul Hopkins

A thesis submitted for the degree of  
Doctor of Philosophy

University of York

Department of Biology

July 2010

## Abstract

Sialic acid utilisation plays an important role in the growth and persistence of the obligate human mucosal pathogen *Haemophilus influenzae*, which causes respiratory tract infections, septicaemia and meningitis. Like many other bacteria, *H. influenzae* can use host-derived sialic acids as carbon, nitrogen and energy sources, but also as a terminal modification on the LPS to better evade the human immune system. *H. influenzae* takes up exogenous sialic acid via a tripartite ATP-independent periplasmic (TRAP) transporter, SiaPQM. This possesses an extracytoplasmic substrate binding protein (SBP), SiaP, which binds the substrate in the periplasm and delivers it to the specific membrane permease, SiaQM. SiaP contains two globular domains, which close around the substrate upon binding. Here, the mechanism of sialic acid binding by SiaP is investigated using site-directed mutagenesis of residues in the ligand binding site and analogues of sialic acid. These, and several mutations on the surface of SiaP, were investigated for their effect on transport by SiaPQM *in vitro*, using SiaQM reconstituted into proteoliposomes, and *in vivo*, using expression of *siaPQM* in an *E. coli* strain lacking its native sialic acid transporter, NanT.

It is demonstrated that stabilisation of the carboxylate group of sialic acid by the totally conserved Arginine-147 is important for high-affinity ligand binding, but is not essential for transport. Mutation of Arginine-150 to Aspartate abolishes the function of the transporter without affecting ligand binding, suggesting the existence of a critical interaction between the components of the transporter. The catabolism of the sialic acid analogues was also examined in *E. coli* expressing different sialic acid transporters. This indicates that a wide variety of sialic acid analogues are potential carbon sources in many pathogenic bacteria.

## List of Contents

Title page	1
Abstract	2
List of contents	3
List of figures	10
List of tables	13
Acknowledgements	14
Chapter 1. Introduction	15
1.1 The requirement for transporters	16
1.2 Active transport	19
1.3 ATP binding cassette (ABC) transporters	20
1.3.1 The <i>E. coli</i> ABC transporter for maltose as a model system	21
1.3.1.1 The substrate-binding protein (SBP)	21
1.3.1.2 The membrane-associated complex	23
1.3.1.3 The proposed mechanism of transport	24
1.3.2 The role of the PBP in transport	28
1.3.3 The mechanism of PBP domain closure	29
1.3.4 Exploiting SBPs from ABC transporters	32
1.3.4.1 Technological exploitation	33
1.3.4.2 Alternative uses of the PBP fold	35
1.4 Secondary active transport	37
1.4.1 The Major Facilitator Superfamily	37
1.4.2 The <i>E. coli</i> lactose permease as a model system	38
1.4.2.1 The structure of the lactose permease	38
1.4.2.2 Proposed mechanism of transport	40
1.4.3 Sodium-driven co-transport	40
1.5 SBP-dependent secondary active transport	43
1.5.1 Features and components of SBP-dependent secondary transporters	43
1.5.1.1 Co-ordination of the ligand by SBPs	46

1.5.1.2	Multimeric SBPs	48
1.5.2	The proposed mechanism of transport	50
1.5.2.1	The <i>R. capsulatus</i> C <sub>4</sub> -dicarboxylate transporter, DctPQM	50
1.5.2.2	The <i>H. influenzae</i> sialic acid transporter, SiaPQM	50
1.6	Sialic acids	52
1.6.1	Utilisation and catabolism of sialic acids	54
1.6.1.2	Use of sialic acid by <i>H. influenzae</i>	56
1.7	Aims of this investigation	57
Chapter 2. Materials and Methods		59
2.1	Media and antibiotics	60
2.1.1	Luria-Bertani broth	60
2.1.2	M9 minimal medium	60
2.1.3	Solid media	60
2.1.4	Antibiotics	60
2.2	Table of primers	60
2.3	Strains and plasmids	60
2.4	General cloning techniques	66
2.4.1	Agarose gel electrophoresis	66
2.4.2	Plasmid preparation	66
2.4.3	Polymerase Chain Reaction (PCR)	66
2.4.4	Site-directed mutagenesis PCR	66
2.4.4.1	Mutagenic primer design	66
2.4.4.2	Mutagenic PCR conditions	67
2.4.5	Megaprimered mutagenic PCR	67
2.4.5.1	Megaprimer primer design	67
2.4.5.2	Megaprimer PCR	67
2.4.5.3	Megaprimered mutagenic PCR	67
2.4.6	Preparation of DNA fragments and cloning	68
2.4.6.1	Restriction enzyme digestions	68
2.4.6.2	Dephosphorylation reaction conditions	68

2.4.6.3	Gel extraction	68
2.4.6.4	PCR clean up	68
2.4.6.5	Ligation of DNA	68
2.4.7	Transformations	69
2.4.7.1	Chemically competent <i>E. coli</i> DH5 $\alpha$ stock	69
2.4.7.2	Small volumes of chemically competent cells	69
2.4.7.3	Heat shock	69
2.4.7.4	Freeze-thaw competency	69
2.5	Growth of bacteria	70
2.5.1	Expression of periplasmic proteins from pET-based constructs	70
2.5.2	Expression of cytoplasmic proteins from pRSET-based constructs	70
2.5.3	Expression of membrane proteins from pBAD-based constructs	70
2.6	Preparation of bacterial extracts	71
2.6.1	Periplasmic fraction preparation	71
2.6.2	Cytoplasmic protein recovery	71
2.6.3	Preparation of cell membranes as total membrane vesicles	71
2.7	Protein purification techniques	71
2.7.1	Hydrophobic Interaction Chromatography (HIC)	71
2.7.2	Size Exclusion Chromatography (SEC)	72
2.7.3	Nickel-affinity chromatography	72
2.7.3.1	FPLC	72
2.7.3.2	Peristaltic pump	73
2.7.4	Nickel-Nitrilotriacetic acid (Ni-NTA) resin – FLIP	73
2.7.4.2	Purification of soluble protein	73
2.7.4.2	Purification of membrane protein	74
2.7.5	Reconstitution of membrane protein into proteoliposomes	74
2.8	Large volume fermentation for the production of SiaP-His <sub>6</sub> :A11N	74
2.8.1	Expression of protein from the pAH55 construct	74
2.8.2	Preparation of bacterial extracts	75
2.9	Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS PAGE)	75

2.10.1	Native conditions	75
2.9.1.1	Buffers and gel	75
2.9.1.2	Sample preparation	75
2.9.1.3	Running conditions	75
2.9.2	Denaturing conditions	75
2.9.2.1	Buffers and gel	76
2.9.2.2	Sample preparation	76
2.9.2.3	Running conditions	76
2.9.3	Staining/destaining	76
2.9.4	Western blotting	76
2.9.4.1	Buffers	76
2.9.4.2	Running conditions	77
2.9.4.3	Sample visualisation	77
2.10	Fluorescence spectroscopy	77
2.11	Isothermal Titration Calorimetry (ITC)	78
2.11.1	ITC protocol	78
2.11.2	ITC for SiaP-His <sub>6</sub> -F170W	78
2.11.3	ITC for higher $K_d$ values	79
2.12	Circular Dichroism (CD)	79
2.12.1	CD spectra determination	79
2.12.2	Thermal stability by CD	79
2.13	<i>In vivo</i> growth assays	79
2.13.1	Growth on solid medium	79
2.13.2	Liquid culture	79
2.13.3	Prototype incubated plate shaker – buffers etc and pre-growth	80
2.13.3.1	24-well plate set up	80
2.13.3.2	The prototype incubated plate shaker	80
2.13.3.3	Standard curve	80
2.14	ELISA-based assays for the detection of SiaP-His <sub>6</sub>	80
2.14.1	Target substrates	80
2.14.2	96-well plate assay	81

2.15	<sup>14</sup> C-radiolabelled sialic acid-based assays	81
2.15.1	Filter binding assay	81
2.15.2	In vitro <sup>14</sup> C-sialic acid transport assay	81
2.15.2.1	Preparation of SiaQM-containing proteoliposomes	81
2.15.2.2	Buffers for gradients	83
2.15.2.3	Standard assay	83
2.15.2.4	Altered <sup>14</sup> C-sialic acid concentration	83
2.15.2.5	Competition assay	83
2.16	Guanidine hydrochloride (GnHCl) denaturation for protein recycling	84
2.16.1	Buffers	84
2.16.2	In-column denaturation and refolding	84
Chapter 3. Examination of the contribution of individual amino acid residues to high-affinity sialic acid binding by SiaP		85
3.1	Strategy for site-directed mutagenesis of the <i>siaP</i> gene to alter sialic acid-binding properties of SiaP	86
3.2	Introduction of Arginine-147 mutants into native SiaP	86
3.3	SiaP-His <sub>6</sub> functions as SiaP both <i>in vitro</i> and <i>in vivo</i>	91
3.4	The introduction of tryptophan residues can improve the fluorescence signal change on ligand binding	96
3.5	Mutations of residues co-ordinating the ligand carboxylate group disrupt ligand binding	100
3.6	An aromatic residue is required at position 170 for high-affinity ligand binding	104
3.7	Co-ordination of the ligand carboxylate group is important for high-affinity binding	106
3.8	No ligand binding can be detected for the Arg-147 mutants <i>in vitro</i>	109
3.9	The Arg-147 residue is not essential for sialic acid transport <i>in vivo</i>	115
3.10	Summary	118

Chapter 4. Investigation and manipulation of the ligand binding properties of SiaP for application as a biosensor	120
4.1 Investigation of the biophysical properties of SiaP-His <sub>6</sub>	121
4.1.1 Ethanol decreases ligand binding affinity and promotes $\alpha$ -helix formation	122
4.1.2 Ligand binding affinity is affected by ionic strength and the presence of sodium ions	125
4.1.3 Decreasing temperature increases ligand binding affinity	129
4.1.4 Sialic acid-binding by SiaP is enthalpically driven and releases water	131
4.1.5 Optimised conditions for high affinity binding of Neu5Ac by SiaP	134
4.2 Detection of sialic acid	136
4.2.1 The Lateral Flow Device	136
4.2.2 Conjugated sialic acid as a basis for a Lateral Flow Device	136
4.2.3 SiaP is specific for monomeric sialic acid	138
4.2.4 An in-solution approach for sialic acid detection	143
4.3 Rational design of SiaP to modulate its sialic acid binding affinity	145
4.3.1 Binding site mutations designed to increase ligand affinity	145
4.3.2 Mutations designed to promote the closed conformation	147
4.4 Summary	151
 Chapter 5. Investigation of the effect of mutations in SiaP upon transport by reconstituted SiaPQM	 152
5.1 An <i>in vitro</i> assay for <sup>14</sup> C-Neu5Ac uptake by SiaP-SiaQM	153
5.2 The Arg-147 mutants catalyse no uptake <i>in vitro</i>	155
5.3 Position 170 is important for ligand binding and transport	155
5.4 Binding site mutations reduce transport <i>in vitro</i> via their reduced occupancy	157
5.5 Surface mutations affect transport by different mechanisms	161
5.6 The Asn-150-Asp mutation abolishes transport without affecting ligand binding	163
5.6.1 SiaP-His <sub>6</sub> : N150D forms a transient, non-productive complex with SiaQM	166

5.6.2	Asn-150 forms a part of a network extending from the bound ligand to the surface of SiaP	171
5.7	Summary	177
Chapter 6. The utilisation of different sialic acid analogues		178
6.1	The ligand <i>N</i> -acetyl group is surrounded by a water network	179
6.2	<i>E. coli</i> can grow on different sialic acids as the sole carbon source using different sialic acid transporters	189
6.3	The catabolism of sialic acids requires relevant components of the sialometabolic pathway in <i>E. coli</i>	196
6.3.1	The catabolism of Neu5Gc requires the <i>nan</i> and <i>nag</i> genes and releases glycolate	196
6.3.2	The catabolism of KDN requires <i>nanA</i>	200
6.4	Summary	202
Chapter 7. Discussion		204
7.1	The binding of sialic acid by SiaP is sensitive to conditions and mutation	205
7.1.1	Can SiaP function without domain closure?	205
7.1.2	The sialic acid binding affinity of SiaP is sensitive to conditions and mutation of the binding site	209
7.2	The transport of sialic acid by SiaPQM is dependent upon ligand binding by SiaP and the correct interaction between the SiaP and SiaQM	213
7.3	Utilisation of sialic acids by bacteria	219
References		223

## List of Figures

<b>Figure 1.1:</b> The cell wall of (A) Gram-positive and (B) Gram-negative bacteria.	17
<b>Figure 1.2:</b> The major types of transporters in bacteria.	18
<b>Figure 1.3:</b> The <i>E. coli</i> maltose transport, MalEFGK <sub>2</sub> .	22
<b>Figure 1.4:</b> The proposed transport mechanism of MalEFGK <sub>2</sub> -like ABC importers.	25
<b>Figure 1.5:</b> Investigations of the mechanism PBP domain closure.	31
<b>Figure 1.6:</b> Exploited periplasmic binding proteins.	34
<b>Figure 1.7:</b> Alternative uses of periplasmic binding proteins.	36
<b>Figure 1.8:</b> Protein crystal structure of the <i>E. coli</i> lactose permease, LacY.	39
<b>Figure 1.9:</b> The structure and mechanism of sodium-coupled symport.	42
<b>Figure 1.10:</b> The structures of DctP-type SBPs.	45
<b>Figure 1.11:</b> The interactions between Neu5Ac and the binding site of SiaP.	47
<b>Figure 1.12:</b> Dimers of the DctP-type SBPs.	49
<b>Figure 1.13:</b> The proposed mechanism of TRAP transporters.	51
<b>Figure 1.14:</b> Nonulosonic acids found in bacteria.	53
<b>Figure 1.15:</b> Sialic acid utilisation by bacteria.	55
<b>Figure 1.16:</b> Sialyltransferase activities characterised in <i>H. influenzae</i> .	58
<b>Figure 2.1:</b> Standard curve for the relationship between OD <sub>650</sub> and G value.	82
<b>Figure 3.1:</b> The interactions between Neu5Ac and the binding site of SiaP.	87
<b>Figure 3.2:</b> The purification of native SiaP, SiaP:R147E and SiaP-R147K.	90
<b>Figure 3.3:</b> Purification and <i>in vitro</i> ligand binding analysis of SiaP-His <sub>6</sub> .	92
<b>Figure 3.4:</b> ITC analysis of 10 μM SiaP-His <sub>6</sub> in 50 mM Tris/HCl pH 8.0 at 37 °C.	94
<b>Figure 3.5:</b> <i>E. coli</i> BW25113 Δ <i>nanT</i> expressing <i>siaP<sub>QM</sub></i> and <i>siaP-His<sub>6</sub>-QM</i> .	95
<b>Figure 3.6:</b> Targets for tryptophan mutants in SiaP.	97
<b>Figure 3.7:</b> Expression and purification of SiaP reporter variants.	98
<b>Figure 3.8:</b> Fluorescence emission spectra of SiaP reporter variants	99
<b>Figure 3.9:</b> Expression and purification of SiaP-His <sub>6</sub> :R147 mutants.	101
<b>Figure 3.10:</b> Fluorescence titration of SiaP-His <sub>6</sub> with mutations as indicated.	102

<b>Figure 3.11:</b> Overlaid CD spectra of SiaP-His <sub>6</sub> and the SiaP-His <sub>6</sub> :R147 mutants.	105
<b>Figure 3.12:</b> Fluorescence titration of the SiaP-His <sub>6</sub> :F170 mutants.	107
<b>Figure 3.13:</b> Fluorescence spectra and titration of SiaP-His <sub>6</sub> :F170W.	108
<b>Figure 3.14:</b> ITC analysis of 13 μM SiaP-His <sub>6</sub> :R147K.	111
<b>Figure 3.15:</b> Native PAGE gel-shift assay with and without sialic acid.	113
<b>Figure 3.16:</b> Thermal denaturation SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :R147K.	114
<b>Figure 3.17:</b> <i>E. coli</i> Δ <i>nanT</i> expressing <i>siaP-His<sub>6</sub>-QM</i> with Arg-147 mutations.	116
<b>Figure 4.1:</b> CD spectra of SiaP-His <sub>6</sub> in the presence of ethanol.	124
<b>Figure 4.2:</b> Effect of pH and ionic strength on the K <sub>d</sub> of SiaP-His <sub>6</sub> :F170W.	128
<b>Figure 4.3:</b> Effect of temperature on the K <sub>d</sub> of SiaP-His <sub>6</sub> :F170W for sialic acid.	130
<b>Figure 4.4:</b> The thermodynamics of sialic acid binding by SiaP-His <sub>6</sub> .	132
<b>Figure 4.5:</b> ITC of SiaP-His <sub>6</sub> in sodium phosphate buffer pH 6.0 15 °C.	135
<b>Figure 4.6:</b> The principle of a lateral flow device using SiaP.	137
<b>Figure 4.7:</b> Detection of SiaP-His <sub>6</sub> with potential substrates.	139
<b>Figure 4.8:</b> Determination of glycan binding using a GlycoArray.	141
<b>Figure 4.9:</b> The interaction of SiaP-His <sub>6</sub> with modified sialic acid.	142
<b>Figure 4.10:</b> Construction of the FLIP-Neu5Ac nanosensor.	144
<b>Figure 4.11:</b> A family of binding site mutations designed to decrease K <sub>d</sub> .	146
<b>Figure 4.12:</b> A family of surface pair-wise targets designed to decrease K <sub>d</sub> .	149
<b>Figure 5.1:</b> <i>In vitro</i> <sup>14</sup> C-Neu5Ac uptake assay.	154
<b>Figure 5.2:</b> Uptake catalysed by SiaP-His <sub>6</sub> and the SiaP-His <sub>6</sub> :R147 mutants.	156
<b>Figure 5.3:</b> Uptake catalysed by the SiaP-His <sub>6</sub> :F170 mutants.	158
<b>Figure 5.4:</b> Uptake catalysed by the binding site mutants.	160
<b>Figure 5.5:</b> Uptake catalysed by the SiaP surface double mutants.	162
<b>Figure 5.6:</b> Uptake catalysed by SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :N150D.	164
<b>Figure 5.7:</b> ITC analysis of SiaP-His <sub>6</sub> :N150D.	165
<b>Figure 5.8:</b> Growth of <i>E. coli</i> Δ <i>nanT</i> expressing <i>siaP-His<sub>6</sub>:N150D-QM</i> .	167
<b>Figure 5.9:</b> Competition uptake assay between SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :N150D.	169
<b>Figure 5.10:</b> <i>E. coli</i> Δ <i>nanT</i> expressing both <i>siaP-His<sub>6</sub></i> and <i>siaP-His<sub>6</sub>:N150D</i> .	170

<b>Figure 5.11:</b> Representation of the Arg-50 – Asn-150 network.	172
<b>Figure 6.1:</b> The thermodynamic cycle of Neu5Ac and KDN binding by SiaP-His <sub>6</sub> .	184
<b>Figure 6.2:</b> Thermodynamics of SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :A11N binding Neu5Ac.	187
<b>Figure 6.3:</b> SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :A11N with bound Neu5Ac and KDN.	188
<b>Figure 6.4:</b> Growth curves of <i>E. coli</i> with Neu5Ac, Neu5Gc and KDN.	191
<b>Figure 6.5:</b> Growth of induced <i>E. coli</i> with Neu5Ac, Neu5Gc and KDN.	193
<b>Figure 6.6:</b> Growth of <i>E. coli</i> $\Delta$ <i>nanT</i> expressing <i>nanT</i> , <i>siaPQM</i> and <i>stm1128</i> .	194
<b>Figure 6.7:</b> Growth of <i>E. coli</i> sialometabolic deletion strains on Neu5Gc.	197
<b>Figure 6.8:</b> Growth of <i>E. coli</i> <i>glcE</i> and <i>yhcH</i> deletion strains on Neu5Gc.	199
<b>Figure 6.9:</b> Effect of sialometabolic deletions on growth of <i>E. coli</i> on KDN.	201
<b>Figure 7.1:</b> Representation of the mutation targets in the binding site of SiaP.	206
<b>Figure 7.2:</b> Binding sites SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :A11N with bound Neu5Ac.	210
<b>Figure 7.3:</b> Conjugated sialic acids could not be bound by SiaP.	212
<b>Figure 7.4:</b> The proposed sialic acid transport scheme of SiaPQM.	215
<b>Figure 7.5:</b> The residues involved in the Arg-50 – Asn-150 network.	218

## List of Tables

<b>Table 2.1:</b> Primer list	61
<b>Table 2.2:</b> Bacterial strains	63
<b>Table 2.3:</b> Plasmid list	64
<b>Table 3.1:</b> Target residues in SiaP and their chosen replacements.	88
<b>Table 3.2:</b> Sialic acid-binding affinities of the ligand-binding mutant proteins.	103
<b>Table 3.3:</b> Binding affinities of the ligand-binding mutants in the reporter variant.	110
<b>Table 3.4:</b> <i>E. coli</i> $\Delta nanT$ expressing <i>siaP-His<sub>6</sub>-QM</i> with Arg-147 mutations	117
<b>Table 4.1:</b> Binding affinities of SiaP-His <sub>6</sub> :F170W in the presence of ethanol.	123
<b>Table 4.2:</b> Secondary structure content of SiaP-His <sub>6</sub> in the presence of ethanol.	126
<b>Table 4.3:</b> Thermodynamic data from the ITC analysis of SiaP-His <sub>6</sub> .	133
<b>Table 4.4:</b> Sialic acid-binding affinities of the pair-wise mutant proteins.	150
<b>Table 5.1:</b> Thermodynamic data from the ITC analysis of SiaP-His <sub>6</sub> :N150D.	174
<b>Table 6.1:</b> The $K_d$ values of the binding pocket mutants for sialic acids	180
<b>Table 6.2:</b> The effect of the binding pocket mutants	180
<b>Table 6.3:</b> Thermodynamic data for Neu5Ac and KDN binding by SiaP-His <sub>6</sub> .	182
<b>Table 6.4:</b> ITC data for Neu5Ac binding by SiaP-His <sub>6</sub> and SiaP-His <sub>6</sub> :A11N.	185
<b>Table 6.5:</b> Growth of <i>E. coli</i> $\Delta nanT$ expressing <i>nanT</i> , <i>siaPQM</i> and <i>stm1128</i> .	195

### **Acknowledgements**

Firstly, I should thank my supervisor, Dr Gavin Thomas, for all of the help, direction and tolerance that he has offered. Thanks also to all of the members of the lab, past and present. I would also like to thank my CASE supervisor, Dr Mark Saw, and CASE sponsor, Authentix, for the amount of time that they have devoted to me and to whom part of this work is especially dedicated. I am indebted to my training committee, Profs Colin Kleanthous and Neil Bruce, who gave guidance and encouragement at regular intervals. I would like to make a specific mention of the time that I took up from Dr Andrew Leech and Berni Strongitharm in the Molecular Interactions lab and Marcus Fischer in the YSBL.

My sincere thanks go to Zoe, my best friend and wife, for her continuing love and support. I would also like to thank my parents, Christine and Paul, my sister, Jill, my aunt and my grandparents. Your encouragement and love are greatly appreciated; I wouldn't be where I am without each of you. Really guys, thanks a lot.

# Chapter One

## Introduction

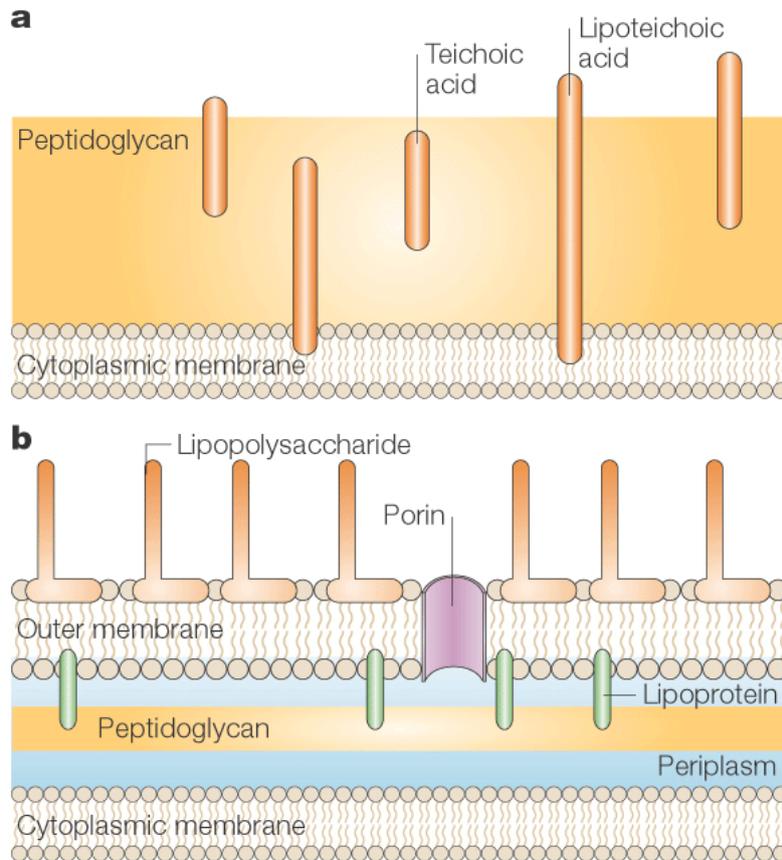
## 1.1 The requirement for transporters

Cells are encased by a phospholipid bilayer, separating the self from the non-self. In Gram-positive bacteria, this membrane is surrounded by a thick layer of protective peptidoglycan, together making up the cell envelope (Figure 1.1a). Gram-negative bacteria possess a second membrane above a thinner peptidoglycan layer, producing another distinct compartment for the cell called the periplasm (Figure 1.1b).

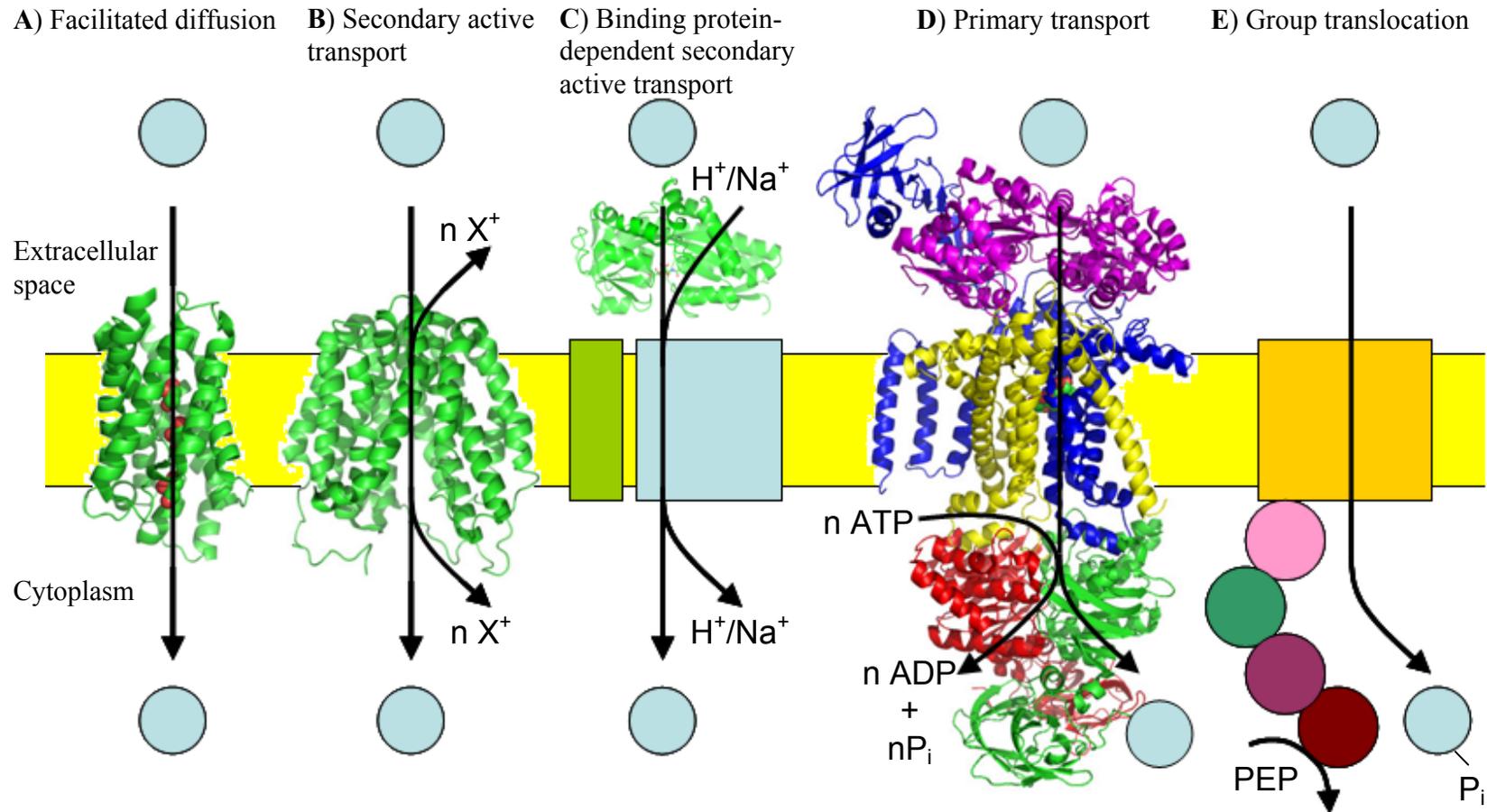
The phospholipid bilayer is made up of amphipathic phospholipids, which orient themselves with hydrophobic lipid groups in the centre of the bilayer and hydrophilic groups on the water-exposed surfaces (Wilkins *et al.*, 1971). This arrangement of hydrophobic and hydrophilic groups impedes the movement of polar and non-polar groups and so, to allow metabolic process to take place, proteinacious transporters have evolved to move exogenous and endogenous substrates across this barrier (Saier, 2000a). In the bilayer, the lipid groups exist in a liquid crystalline state, in which transmembrane proteins can diffuse laterally as both individual proteins and in large complexes (Singer & Nicolson, 1972, Steim *et al.*, 1969).

Various methods of membrane transport have evolved to suit environment and function. Conceptually, the least complex method is that of a channel facilitating the diffusion of a substrate to equalise its concentration on either side of the membrane (Transporter Classification TC1.A), as in the example of the tetrameric *Escherichia coli* glycerol facilitator, GlpF (TC1.A.8.1.1) (Figure 1.2a) (Sweet *et al.*, 1990).

It is also worth mentioning that different classes of transporters exist for the inward translocation of substrates through the outer membrane of Gram-negative bacteria. These include porins (non specific channels), specific channels and TonB-dependent receptors, which are powered transporters and are capable of active transport, accumulating their substrate against a concentration gradient (Nikaido, 2003).



**Figure 1.1:** The cell wall of (A) Gram-positive and (B) Gram-negative bacteria. In Gram-positive bacteria, the cell membrane is coated with a thick layer of protective peptidoglycan containing teichoic acid and lipoteichoic acid, which are critical for cell shape and survival. In Gram-negative bacteria, the peptidoglycan layer is much thinner and surrounded by the outer membrane, which are anchored together by embedded lipoproteins. The outer membrane is permeabilised by porins and coated by lipopolysaccharide, which is critical for cell survival and integrity. Taken from Cabeen and Jacobs-Wagner (2005).



**Figure 1.2:** The major types of transporters in bacteria (with substrate represented by blue circles). **A)** Substrate specific channels that equalise the concentration of the substrate across the membrane. Shown here is a monomer of the tetrameric *E. coli* glycerol facilitator, GlpF, with three molecules of bound glycerol as atom-coloured spheres (Fu *et al.*, 2000). **B)** Secondary active transporters couple translocation of the substrate with ion flow down an electrochemical gradient. The example given here is the *E. coli* lactose permease, LacY (Guan *et al.*, 2007). **C)** Binding protein-dependent secondary active transporters use two integral membrane proteins, the larger shows some similarity to secondary active transporters and the smaller (green) has an essential, but unknown function. Substrate is delivered to this complex by a periplasmic substrate binding protein (SBP). The example SBP shown here in green ribbons is SiaP (Johnston *et al.*, 2008). **D)** Primary transporters use energy directly to translocate the substrate. The example given here is a catalytic intermediate of the *E. coli* maltose transporter, MalEFGK<sub>2</sub> (Oldham *et al.*, 2007), where ATP hydrolysis drives substrate translocation across the membrane. For the ABC importer, the substrate is bound by a SBP and delivered to the permease. **E)** Group translocators. This group catalyse transport of the substrate at the same time as its phosphorylation. The example shown here is a representation of a general phosphotransferase system (Saier *et al.*, 2005).

## 1.2 Active transport

Active transport refers to systems that transport their substrate against a concentration gradient, powered at the expense of an energy source such as ATP, an electrochemical gradient or light. These active transporters can be grouped into three major types: secondary active (TC 2), primary active (TC 3) and group translocators (TC 4) (Figure 1.2b-f).

In secondary active transporters, transport of the solute is driven by its coupling with ion flow down an electrochemical gradient, such as  $H^+$  or  $Na^+$  (Saier, 2000b) (Figure 1.2b). The classical example of this is the *E. coli* lactose permease (TC 2.A.1.5.1) from the Major Facilitator superfamily (MFS; TC 2.A.1). This is a lactose: $H^+$  symporter where the translocation of extracellular protons into the cytoplasm drives the accumulation of intracellular lactose (Wong *et al.*, 1970, Wong & Wilson, 1970).

A distinct class of secondary transporters are the substrate-binding protein (SBP)-dependent secondary transporters (Figure 1.2c). The founding member of this class of transporter is the *Rhodobacter capsulatus*  $C_4$ -dicarboxylate tripartite ATP-independent periplasmic (TRAP) transporter, DctPQM (TC 2.A.56.1.1), which couples solute transport with ion flow down an electrochemical gradient (Forward *et al.*, 1997). The DctM integral membrane proteins of these are part of the Ion Transporter (IT) superfamily and share homology with the DcuC family of  $C_4$ -dicarboxylate secondary active transporters (TC 2.A.61) (Prakash *et al.*, 2003). This permease is associated with an essential, smaller integral membrane protein of unknown function and a periplasmic SBP (Forward *et al.*, 1997, Rabus *et al.*, 1999).

Primary active transporters use an energy source, such as ATP hydrolysis, to drive substrate translocation across the membrane. In the case of ABC importers, the substrate is delivered to the permease by a substrate binding protein (SBP; also known as PBP or ESR), which binds the substrate with high specificity and affinity (Figure 1.2d) (Ames & Nikaido, 1978). ATP is hydrolysed by ATP-binding cassettes, hence the superfamily

name of ABC transporters. The classical example of this family is the *E. coli* maltose transporter (TC3.A.1.1.1) from the carbohydrate uptake transporter-1 (CUT-1) family.

The final group of transporters considered here are the Group Translocators, such as the bacterial phosphotransferase system (PTS; TC 4.A) (Figure 1.2f) (Saier *et al.*, 2005). In these, translocation of the substrate is concomitant with its phosphorylation. The phosphate group is transferred through cytosolic members of the complex from phosphoenolpyruvate. A major role for these transporters is the uptake of sugars, generating cytosolic phosphorylated sugars.

Of these groups, ABC transporters will be discussed in detail in terms of the SBP and the mechanism of transport. There have recently been critical developments in this field in describing the mechanism of this SBP-dependent transporter. Secondary active transporters will also be discussed, in line with the TRAP transporters being a class of this transporter type.

### **1.3 ATP binding cassette (ABC) transporters**

ABC transporters make up a large proportion of all transporters found in nature and are a common bacterial import system. The *E. coli* genome contains 65 ABC transport systems, taking up over 5% of the total genome (Moussatova *et al.*, 2008). Fifteen of these lack an apparent SBP and so are presumed to be ATP-driven export systems, leaving the remaining 50 as ABC uptake systems. Their use of ATP hydrolysis to drive transport has been shown in many systems both *in vivo* and *in vitro* by ATP depletion or removal and analogue replacement (Ames *et al.*, 1989, Dean *et al.*, 1989a, Dean *et al.*, 1989b, Prossnitz *et al.*, 1989). Substrate-specific SBPs, previously discovered by their role in chemotaxis (Adler *et al.*, 1973, Hazelbauer & Adler, 1971), were shown to be essential for transport by deletion studies, removal of the periplasm by osmotic shock and transport in membrane vesicles (Ames & Lever, 1970, Wandersman *et al.*, 1979, Ames *et al.*, 1989, Dean *et al.*, 1989a, Dean *et al.*, 1989b).

### 1.3.1 The *E. coli* ABC transporter for maltose as a model system

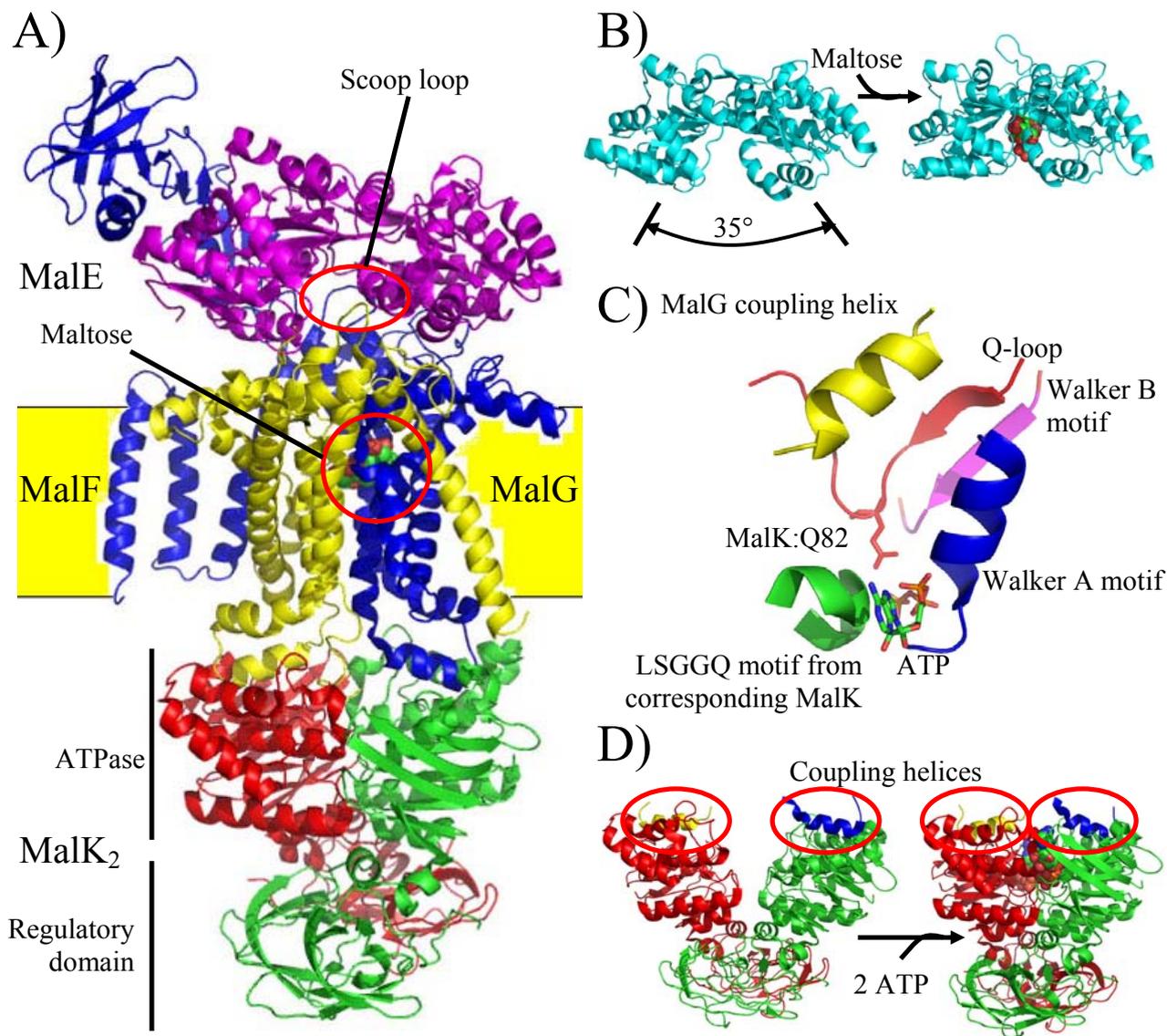
By the end of the 1970's, it had been shown that maltose transport in *E. coli* was dependent upon a small number of genes including *malE* that encodes a soluble periplasmic maltose chemotaxis protein, which was purified to homogeneity and found to bind to malto-oligosaccharides (Hofnung, 1974, Hofnung *et al.*, 1974, Kellermann & Szmecman, 1974, Wandersman *et al.*, 1979). Following this, the maltose transport system was described in detail using *E. coli* membrane vesicles (Dean *et al.*, 1989a, Dean *et al.*, 1989b). This showed conclusively that ATP-driven transport was dependent on MalE, whether this was supplied in the reaction mix or tethered to the membrane by an un-cleavable signal sequence.

Like all other ABC transporters, the *E. coli* maltose transporter, MalEFGK<sub>2</sub>, is made up of a SBP (maltose binding protein; MBP, MalE), two integral membrane domains (MalF and MalG) and an ATPase dimer (MalK<sub>2</sub>) (Figure 1.3a). The ligand-bound MBP interacts with the periplasmic surface of MalFG to deliver the substrate for transport through MalFG (Daus *et al.*, 2007a, Oldham *et al.*, 2007). The ATPase dimer interacts with the cytoplasmic surface of MalFG and, on binding of two molecules of ATP, the dimer and then reopens following ATP hydrolysis (Daus *et al.*, 2007b, Lu *et al.*, 2005).

#### 1.3.1.1 The substrate-binding protein (SBP)

Periplasmic substrate-binding proteins are soluble proteins made up of two  $\alpha/\beta$  globular domains connected by a flexible hinge region (Davidson *et al.*, 2008, Quioco, 1990, Quioco & Ledvina, 1996). SBPs from different systems have low sequence homology but highly similar tertiary structure (Quioco & Ledvina, 1996). Ligand binding occurs by closure of the two domains in a Venus fly trap-like mechanism (Kellermann & Szmecman, 1974, Mao *et al.*, 1982, Newcomer *et al.*, 1981, Sharff *et al.*, 1992, Spurlino *et al.*, 1991).

Previously, SBPs were classified based upon their secondary structure arrangement (Davidson & Chen, 2004). Class I is believed to have arisen from the fusion of a substrate-binding CheY-like protein dimer. From this ancestral class I SBP, a helix swap



**Figure 1.3:** The *E. coli* maltose transport, MalEFGK<sub>2</sub> (Oldham *et al.*, 2007). **A)** Protein crystal structure of a catalytic intermediate of the entire complex. MalE (magenta) delivers maltose to the TMDs, MalFG (blue and yellow). In this intermediate, maltose (atom coloured spheres) is held in a binding site in MalF within the membrane. The motion of the TMDs that allow transport of maltose (atom coloured spheres) is driven by ATP hydrolysis by the ATPase dimer, MalK<sub>2</sub> (red and green). **B)** MalE is made up of two domains that close around the ligand. This motion is made up of a 35° hinge bend and an 8° rotation around the hinge. **C)** One of the ATP binding sites. ATP (atom coloured sticks) is held between the Walker A motif (blue) of MalK and the LSGGQ motif (green) of the opposite MalK with a glutamine residue (red sticks) from the Q-loop (red). The Q-loop has extensive contacts with the TMD coupling helix (yellow) and is proposed to signal between the domains. The Walker B motif is shown in pink. **D)** The NBDs close around two ATP molecules. It is proposed that this motion is transmitted to the motion of the TMDs via the coupling helices in MalFG (circled).

(“domain dislocation”) occurred between the domains to give the ancestral class II SBP (Fukami-Kobayashi *et al.*, 1999). Class III SBPs have a more rigid hinge structure and undergo a smaller conformational change upon ligand binding (Borths *et al.*, 2002, Davidson & Chen, 2004, Karpowich *et al.*, 2003).

Recently, a new scheme has been proposed that is based upon the high resolution structural data available for 107 SBPs (Berntsson *et al.*, 2010). In this, SBPs were reclassified based upon distinct, defining structural features into six clusters, A-F, all of which can be associated with ABC importers except for cluster E, which is specific to TRAP transporters.

In the maltose binding protein (MBP), the extent of domain opening between the two conformations is 35° hinge bending and 8° rotation (Figure 1.3b) (Sharff *et al.*, 1992). This large change in the presented surface of the protein between these two conformations is the basis for the proposed mechanism for discrimination of the ligand bound versus unbound PBP.

#### **1.3.1.2 The membrane-associated complex**

The membrane-associated complex of an ABC transporter is made up of two NBDs and two TMDs, with a total of 10-20 TMHs (Figure 1.3a) (Hollenstein *et al.*, 2007). The TMDs can form a homo- or heterodimer and can often be fused to a NBD monomer, resulting in a half-transporter that can form a homo- or heterodimeric whole transporter (Dawson *et al.*, 2007). MalEFGK<sub>2</sub> is made up of a heterodimer of the transmembrane proteins MalF and MalG associated with a NBD homodimer, MalK<sub>2</sub>.

MalF and MalG are made up of 8 and 6 TMHs, respectively, which form corresponding crescent shapes around a central opening (Oldham *et al.*, 2007). At the base of each TMD, there is a conserved loop containing the Gln-Ala-Ala (EAA) motif, which is critical in the interaction with the NBD (Bluschke *et al.*, 2007, Daus *et al.*, 2007b). In the MalEFGK<sub>2</sub> crystal structure (Figure 1.3a), it can be seen that each of these loops fit into a

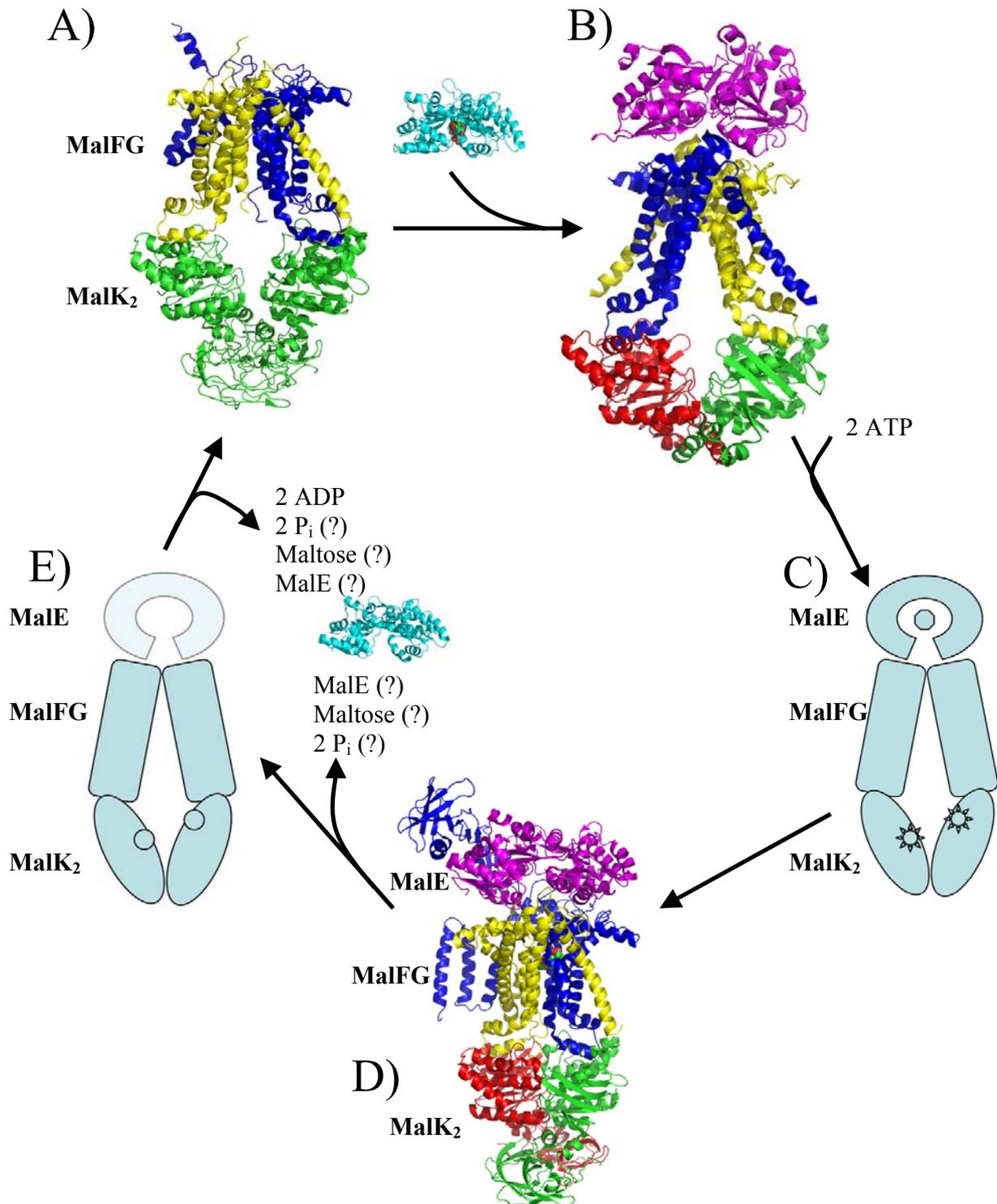
groove between lobes 1 and 2 of each MalK monomer and interact with a conserved region, called the Q-loop, which also contributes Gln-82 to the co-ordination of ATP (Chen *et al.*, 2003). As shown in Figure 1.3d, ATP is bound between the Walker A and Walker B motifs of the NBD and the LSGGQ motif of the opposite NBD (Chen *et al.*, 2003, Oldham *et al.*, 2007). It has been proposed that this arrangement with the Q-loop allows signalling and the application of mechanical force between the TMDs and the NBDs (Chen *et al.*, 2003).

In the protein crystal structure of the MalK<sub>2</sub> dimer in isolation, Chen *et al.* (2003) found that in the absence of ATP, the dimer could adopt a fully open conformation (Figure 1.3d) and a semi-closed conformation, which were both suggested to be mechanistically relevant to transport.

MalF is unusual in that it has an extended periplasmic loop that reaches around MalE in the protein crystal structure. This interaction with MalE is apparently stable, since this loop, when expressed alone, can associate with MalE in solution (Jacso *et al.*, 2009).

### **1.3.1.3 The proposed mechanism of transport**

The recent protein crystal structure of MalFGK<sub>2</sub> in its resting state shows MalFG in a periplasm-closed, cytoplasm-open conformation with the NBDs in a conformation between the fully open and semi-open forms (Figure 1.4a) (Chen *et al.*, 2003, Khare *et al.*, 2009). The periplasmic loop of MalF is not represented in the structure due to its disordered state. This structure is very similar to those of *E. coli* MetN<sub>2</sub>I<sub>2</sub> and ModB<sub>2</sub>C<sub>2</sub> from *Methanosarcina acetivorans* and *Archaeoglobus fulgidus*, in that the core TMDs adopt a curved shape (Gerber *et al.*, 2008, Hollenstein *et al.*, 2007, Jones *et al.*, 2009, Kadaba *et al.*, 2008). It is worth mentioning here that the NBD extended C-termini of *E. coli* MetN<sub>2</sub>I<sub>2</sub> and *M. acetivorans* ModB<sub>2</sub>C<sub>2</sub> have an extended C-terminus that are involved in a transinhibition regulatory mechanism, where cytoplasmic substrate binds to this regulatory domain and result in the NBDs moving further apart (Gerber *et al.*, 2008, Kadaba *et al.*, 2008).



**Figure 1.4:** The proposed transport mechanism of MalEFGK<sub>2</sub>-like ABC importers. The binding protein is shown in magenta, the TMDs in blue and yellow, the NBDs in green or green and red. **A)** The resting state of MalFGK<sub>2</sub>. MalFG is cytoplasm-open, periplasm-closed and the NBDs are in a conformation between their fully open and semi-open forms. **B)** The interaction of the PBP with the TMDs is represented by the protein crystal structure of *A. fulgidus* molybdate transporter, ModAB<sub>2</sub>C<sub>2</sub>. **C)** The proposed pre-hydrolytic transition state complex of maltose-MalEFGK<sub>2</sub>-2ATP. **D)** The transport intermediate complex, stabilised by vanadate trapping. **E)** The proposed post-hydrolytic transition state with semi-open, ADP-bound NBDs and could still involve the binding protein. The dissociation of maltose, MalE, ADP and P<sub>i</sub> would lead back to the resting state of the transporter.

When considering the mechanism of transport, it would seem apparent that the MalEFGK<sub>2</sub>·2ATP complex must adopt an unstable conformation so that the NBD closure and the ATP hydrolysis can occur (Shilton, 2008). This state would be analogous to a transition state in enzyme catalysis. Since the closed MalE and the resting MalFGK<sub>2</sub> are in their lowest energy conformations, their interaction, along with ATP, must produce a relatively unstable complex to drive the reaction forwards.

The ligand-bound SBP is able to interact with MalFGK<sub>2</sub> and, in doing so, has been shown to cause a change in the closure of the NBDs to their semi-open positions (Austermuhle *et al.*, 2004, Hollenstein *et al.*, 2007, Orelle *et al.*, 2008). This conformation is represented in Figure 1.4b by the crystal structure of *A. fulgidus* ModAB<sub>2</sub>C<sub>2</sub>. Since ATP would be present in the nucleotide binding sites of MalK<sub>2</sub> (intracellular ATP concentration > K<sub>m</sub>), it has been proposed that, with ATP, the conformation that this complex would actually be found in is the ATP-loaded, SBP-absent conformation (Khare *et al.*, 2009, Orelle *et al.*, 2008, Shilton, 2008).

This complete closure of the NBDs and concerted closure of the cytoplasmic side of the TMDs is believed to be transmitted via the TMD coupling helices (Figure 1.3 ac) (Chen *et al.*, 2003, Khare *et al.*, 2009, Oldham *et al.*, 2007). The interaction between the coupling helix and the Q-loop is extensive, but mediated mainly via Van der Waal's interactions, and so can be satisfied throughout the mechanistic cycle.

The closure of the NBDs appears to result in a rigid body movement of MalFG to the cytoplasm-closed, periplasm-open conformation with a concerted opening of the associated MalE (Figure 1.4b-d) (Khare *et al.*, 2009, Oldham *et al.*, 2007). In this structure, maltose is held in a binding site within the TMDs and not by the SBP, where a loop from MalF partially occupies the SBP ligand binding site (Figure 1.3a) (Oldham *et al.*, 2007). This loop occupies the maltose binding site of MalE, preventing the ligand from maintaining an interaction with the open conformation. The concentration of maltose in this occluded cavity is approximately 250 mM and so the maltose is bound by the MalF binding site despite its low affinity for maltose, which was revealed by the low

affinity of SBP-independent mutants of MalFGK<sub>2</sub> (Khare *et al.*, 2009, Oldham *et al.*, 2007).

In this intermediate state, the NBDs are closed and so the spontaneous hydrolysis of ATP to ADP and P<sub>i</sub> would occur and is, in fact, necessary to reopen the NBDs in solution (Lu *et al.*, 2005). With ADP bound, this state is extremely unstable and so adopts the cytoplasm-open, periplasm-closed TMD conformation, followed by the dissociation of maltose, MalE, ADP and P<sub>i</sub> to give the resting state of MalFGK<sub>2</sub> (Figure 1.4d-a). However, it is possible that there is a post-hydrolytic transition state that can be adopted. In their work on the motion of the NBDs, Orelle *et al.* (2008) detected the presence of a semi-open NBD following ATP hydrolysis. This state in the transport cycle could follow the dissociation of P<sub>i</sub> and/or maltose. It is reasonable to assume that this transition state would be very unstable, particularly if the closed, unliganded MalE is still associated. Whatever the order, the dissociation of MalE and ADP may allow the NBDs to open further to allow the binding of ATP (Orelle *et al.*, 2008).

The stoichiometry of ATP hydrolysed per molecule of substrate transported is, in most cases, still up for debate. In the case of the *Lactococcus lactis* OppA ABC transporter (TC 3.A.1.5.10), Patzlaff *et al.* (2003) suggested that 2 ATP are hydrolysed per cycle of transport.

Recently, the protein crystal structure of the *E. coli* vitamin B<sub>12</sub> transporter, BtuC<sub>2</sub>D<sub>2</sub>F (TC 3.A.1.13.1), has been used to describe a different TMD-fold family and propose a different transport cycle (Hvorup *et al.*, 2007, Lewinson *et al.*, 2010). Along with the *H. influenzae* putative metal transporter HI1470/1471-HI1472 (TC 3.A.1.14.11), BtuC<sub>2</sub>D<sub>2</sub>F has a different transmembrane fold to the molybdate, methionine and maltose transporters, forming a narrower channel surrounded by densely-packed TMDs (Jones *et al.*, 2009). The most illustrative difference is that, for the MalEFGK<sub>2</sub>-like transporters, the highest affinity of the TMDs for the SBP occurs in the transport intermediate state, whereas this is the state with the least stable interaction in the BtuC<sub>2</sub>D<sub>2</sub>F-like transporters (Lewinson *et al.*, 2010). The resulting effect is that the ground state of the transporter is

SBP-bound with closed NBDs and a significant basal ATPase activity that is poorly stimulated by liganded SBP-binding. Quite how this mechanism is advantageous is unclear. The use of such quantities of ATP might be an acceptable loss under metal-limiting conditions and might be the cost of transporting high affinity-bound metal ions or large metal chelating groups from class III SBPs.

### **1.3.2 The role of the PBP in transport**

The use of a PBP in these systems confers not only high affinity and specificity in transport, but can also reduce non-productive consumption of ATP. Binding protein-independent mutants of MalFGK<sub>2</sub> (iMalFGK<sub>2</sub>) all show a high level of constitutive ATPase activity (Davidson *et al.*, 1992, Shuman, 1982). All of these mutants have been shown to bind maltose with a similar, low affinity, but allow growth at different rates, which are directly proportional to their rate of ATP hydrolysis (Davidson *et al.*, 1992).

The substrate specificity of the transporter is dictated by the specificity of the PBP. The *L. lactis* oligopeptide binding protein, OppA, has little specificity for the sequence of its ligands except for a preference for the presence of a leucine residue, *yet all* peptides that could be bound by OppA were transported by OppBCDF (Berntsson *et al.*, 2009, Doeven *et al.*, 2004). A similar effect was also seen where a sucrose-binding variant of MalE was found to interact with MalFGK<sub>2</sub> and stimulate ATPase activity (Gould & Shilton, 2010). The *Pseudomonas* ABC transporter for choline, betaine and carnitine has recently been reported to use separate, specific PBPs for betaine and carnitine that are encoded separately on the genome (Chen *et al.*, 2010).

It has been shown repeatedly that the opening and closing mechanism of the PBP is important in transport; mutations in the PBP that trap the domains closed cannot catalyse transport (Jacobson *et al.*, 1991, Sjoelund & Kaltashov, 2007, Zhang *et al.*, 1996) and a similar observation has been made for mutations that appeared to be on the interface of the PBP domains and the TMDs (Hor & Shuman, 1993, Shilton *et al.*, 1996).

Previously, it was believed that only the closed conformation of the PBP interacted with the membrane components. It was shown that this was not necessarily the case when Ames *et al.* (1996) showed that both the liganded and unliganded forms of the *Salmonella typhimurium* histidine-PBP could interact with the transporter with similar, micromolar affinity. However, the interaction was different for the open conformation, where one domain would bind preferentially and the ATPase would not be highly stimulated. Their results also suggested that the “de-liganded” PBP did not leave the complex following transport, resulting in transport depending on the diffusion of the substrate and not the slower-diffusing PBP. In this proposed model, the large excess of PBP in the periplasm (often greater than 1 mM) would store extracellular substrate and decrease the concentration of free substrate, so increasing the amount that entered the periplasm.

In some ABC import systems, binding proteins are found fused to the TMDs. Where multiple domains are found that bind the same ligand, two models have been proposed (van der Heide & Poolman, 2002). The first proposes that the second domain could deliver substrate to the first domain, which cycles through transport. In the second, the domains interact alternately so that one binds substrate whilst the other is associated to the TMDs. It could be that the faster diffusion of small molecules is being taken advantage of, as in the proposed model for the *S. typhimurium* histidine ABC transporter. If the large PBP concentration has an important role in transport, then this effect could be amplified in multiple binding domain systems, where the apparent concentration of the binding domain would approach 100 mM (van der Heide & Poolman, 2002).

### **1.3.3 The mechanism of PBP domain closure**

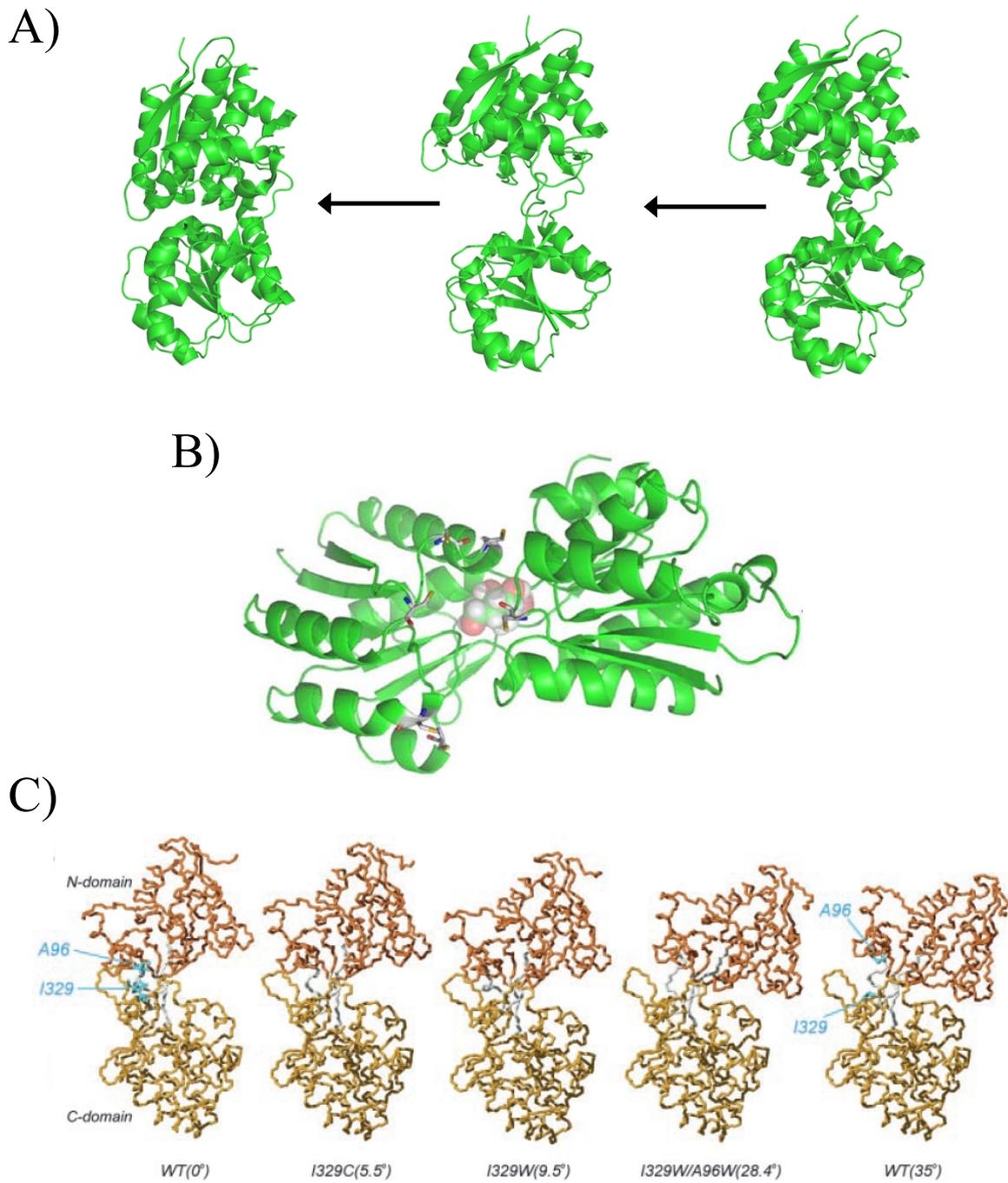
The ligand-induced conformational change of the PBP must occur following recognition of the ligand by the open conformation, or at least a semi-open conformation. The protein crystal structures of PBPs such as MalE show the networks of amino acid residues in their binding sites that recognise the cognate ligand. The binding site of the oligopeptide binding protein OppA manages to combine high affinity for its oligopeptide ligands with low specificity for their sequence inside a voluminous cavity by interacting with the

peptide backbone and a single isoleucine side chain (Berntsson *et al.*, 2009, Doeven *et al.*, 2004).

Some evidence for the mechanism of ligand recognition has come from protein crystal structures of PBPs bound to ligands that induce only partial or no closure of the domains. For MalE, these ligands include maltose, maltotriitol and maltotetraitol (Duan *et al.*, 2001, Spurlino *et al.*, 1991). Different conformations of the *E. coli* Leu/Ille/Val binding protein have also been compared (Trakhanov *et al.*, 2005). In this case, a more open form was also discovered along with the open and closed, all of which lay along the same trajectory between fully closed and fully open (Figure 1.5a). In both of these structures, the binding site contributed both hydrophobic and polar interactions to the ligand, arranged so that the majority of the non-polar interactions were made by aromatic residues donated from one domain. It was found that in the ligand-bound open conformations, the oligosaccharides were almost always bound to the aromatic residue-containing domain (Duan *et al.*, 2001, Duan & Quioco, 2002).

It is widely accepted that SBPs in solution exist in an equilibrium between the open and closed forms and that the presence of ligand stabilises the closed form. Using paramagnetic relaxation enhancement (PRE) NMR to probe the proximity of paramagnetic probes in the cleft of MalE, Tang *et al.* (2007) showed that, in solution, 5% of MalE can be found in the closed form in dynamic equilibrium with the open form. However, this does not always occur, as shown for the *E. coli* glutamine binding protein, all of which is found in a fully open conformation using similar methods (Bermejo *et al.*, 2010). While investigating the closing mechanism of MalE using molecular dynamics simulations, Stockner *et al.* (2005) suggested that a hydrogen bond network around Lys-15 and Glu-111 altered its contacts throughout the domain closure, creating a strong salt-bridge that restricted re-opening of the cleft.

Like MalE, the closure of the *E. coli* glucose/galactose binding protein (GGBP) is made up of a hinge-bending motion and a rotation, which was investigated by Careaga *et al.* (1995). The rotation component was investigated by disulphide trapping at six positions



**Figure 1.5:** Investigations of the mechanism PBP domain closure. **A)** The protein crystal structures of three conformations of the *E. coli* Leu/Ile/Val binding protein showing the closure of the domains. **B)** The positions of cysteine mutants in the *E. coli* glucose/galactose binding protein used in the investigation of domain rotation. **E)** The degree of closure caused by mutations in MalE. Taken from Millet *et al.* (2003).

around the lip of the cleft that covered 142° rotation (Figure 1.5b). They found that GGBP was sampling 120° rotation, which reduced to 36° on binding of glucose with a reduced rate of twisting. It was suggested that this could be a method for disrupting the interaction of the unliganded binding protein with the transport complex, although it is also possible that it is either an inherent property of a mobile system or is critical for the mechanism of ligand binding and release.

Since ligand binding by PBPs is a function of the equilibrium between the open and closed conformations, it would seem that destabilisation of the open form would move the equilibrium towards the closed, ligand-bound conformation. Various groups have approached this and have found differences in their results. In MalE, the first example of this was by Marvin & Hellinga (2001), who targeted a residue in a crevice on the opposite side of the hinge to the ligand binding site. This position was mutated to various residues including cysteine, to which modifying groups were attached. When the size of this side chain was increased, the affinity for the ligand was increased. These variants had caused a different degree of closure of the protein, where the increase in affinity was proportional to the degree of closure (Figure 1.5e) (Millet *et al.*, 2003). In the work of Telmer & Shilton (2003), the region on the opposite side of the hinge to the binding site was found to have interactions that stabilised the open conformation. When removed, this ‘balancing interface’ was disturbed and the affinity for the ligand was dramatically increased, mainly by a decrease in the off-rate of the bound ligand. In solution, these mutations were found to exist in the same proportion of the open and closed forms and their crystal structures showed a similar extent of opening. These results showed that the PBP bending around its hinge with stabilising interactions that can occur on both sides and that it was possible to alter this equilibrium via these or by blocking the extent of opening.

#### **1.3.4 Exploiting SBPs from ABC transporters**

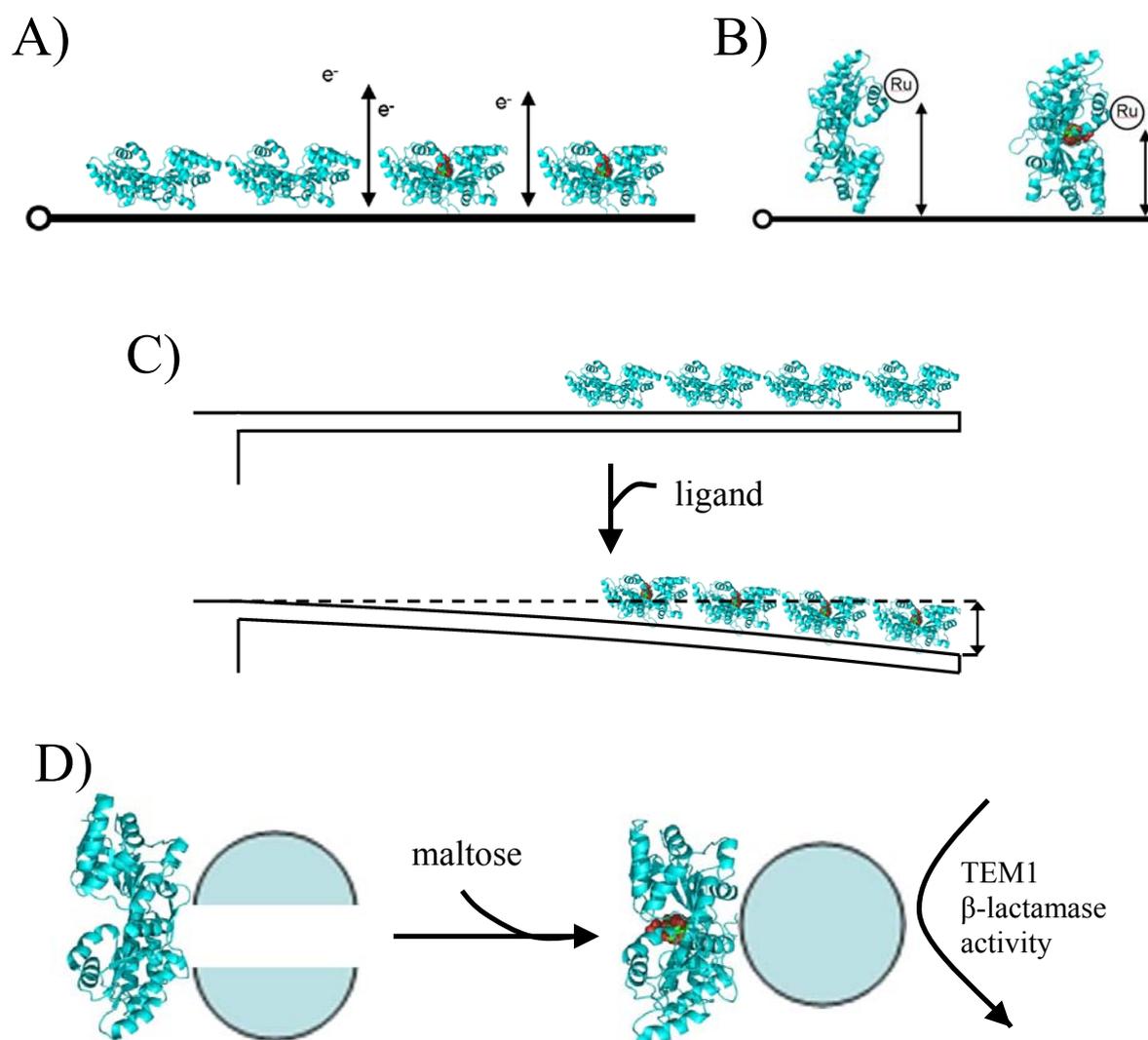
The mechanism of periplasmic binding proteins is one that is easily and frequently exploited for purposes other than transport. The most common is as a detector for its

substrate, such as in chemotaxis. This section is most easily separated into designed systems, exploited for technological gain, and natural systems, exploited for survival.

#### **1.3.4.1 Technological exploitation**

Most designed exploitation of the SBP fold is for the generation of biosensors for the SBP substrate, taking advantage of the high specificity for, and strong response to the presence of, the substrate. Almost all systems so far are based on designing an electronic interface to detect the conformational change in the PBP to create a sensor for a specific substrate. The major variance in this is the methods used to convert the conformational change into a signal. This ranges from the production of a fluorescence change to a more direct interaction with an electrode (Figure 1.6). Fluorescence-based methods involve the addition of a chemical fluorophore to the protein, usually around the cleft or the hinge so that the change in environment on domain closure causes a large change in fluorescence (de Lorimier *et al.*, 2006, Rizk *et al.*, 2006, Thomas *et al.*, 2006). The technology for the construction of genetically-encoded nanosensors has been advanced recently and involves the fusion of two green fluorescent proteins (GFPs) to either end of the PBP so that fluorescence resonance energy transfer (FRET) occurs between them (Deuschle *et al.*, 2005, Fehr *et al.*, 2002). This FRET signal is very sensitive to changes in the distance and orientation between the two GFPs and so can give a strong signal for the domain closure of the PBP. An advance from this is the circular permutation of the PBP (Okada *et al.*, 2009); this effectively rotates the secondary structure of a class II PBP so that each domain is made of a continuous polypeptide chain and there is just one strand to the hinge. This means that the termini, and so the GFPs, are at the maximum possible distance from each other in the open conformation, making the signal change much larger.

More direct interfaces with the conformational change usually involves surface immobilisation of the PBP to an electrode. From this, the closing of the PBP can be detected as altered electron transfer of the protein-coated electrodes or the motion of a conjugated heavy metal probe such as ruthenium (Figure 1.6ab) (Andreescu & Luck, 2008, Benson *et al.*, 2001). Cantilever-based sensors have been proposed that would take



**Figure 1.6:** Exploited periplasmic binding proteins (cyan ribbons). **A)** PBP-coated electronic probe. The binding of ligand (atom coloured spheres) affects the transfer of electrons around the electrode. The positions of cysteine mutants in the *E. coli* glucose/galactose binding protein used in the investigation of domain rotation. **B)** A heavy metal (e.g. ruthenium<sup>2+</sup>) cofactor is added to the protein, which is attached to an electrode in a specific orientation. The binding of ligand (atom coloured spheres) decreases the distance between the electrode and the cofactor, altering the potential of the electrode. **C)** Cantilever-based detection. The binding of heavy metal ions to the PBP would cause a deflection of the cantilever in an electric field. **D)** Closure of the PBP domains can be linked to the activation, or inactivation, of enzyme activity, such as TEM1  $\beta$ -lactamase.

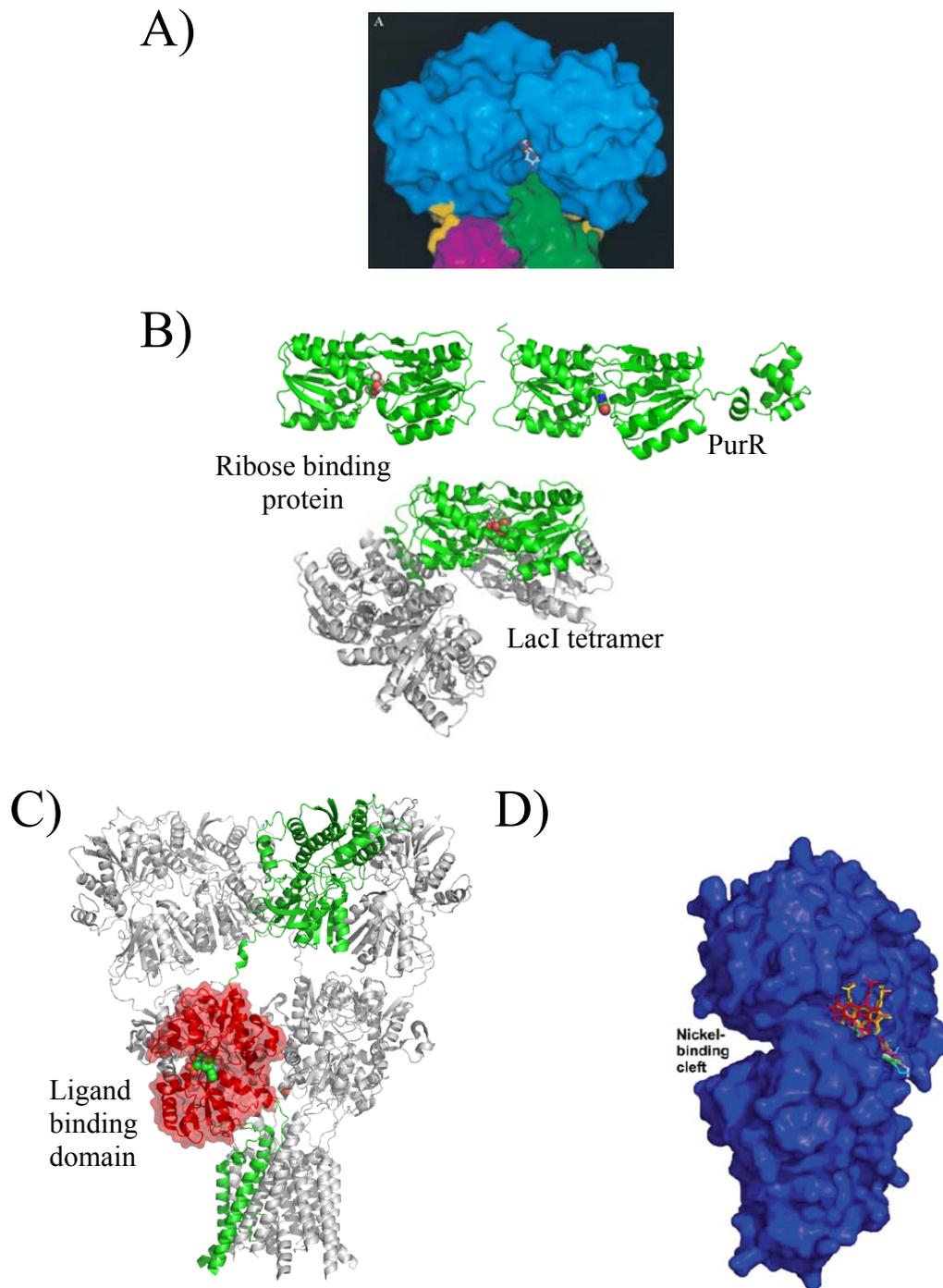
advantage of recent advances in scanning force microscopy, effectively producing a microbalance (Ziegler, 2004). These would be based on detecting the small movements of the cantilever following binding of the ligand and have, thus far, only been seriously suggested for heavy metal-binding proteins, which would be detected via their response to an electric field (Figure 1.6c).

Exploitation methods that link the PBP mechanism to another function, such as an enzyme, are exemplified by the work of Guntas *et al.* (2004), who generated a maltose-dependent TEM1  $\beta$ -lactamase (Figure 1.6d). This was achieved by random circular permutation and random insertion of the TEM1  $\beta$ -lactamase gene into *malE* and screening for maltose-sensitive  $\beta$ -lactamase activity.

#### **1.3.4.2 Alternative uses of the PBP fold**

The use of PBPs as detectors for the presence of their substrates is a common one in bacteria; PBPs were originally characterised from their role in chemotaxis (Hazelbauer & Adler, 1971, Adler *et al.*, 1973). For example, maltose-bound MalE interacts with the signal transducer tar and induces a chemotactic response (Figure 1.7a) (Manson & Kossmann, 1986, Zhang *et al.*, 1999). Similar exploitation also occurs intracellularly, for example the *E. coli* DNA-binding repressors for the metabolism of lactose and purines, LacI and PurR, which bind both their effector and their DNA target (Figure 1.7b) (Mowbray & Bjorkman, 1999).

In the ionotropic glutamate receptor, an extracellular PBP-like domain has been inserted that regulates the ion channel (Figure 1.7c) (Sobolevsky *et al.*, 2009). *Mycobacteria* have exploited a PBP in the synthesis of their cell wall (Marland *et al.*, 2006). This essential component binds phosphatidylinositol mannoside before its inclusion into the cell wall. In a similar manner, a PBP-like protein from *Pseudomonas aeruginosa* has gained cyclohexadienyl hydratase activity and is used in the synthesis of phenylalanine from chorismate in the periplasm (Tam & Saier, 1993).



**Figure 1.7:** Alternative uses of periplasmic binding proteins. **A)** MaleE (blue surface representation) interacts with the Tar chemotaxis signalling complex (green, magenta and yellow surface representation). Taken from Zhang *et al.* (1999). **B)** The PBP structure of the ribose binding protein is maintained in the transcription repressors PurR and LacI. **C)** Structure of the Rat ionotropic glutamate receptor (ribbons, one monomer in green) with the ligand binding domain highlighted as a red surface representation. **D)** The *E. coli* nickel binding protein (Blue surface representation) also binds periplasmic heme (different coloured cylinders). Taken from Shepherd *et al.* (2007).

In some cases, the original function of the PBP has been maintained, yet provides another function. *Synechocystis* PCC 6803 has two PBPs for  $\text{Fe}^{3+}$ , one of which, FutA2, is redundant in  $\text{Fe}^{3+}$  transport but involved in  $\text{Cu}^{2+}$  import (Waldron *et al.*, 2007). It has been shown that FutA2 sequesters excess  $\text{Fe}^{3+}$  in the periplasm, preventing its interaction with other metal binding proteins. Another example is *E. coli* nickel-PBP, NikA, which has been found to have a second, independent binding site for heme and is involved in its transport (Figure 1.7d) (Shepherd *et al.*, 2007).

#### 1.4 Secondary active transport

Secondary active transporters are common to all domains of life and can be found transporting almost any small molecule family (Saier, 2000b, Sobczak & Lolkema, 2005). Transport usually occurs at the cost of depleting an electrochemical gradient, usually  $\text{H}^+$ ,  $\text{Na}^+$  or  $\text{K}^+$ , which had been built up by transport of these ions at the cost of the cell's chemical energy, giving rise to the name of secondary transport (Harold & Kakinuma, 1985).

These transporters can catalyse transport by one of four different methods: Uniport (a single translocated substrate); Symport (co-transport of substrate and cations); Antiport (counter-transport of substrate and cations); or Solute-Solute exchange (exchange of one solute for another). It has also been reported that more complex combinations of these can occur including co- and counter-transport of multiple cations (Saier, 2000b). Secondary active transporters can be split into many families. The largest is the major facilitator superfamily (MFS; TC 2.A.1).

##### 1.4.1 The Major Facilitator Superfamily

The largest of these four groups of secondary transporters is the MFS, previously known as the uniporter-symporter-antiporter superfamily. It is itself split up into 29 different families (Pao *et al.*, 1998, Saier *et al.*, 1999). Secondary active transporters from the MFS are the most commonly used transporter for sugars and up to 8 MFS families have members involved in sugar transport (Kaback *et al.*, 2001, Saier, 2000b, Saier *et al.*,

1999). Two of these especially worth noting are the sugar porter (SP) family, which contains the largest number of members, and the oligosaccharide:H<sup>+</sup> symporters (OHS), which contains the *E. coli* lactose permease, LacY.

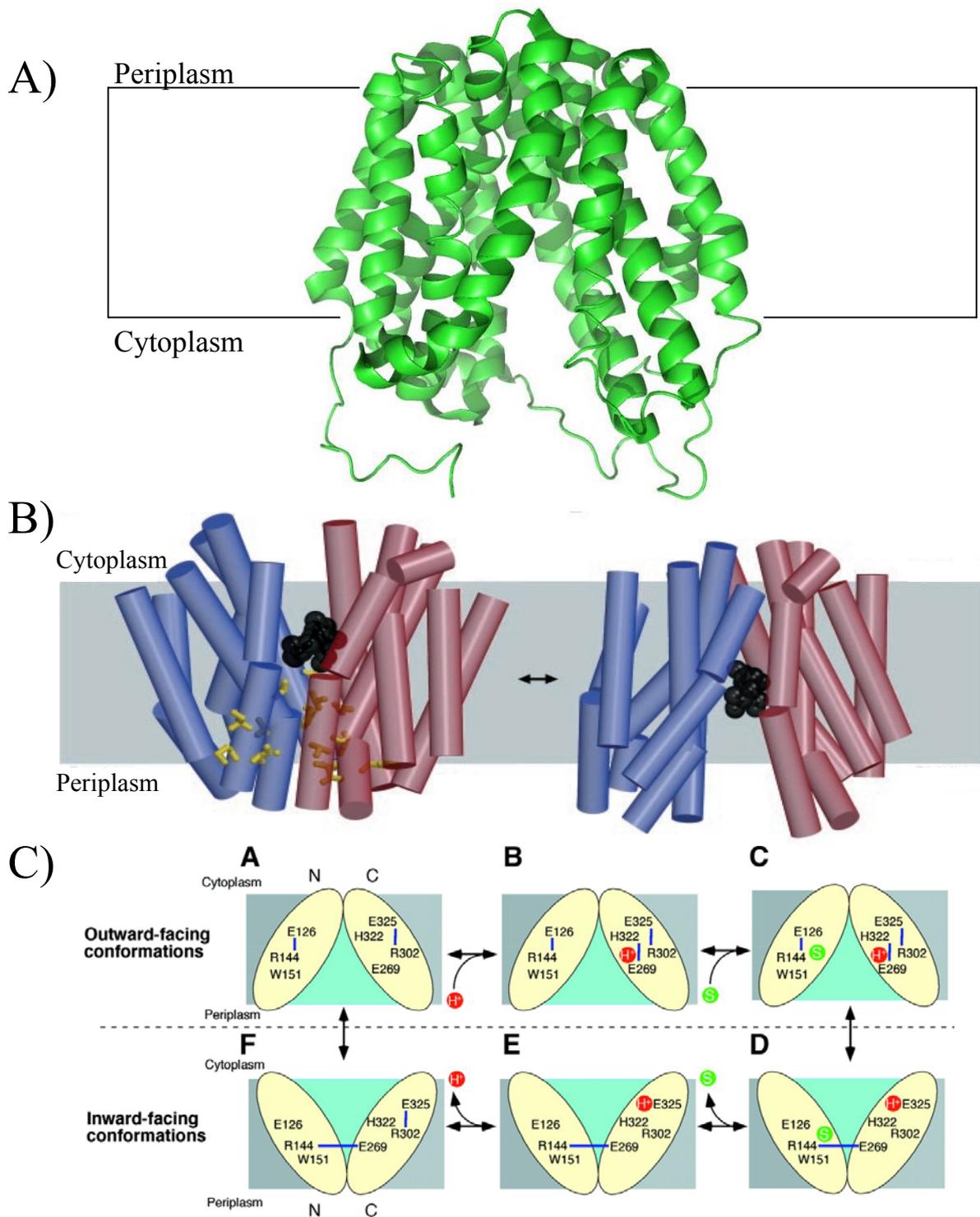
Many of these MFS families, such as the OHS, are restricted to bacteria and are 400-500 amino acid residues in length, containing 12 TMHs. Sugar porters also contain 12 TMHs, but are found in bacteria, archaea and eukarya and so their size range is larger, from 400-800 amino acids, with bacterial members generally being smaller. In MFS transporters, the first six TMHs show some homology to the last six, suggesting that the MFS transporter superfamily arose from a gene duplication and fusion event (Pao *et al.*, 1998).

#### **1.4.2 The *E. coli* lactose permease as a model system**

Since the reported discovery of the *E. coli lac* operon, the lactose permease LacY (FHS family) could be described as one of the most intensively studied proteins in history, with over 500 mutants constructed of the 417 amino acid polypeptide (Varela & Wilson, 1996). Like MalEFGK<sub>2</sub>, LacY has lent itself particularly well to analysis, in part due to the convenience of transport assay development.

##### **1.4.2.1 The structure of the lactose permease**

The protein crystal structure of LacY in a protonated, cytoplasm-open, periplasm-closed conformation was reported in 2007 (Figure 1.8a), which adopted the same conformation as the non-transporting Cys-154-Gly mutant that had been crystallised previously (Abramson *et al.*, 2003a, Abramson *et al.*, 2003b, Guan *et al.*, 2007). This showed the presence of the expected 12 TMHs and homology between the N- and C-terminal domains, which have approximately two-fold rotational symmetry and can be superimposed with an RMSD of 2.2 Å. Between these two domains is a water-filled cavity that contains the binding site, where the substrate analogue is bound. The protein crystal structure of the *E. coli* glycerol-3-phosphate transporter, GlpT was simultaneously reported and adopted a very similar cytoplasm-open, periplasm-closed conformation (Huang *et al.*, 2003).



**Figure 1.8:** **A)** Protein crystal structure of the *E. coli* lactose permease, LacY (Abramson *et al.*, 2003b). The 12 TMHs form two 6 TMH bundles with an approximate two-fold symmetry. These two bundles form a crevice extending into the membrane with the substrate binding site located approximately halfway through the membrane. **B)** Global changes between the cytoplasm-open structure and the periplasm-open model, *taken from Abramson et al. (2003b)*. *Left*, the cytoplasm-open structure of LacY with helical content represented as cylinders (N-terminal domain in blue and C-terminal domain in red). *Right*, the putative periplasm-open model of LacY. This was produced by a domain-domain rotation of 60°, supported by cysteine reactivities (residues shown in yellow on the *left*) and cross-linking studies. **C)** The proposed alternating access model for transport, *taken from Abramson et al. (2003b)*. The periplasm-open conformation, *A*, becomes protonated as shown in *B*. The substrate is bound, *C*, and the global conformation change produces the cytoplasm-open form, *D*, which dissociates the substrate, *E*, and then the proton. The substrate-free cytoplasm-open conformation, *F*, then flips back to the periplasm-open conformation.

#### 1.4.2.2 Proposed mechanism of transport

Both the LacY and GlpT structures support the proposed Alternating Access mechanism for transport (Figure 1.8b) (Locher *et al.*, 2003, Tanford, 1983). The putative periplasm-open, cytoplasm-closed structural model of LacY is shown in Figure 1.8b (Abramson *et al.*, 2003b). This was produced by a domain-domain rotation of 60° and is supported by cross-linking studies and by increased solvent exposure of the cysteine residues highlighted in the figure.

In this alternating access mechanism (Figure 1.8c), the deprotonated, periplasm-open conformation is very unstable and immediately binds a proton via the Glu-269–His-322 pair, producing the ground state of the transporter, to which substrate is bound by salt-bridged Glu-126 and Arg-144 pair (Abramson *et al.*, 2003b). To cause transport, this complex must adopt a highly unstable conformation, much like the maltose ABC transporter, previously. It is suggested that this transition is coupled to the formation of an inter-domain salt-bridge between Arg-144 and Glu-269. This then causes the release of the substrate into the cytoplasm followed by dissociation of the proton. The deprotonated, substrate-free, cytoplasm-open conformation then returns to the periplasm-open conformation, probably via breaking of the Arg-144–Glu-269 salt bridge and reformation of the Glu-269–His-322 pair.

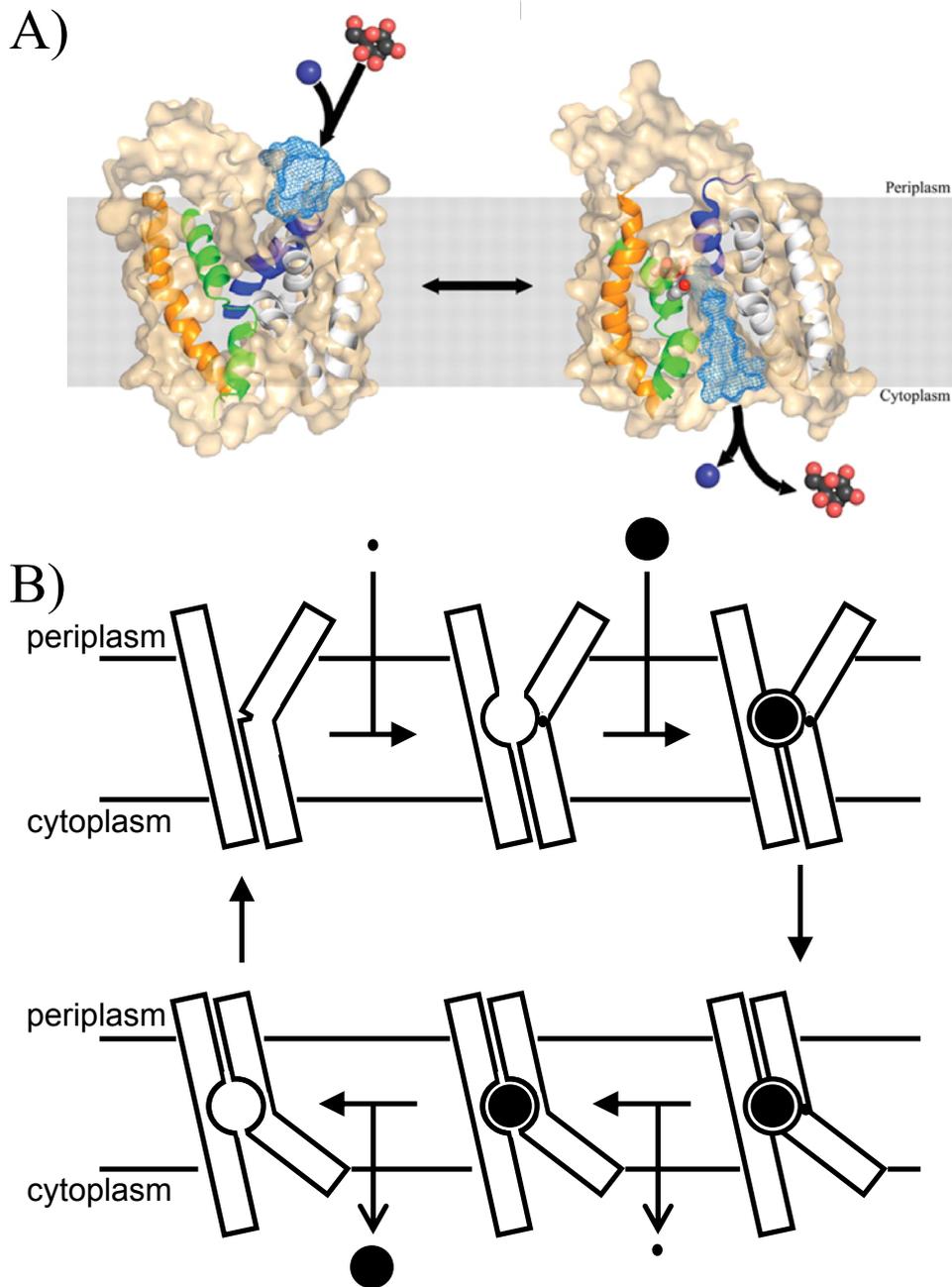
#### 1.4.3 Sodium-driven co-transport

As stated previously, secondary transporters can also utilise Na<sup>+</sup>-gradients to drive substrate accumulation. A major group of these is the sodium:solute symporter (SSS) family (TC2.A.21). These are ubiquitous, ranging in size from 400-700 amino acid residues, containing 12-15 TMHs and they transport a wide variety of substrates coupled with either 1 or 2 Na<sup>+</sup> ions (Faham *et al.*, 2008, Saier, 2000b). The protein crystal structure of the *Vibrio parahaemolyticus* sodium:galactose symporter (vSGLT; TC 2.A.21.3.2) was determined in a periplasm-open conformation with bound galactose and its single co-transported sodium ion (Faham *et al.*, 2008). This structure was found to share a similar 10 TMH-core topology with the previously crystallised *Aquifex aeolicus* leucine:2Na<sup>+</sup> symporter, LeuT (TC2.A.22.4.2), of the neurotransmitter:sodium symporter

(NSS) family, which was in the cytoplasm-open conformation and bound to leucine and two Na<sup>+</sup> ions (Yamashita *et al.*, 2005). This similar topology allowed the core structure of vSGLT to be mapped onto that of LeuT giving a structural basis for the alternating access mechanism (Figure 1.9a).

The alternating access model here is similar to that of LacY and is shown in Figure 1.9b (Faham *et al.*, 2008, Jung, 2001, Singh *et al.*, 2008, Yamashita *et al.*, 2005). In this, the periplasm-open conformation would first be protonated, where the binding of the proton would cause structural changes that form the substrate binding site (Faham *et al.*, 2008, Jung, 2001). Binding of the substrate would cause the closure of a hydrophobic periplasmic gate, made up of phenylalanine and tyrosine residues, to produce a periplasm-facing, occluded state. For the transport to progress, this periplasm-facing occluded state must be an intermediate state that moves through a transition state for global conformational change to produce the cytoplasm-facing occluded state. In this state, Na<sup>+</sup> is lost to the low sodium environment of the cytoplasm and the periplasmic hydrophobic gate (phenylalanine) opens, allowing the substrate to diffuse into the cell. The substrate-free, deprotonated, cytoplasm-open form of the transporter would then return to the periplasm-open conformation to complete the cycle.

It is worth noting that, as in MalEFGK<sub>2</sub> transport, the whole transport process, and each step in it, is akin to a series of intermediate and transition states of enzymatic catalysis, where substrates are bound so as to reduce the energy of activation for the process. For LeuT, this is highlighted by the lower rate of transport of its native substrates, leucine and methionine, when compared to non-native substrates, such as alanine, under saturating conditions (Singh *et al.*, 2008). This is explained by the high affinity of the periplasm-open, occluded intermediate state for the native substrate, which stabilises this transport intermediate, reducing the rate of transport.



**Figure 1.9:** The structure and mechanism of sodium-coupled symport. **A)** Global changes between the cytoplasm-open structure and the periplasm-open model, taken from Faham *et al.* (2008). Surface representations of the periplasm-open model (*left*) and the cytoplasm-open structure (*right*) of vSGLT with helices that undergo structural rearrangement shown as ribbons, galactose and sodium shown as coloured spheres and extracellular, water-filled cavities shown as blue mesh. **B)** The proposed mechanism for sodium-coupled symport. An extracellular sodium ion associates near the substrate binding site, causing structural rearrangements that allow substrate binding. The substrate is located near the centre of the membrane, held between a periplasmic and a cytoplasmic gate, both made up of hydrophobic residues, and represented by a complete circle when they have been shown to be involved. The transporter undergoes a global conformational change from the periplasm-facing to the cytoplasm-facing form. The sodium ion dissociates, followed by the substrate, and the transporter then returns to the periplasm-open conformation.

## 1.5 SBP-dependent secondary active transport

The use of soluble binding proteins in transport is not restricted to primary active transporters, but has also been found as part of secondary active transport mechanisms that are widespread in prokaryotes and archaea, but absent from eukaryotes (Forward *et al.*, 1997, Jacobs *et al.*, 1996, Kelly & Thomas, 2001, Mulligan *et al.*, 2010).

In these, a periplasmic substrate binding protein (SBP) is associated with two membrane proteins of 4 and 12 TMHs, which can often be found fused into a single polypeptide (Figure 1.2) (Kelly & Thomas, 2001, Rabus *et al.*, 1999). These types of transporter are split into two classes. The largest is the tripartite ATP-independent periplasmic (TRAP) transporters (TC 2.A.56), after the three components involved. The second, smaller class is the tripartite tricarboxylate transporters (TTT; TC 2.A.80), which share the same three-component structure but no sequence homology (Mulligan *et al.*, 2010).

### 1.5.1 Features and components of SBP-dependent secondary transporters

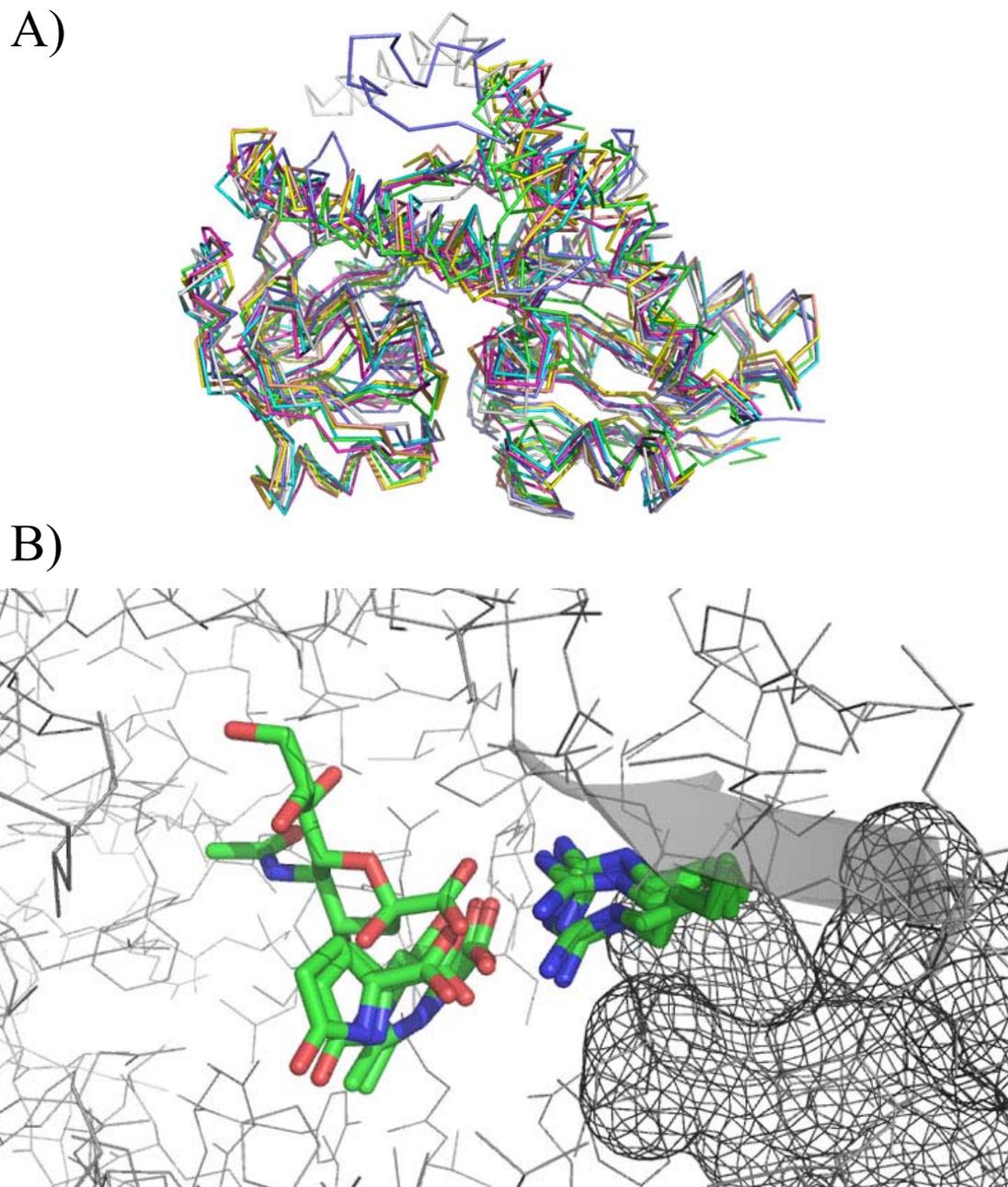
The founding member of this class of transporter was the *R. capsulatus* C<sub>4</sub>-dicarboxylate transporter, which is responsible for the accumulation of malate, succinate and fumarate (Forward *et al.*, 1997). The components of this TRAP transporter system are DctP (the SBP), DctQ (the small integral membrane protein) and DctM (the large integral membrane protein).

The DctM family of proteins, believed to contain the translocation pores, is part of the Ion Transporter (IT) superfamily and shares closest homology with the DcuC family of C<sub>4</sub>-dicarboxylate secondary active transporters (TC2.A.61) (Prakash *et al.*, 2003). The function of the DctQ family of smaller integral membrane proteins is unknown and has no homology with any known protein family (Kelly & Thomas, 2001, Rabus *et al.*, 1999). It is suspected that this membrane subunit is involved in mediating the interaction or initial docking of the SBP, since these two proteins can sometimes be found fused together (Kelly & Thomas, 2001, Rabus *et al.*, 1999).

DctP family SBPs share low sequence homology with each other, but their 3D structures, of which there are seven thus far, are highly similar with an R.M.S.D. of less than 3 Å (Figure 1.10a) (Fischer *et al* 2010). Like the binding proteins of ABC transporters, these SBPs are made up of two  $\alpha/\beta$  globular domains that are linked by a flexible hinge region (Muller *et al.*, 2006). However, they also contain a long  $\alpha$ -helix that stretches the full length of the protein and bends in the closed conformation.

There is an additional family of binding proteins associated with TRAP transporter DctQM homologs called TRAP associated extracytoplasmic immunogenic (TAXI) proteins, which show sequence homology to the *E. coli* glutamine binding protein family from primary active transporters and were originally annotated as immunogenic proteins (Kelly & Thomas, 2001). The one example of these that has been crystallised is structurally similar to type II ABC PBPs not DctP family binding proteins. These TAXI SBPs are always associated with a DctQM fusion protein and all TRAP transporters in archaea are TAXI-DctQM.

The tripartite tricarboxylate transporter (TTT) family is the second class of SBP-dependent secondary transporters and share similar architectural subunits, but no sequence homology with TRAP transporters (Winnen *et al.*, 2003). The classical example of these is the *Salmonella typhimurium* citrate transporter, TctABC (Kay & Cameron, 1978, Widenhorn *et al.*, 1988). Several structures of TTT binding proteins have been determined and are also structurally related to type II PBPs (Huvent *et al.*, 2006).

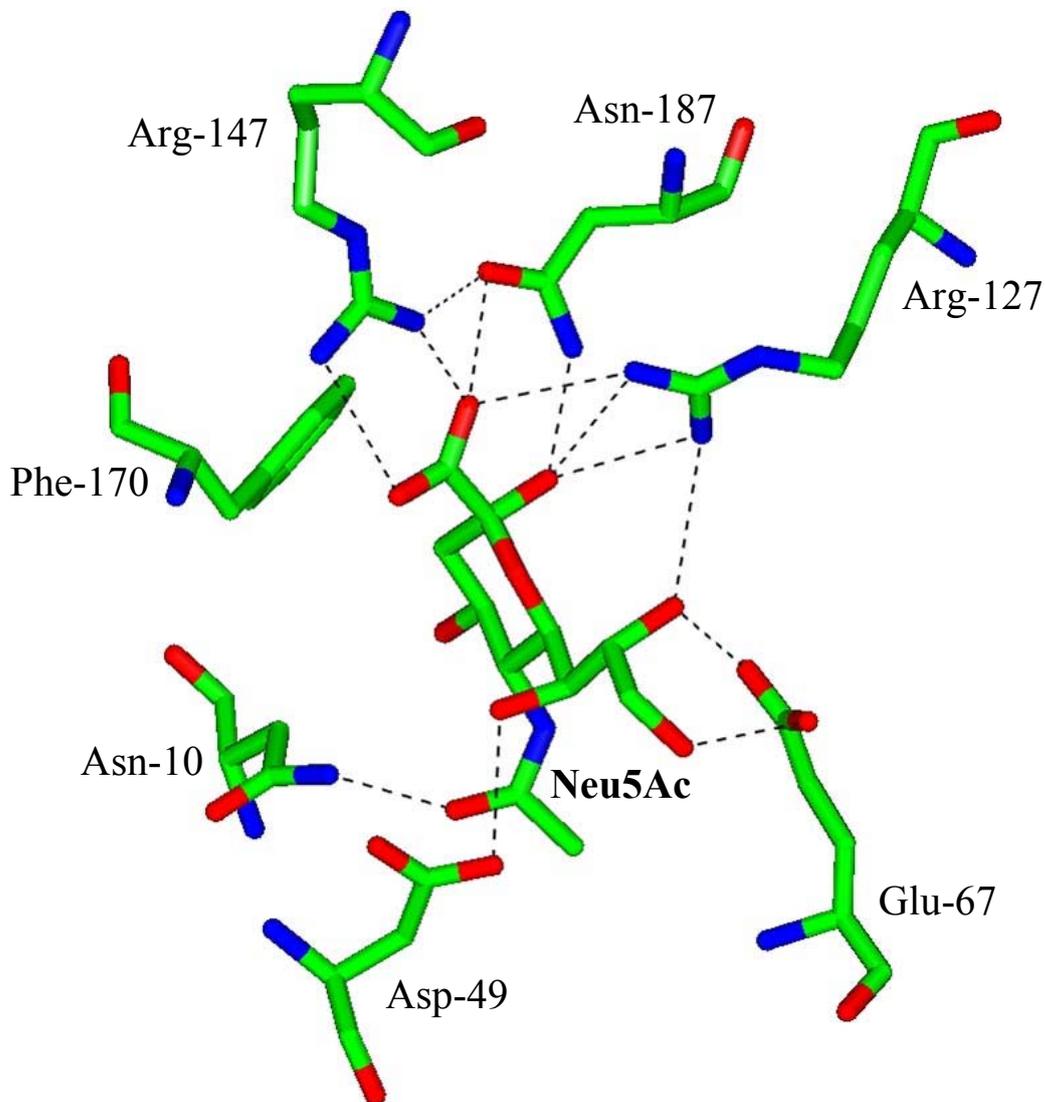


**Figure 1.10:** The structures of DctP-type SBPs. **A)** Superposition of the closed conformation of the SBPs DctP6, DctP7, SiaP, TakP, LakP, UehA, TeaA. The average sequence identity between these proteins is about 20%, while their structures have an R.M.S.D. of less than 3 Å. **B)** The ligand co-ordination by the conserved arginine residue. The arginine residue (*right*) and the ligands (*left*) are shown as atom-coloured cylinders within the SiaP binding site (grey lines). The  $\beta$ -sheet secondary structure of the chain containing the arginine residue is shown as a grey ribbon and the hydrophobic patch from SiaP is shown as a black mesh.

### 1.5.1.1 Co-ordination of the ligand by SBPs

At present, seven DctP-family SBPs have been crystallised and their structures published. All of these bind organic acids via a mix of polar and non-polar interactions; the only common feature is the use of a totally conserved arginine residue to form a bipartite salt bridge to the carboxylate group of the ligand (Figure 1.10b). The conserved arginine residue extends into the binding site in both the open and closed conformations, supported by stable interactions around its  $\beta$ -sheet secondary structure and its proximity to a conserved group of hydrophobic residues (Figure 1.10b) (Fisher *et al.*, in publication). The ligands represented in this group of structures vary in size between sialic acid (MW 309 Da) and pyruvate (MW 88 Da) (Lecher *et al.*, 2009).

*Haemophilus influenzae* SiaP binds the most common form of sialic acid, *N*-acetylneuraminic acid (Neu5Ac), inside its binding pocket (Figure 1.11) with a  $K_d$  of  $0.12 \pm 0.01 \mu\text{M}$  (Severi *et al.*, 2005). The previously determined crystal structures of SiaP with bound Neu5Ac and 2,3-didehydro-2-deoxy-*N*-acetylneuraminic acid (Neu5Ac2en) (Muller *et al.*, 2006, Johnston *et al.*, 2008) showed a number of direct interactions between the protein and the ligand. Firstly, the carboxylate group of the sialic acid ligand is coordinated by a salt-bridge network to Arginine-147, Asparagine-187 and Arginine-127. It is Arg-147 which is the most highly conserved residue in the DctP-type TRAP SBP family (Muller *et al.*, 2006) and it is this residue that donates a bipartite salt bridge between its two terminal nitrogen atoms and the oxygen atoms of the ligand carboxylate group. Secondly, the ligand glycerol group interacts with Aspartate-49 and Glutamate-67 and its C<sub>3</sub>-chain is subject to hydrophobic packing interactions with Alanine-151. Thirdly, the *N*-acetyl group of the ligand forms an H-bonding interaction with the side chains of Asparagine-10 and a hydrophobic interaction with Phe-65. Finally, Phenylalanine-170 forms a stacking interaction with the sugar ring (Muller *et al.*, 2006, Johnston *et al.*, 2008).



**Figure 1.11:** Representation of the interactions between Neu5Ac and the binding site of SiaP, based on 3B50.pdb (Johnston et al., 2008). Amino acid residues and the sialic acid ligand are shown as atom coloured cylinders; hydrogen bonds are shown as dashed black lines; Phenylalanine-65 and Alanine-151 are not shown for clarity.

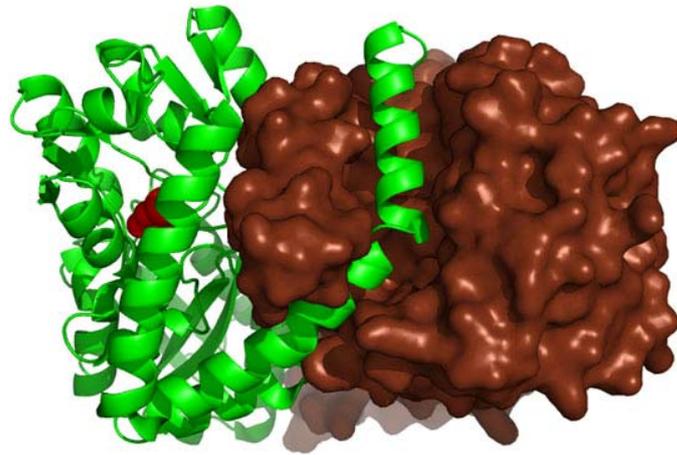
### 1.5.1.2 Multimeric SBPs

The protein crystal structures of the *Rhodobacter sphaeroides* SBP for sodium pyruvate, TakP, and the *Thermus thermophilus* SBP for calcium lactate, LakP, revealed that these exist as back-to-back dimers in solution (Figure 1.12) (Akiyama *et al.*, 2009, Gonin *et al.*, 2007). The first, TakP, swaps an extended C-terminal  $\alpha$ -helix with its corresponding monomer, generating an interface of  $\sim 3600 \text{ \AA}^2$ . The second, LakP, has this C-terminal extension but a helix swap does not occur in the dimer, yet this complex still forms in solution.

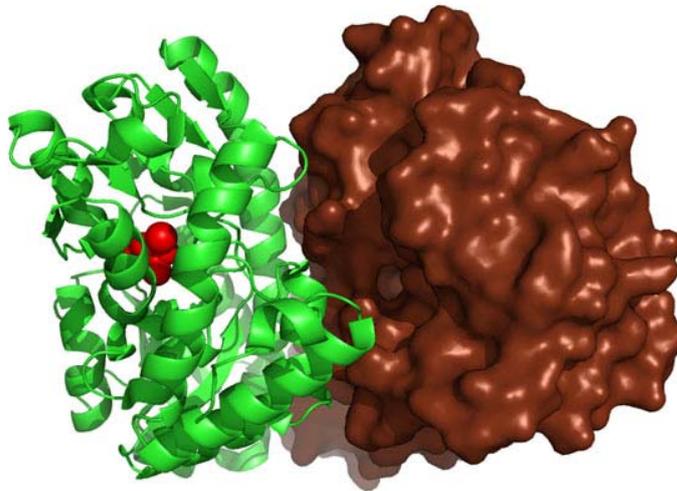
The role of the dimer is not clear, since current models for TRAP transport are based on monomeric binding proteins, as in ABC transport (Mulligan *et al.*, 2009). On the discovery of the TakP dimer, a transport model was proposed where this would remain associated with the permease and the substrate would be bound by the topmost exposed SBP, diffuse through a channel in the dimer to the second SBP and onto the permease (Gonin *et al.*, 2007). However, this supposed channel appears to be blocked by several hydrophobic residues (Fischer *et al.*, in publication).

It is also worth noting that a multimer of the *Thermatoga maritima* SBP TM0322 has been reported; however, these were found in addition to observable monomeric TM0322 (Cuneo *et al.*, 2008). The multimers here were reported to be front-to-front dimers and tetramers, forming an occluded cavity containing the binding sites. It has since been suggested that these complexes are artefacts and non-physiological (Fischer *et al.*, in publication).

A)



B)



**Figure 1.12:** Dimers of the DctP-type SBPs (A) TakP and (B) LakP. One monomer is shown in surface representation and the second as ribbons. The ligands are shown as red spheres.

## 1.5.2 The proposed mechanism of transport

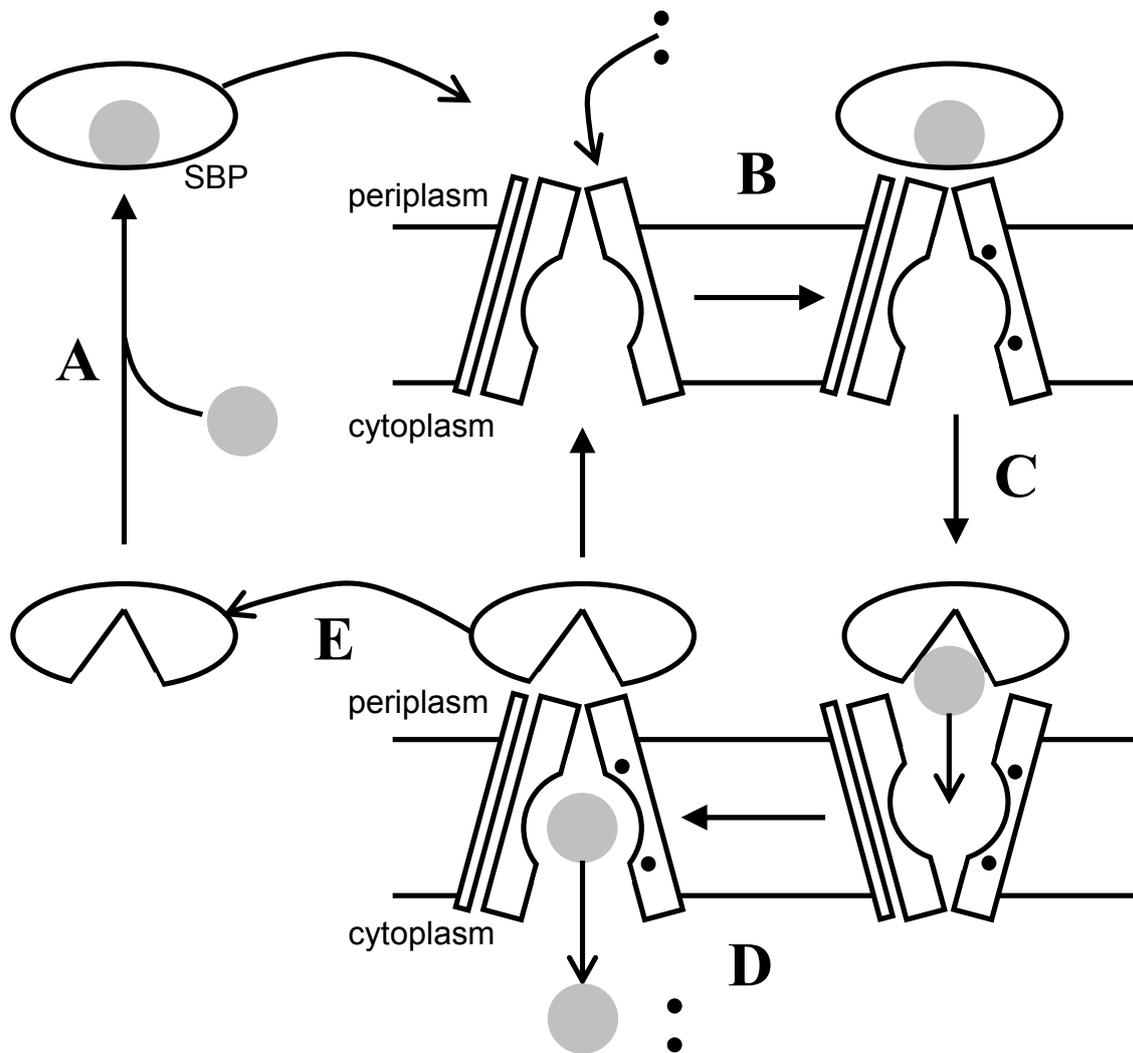
### 1.5.2.1 The *R. capsulatus* C<sub>4</sub>-dicarboxylate transporter, DctPQM

The purple photosynthetic *Rhodobacter capsulatus* can utilise the C<sub>4</sub>-dicarboxylate compounds malate, succinate and fumarate as the sole source of carbon (Stahl & Sojka, 1973). This was shown to be via a single periplasmic binding protein-dependent transport system that was encoded in the *dct* locus (Hamblin *et al.*, 1990). However, instead of an ABC transporter in this region, *dctQ* and *dctM* were discovered (Forward *et al.*, 1997). The further investigation of this transporter made this the first genetically and biochemically characterised binding protein-dependent secondary transporter and the founding member of this class. As in the cases of MalEFGK<sub>2</sub> and LacY, this system lent itself well to analysis, in part due to the traditional use of *R. capsulatus* as a model organism.

It has been proposed that the closed, substrate-bound SBP interacts with the integral membrane proteins to deliver the substrate in much the same way as in ABC transport (Mulligan *et al.*, 2009). At this point, like ABC transporters, the interaction of the closed SBP would transfer the substrate to the permease and catalyse translocation into the cytoplasm. The energy coupling mechanism of transport is suggested to be co-transport of two or more H<sup>+</sup> or Na<sup>+</sup> ions (Kelly & Thomas, 2001). It has been noted that TRAP transporters are enriched in marine organisms, suggesting the use of Na<sup>+</sup> ions (Mulligan *et al.*, 2007).

### 1.5.2.2 The *H. influenzae* sialic acid transporter, SiaPQM

The obligate human mucosal pathogen *Haemophilus influenzae* uses a TRAP transporter for the uptake of exogenous sialic acid, an important compound in its virulence (Bouchet *et al.*, 2003, Severi *et al.*, 2005). This transporter, SiaPQM, contains fused membrane subunits. The characterisation of this transporter has recently moved forward due to the development of an *in vitro* transport assay of the reconstituted system (Mulligan *et al.*, 2009). This has shown that transport of the substrate is coupled to co-transport of two or more Na<sup>+</sup> ions. The proposed mechanism is shown in Figure 1.13. The closed, substrate-bound SBP, SiaP, interacts with the SiaQM, which is probably in a periplasm-closed conformation, as in ABC transporters. SiaQM also binds the Na<sup>+</sup> ions that drive



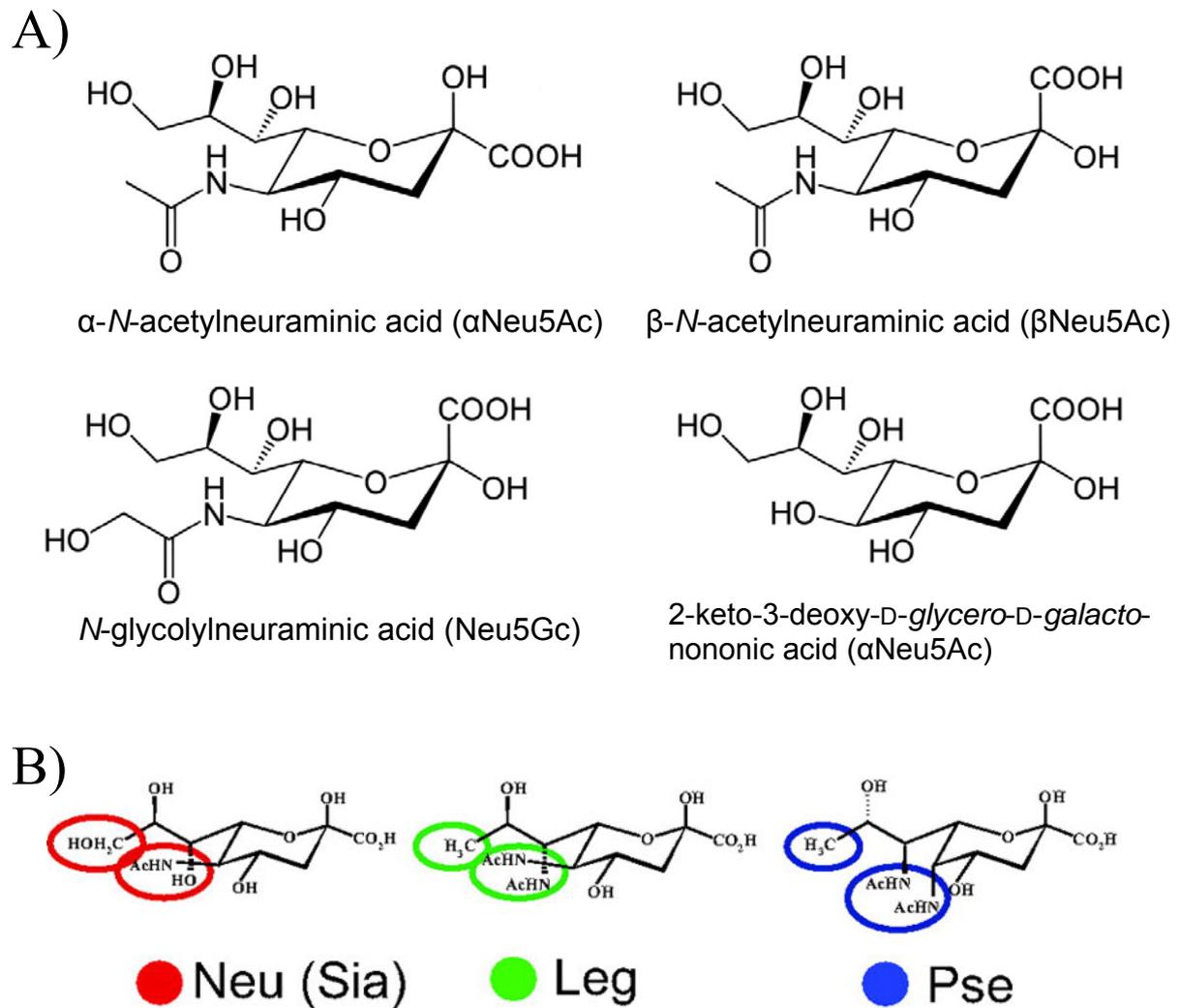
**Figure 1.13:** The proposed mechanism of tripartite ATP-independent periplasmic (TRAP) transporters (Mulligan *et al.*, 2009). **A)** Substrate (grey circle) is bound by the SBP. **B)** The liganded SBP and at least two sodium ions (black dots) bind (open arrows) to the permease. **C)** The membrane complex adopts a periplasm-open conformation (solid arrows), opening the SBP and accepting the substrate into an internal binding site. **D)** The conformation of permease resets to its cytoplasm-open form, releasing the substrate and  $\text{Na}^+$  ions into the cell. **E)** The SBP at this point is ligand-free, but could be in an open or closed conformation. At this point, it would dissociate from the complex.

transport; however, the order of these two events is not certain. Substrate translocation through SiaQM is presumably via an alternating access mechanism, involving the opening of SiaP, in a similar manner to MalE in MalEFGK<sub>2</sub>, and an intramembrane substrate binding site. Transport of the substrate by SiaQM cannot be reversed, as many traditional secondary transporters can. However, a large excess of unliganded SiaP can cause some efflux of a high concentration of substrate. For this to occur, unliganded, open-conformation SiaP would have to interact with SiaQM, which would be saturated with substrate in the cytoplasm-open form. This shows that binding of the SBP is coupled to the conformational changes associated with substrate translocation.

## 1.6 Sialic acids

As discussed in the previous section, the *H. influenzae* transporter SiaPQM is responsible for the transport of the virulence factor and carbon source sialic acid, which is also known as *N*-acetylneuraminic acid (Neu5Ac). This is an  $\alpha$ -keto acidic sugar that adopts a chair-configuration pyranose ring through a hemiketal condensation (Angata & Varki, 2002). Since its discovery, it has been found that this is just one member of a family of similar structures, which are known as Sialic acids and total over 50 different members (Angata & Varki, 2002, Vimr *et al.*, 2004). The most common member, *N*-acetylneuraminic acid (Neu5Ac; 5-acetylamino-3,5-dideoxy-D-glycero- $\alpha$ -D-galactonulosonic acid), usually referred to as sialic acid (Figure 1.14), can be found modified at every available position to generate a large range of compounds. Three important members of this family, shown in Figure 1.14, are *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc) and a de-aminated form, 2-keto-3-deoxy-D-glycero-D-galactononic acid (KDN).

Two other groups related to the Sialic acid family are Legionaminic (Leg) acid from *Legionella* and Pseudominic (Pse) acid from *Pseudomonas*, shown in Figure 1.14. Collectively, these complex, similar structures are nonulosonic acids and are all synthesised and utilised in similar manners (Lewis *et al.*, 2009). Recently, it has been proposed that sialic acid, rather than having arisen in animals and transferred to bacteria,



**Figure 1.14:** Nonulosonic acids found in bacteria. **A)** Major members of the Sialic acids. **B)** The three different families of nonulosonic acids, taken from Lewis et al. (2009).

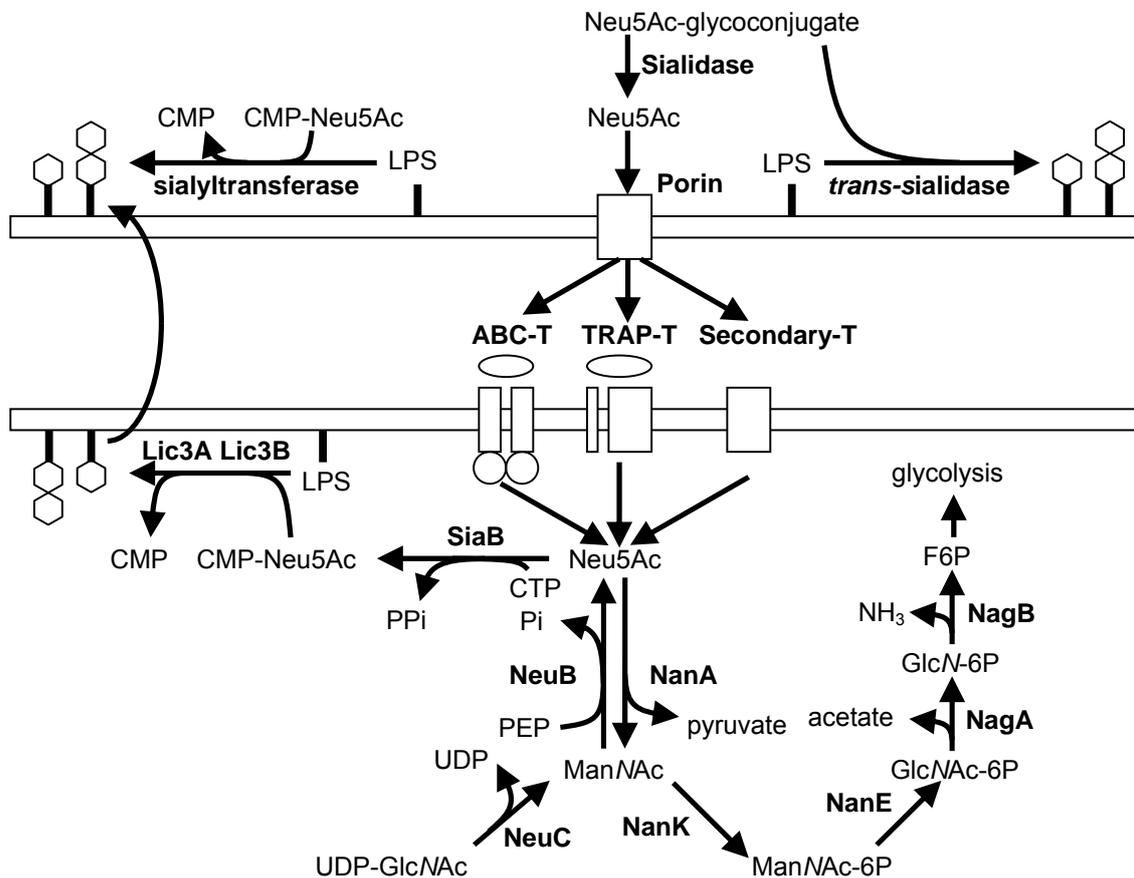
arose from the ancient evolution of nonulosonic acids to be adopted by animals (specifically, deuterostomes) and lost by other eukaryotes (Lewis *et al.*, 2009).

### 1.6.1 Utilisation and catabolism of sialic acids

Sialic acids, widely used in eukaryotes in cell-cell and cell-molecule signalling, are also expressed on the surfaces of many pathogens, mimicking the mucins and glycoconjugates of their host (Severi *et al.*, 2007, van der Merwe *et al.*, 1996, Vimr & Lichtensteiger, 2002). In many cases, the display of sialic acid on the LPS of these pathogens is crucial for infection and survival in the host (Vimr *et al.*, 2004). This can be mediated via different mechanisms, including resistance to serum and the innate immune response (Hood *et al.*, 1999, Ogasawara *et al.*, 2007), dampening of the host immune response (Carlin *et al.*, 2009) and mimicry of the host glycoforms (Mandrell *et al.*, 1992, Xiang *et al.*, 2006). This mimicry can cause health problems beyond the infection, giving rise to autoimmune conditions such as Guillain-Barre syndrome, where *Campylobacter jejuni* mimics peripheral nerve gangliosides and antibodies against this also target peripheral nerves (Xiang *et al.*, 2006).

Host glycoconjugates can also mediate interactions with pathogens and their attachment to host cells (Karlsson, 1998). This is demonstrated by a proposed inhibitor for influenza virus infection, which contains a sialic acid-cleaving enzyme (sialidase) and a cell surface-anchoring sequence (Malakhov *et al.*, 2006). When delivered to the airway epithelium, this causes a non-toxic stripping of cell surface sialic acid, reducing the infectivity of the virus.

In bacteria, sialic acid can be synthesised *de novo* or can be acquired from the environment. As shown in Figure 1.15, sialic acid is synthesised from the common metabolite UDP-*N*-acetylglucosamine (UDP-GlcNAc) via *N*-acetylmannosamine (ManNAc) (Severi *et al.*, 2007, Vimr *et al.*, 2004). To be acquired from the environment, sialic acid must be released from the host sialoglycoconjugates by sialidases, which are released by many pathogens, possibly supplying those that lack a sialidase and are also proposed to be released by the host as a part of sialic acid recycling (Corfield *et al.*, 1992, Corfield, 1992, Severi *et al.*, 2007). This free sialic acid is then transported in to the cell



**Figure 1.15:** Sialic acid utilisation by bacteria. Sialic acid (Neu5Ac) is released from sialoglycoconjugates by a sialidase and reaches the transporter through a porin in Gram-negative bacteria. Sialic acid is then transported into the cytoplasm through an ABC, secondary or TRAP transporter. Sialic acid catabolism uses the *nan* and *nag* genes as shown. *De novo* sialic acid is synthesised from UDP-GlcNAc. Sialylation of the LPS follows activation of sialic acid as CMP-sialic acid (CMP-Neu5Ac) and transfer to the LPS molecule. Lic3A and Lic3B are responsible for mono- and di-sialylation. *N. gonorrhoeae* uses an extracellular sialyltransferase and host-derived CMP-Neu5Ac to transfer sialic acid to its LPS. In some Trypanosome species, a *trans*-sialidase directly attaches Neu5Ac from sialoglycoconjugates to its LPS.

using ABC, secondary or SBP-dependent secondary transporters (Severi *et al.*, 2007). In some Gram-negative bacteria, a sialic acid-specific porin, NanC, is present; however, this is only essential in the absence of the general porins, OmpC and OmpF (Condemine *et al.*, 2005).

Once transported, sialic acid can be catabolised by the *nan* genes to *N*-acetylglucosamine-6-phosphate (GlcNAc6P) and on to fructose-6-phosphate (F6P) by the *nag* genes (Figure 1.15) (Plumbridge & Vimr, 1999, Vimr & Troy, 1985). To be incorporated into the LPS, sialic acid must be activated by CMP-sialic acid synthetase with CTP to give CMP-sialic acid and PP<sub>i</sub> (Bouchet *et al.*, 2003, Severi *et al.*, 2007, Vimr *et al.*, 2004). This is then added to the terminal positions of the LPS acceptors by sialyltransferases, such as the *H. influenzae* Lic3A, which transfers sialic acid to a lactose-glycoform (Hood *et al.*, 2001, Kalovidouris *et al.*, 2003, Severi *et al.*, 2007).

It is worth noting that there are several examples where sialylation of the LPS occurs extracellularly (Figure 1.15). In some *Trypanosoma* species and *Corynebacterium diphtheriae*, an extracellular *trans*-sialidase, transfers a host sialic acid residue to its own LPS in a reaction similar to the hydrolysis that would release free sialic acid (Agusti *et al.*, 2007, Vimr *et al.*, 2004). In *Neisseria gonorrhoeae*, transport of sialic acid does not occur; instead, this bacterium has a membrane-associated, surface-exposed sialyltransferase, which would directly transfer sialic acid from the host's CMP-sialic acid (Severi *et al.*, 2007, Vimr *et al.*, 2004).

#### **1.6.1.2 Use of sialic acid by *H. influenzae***

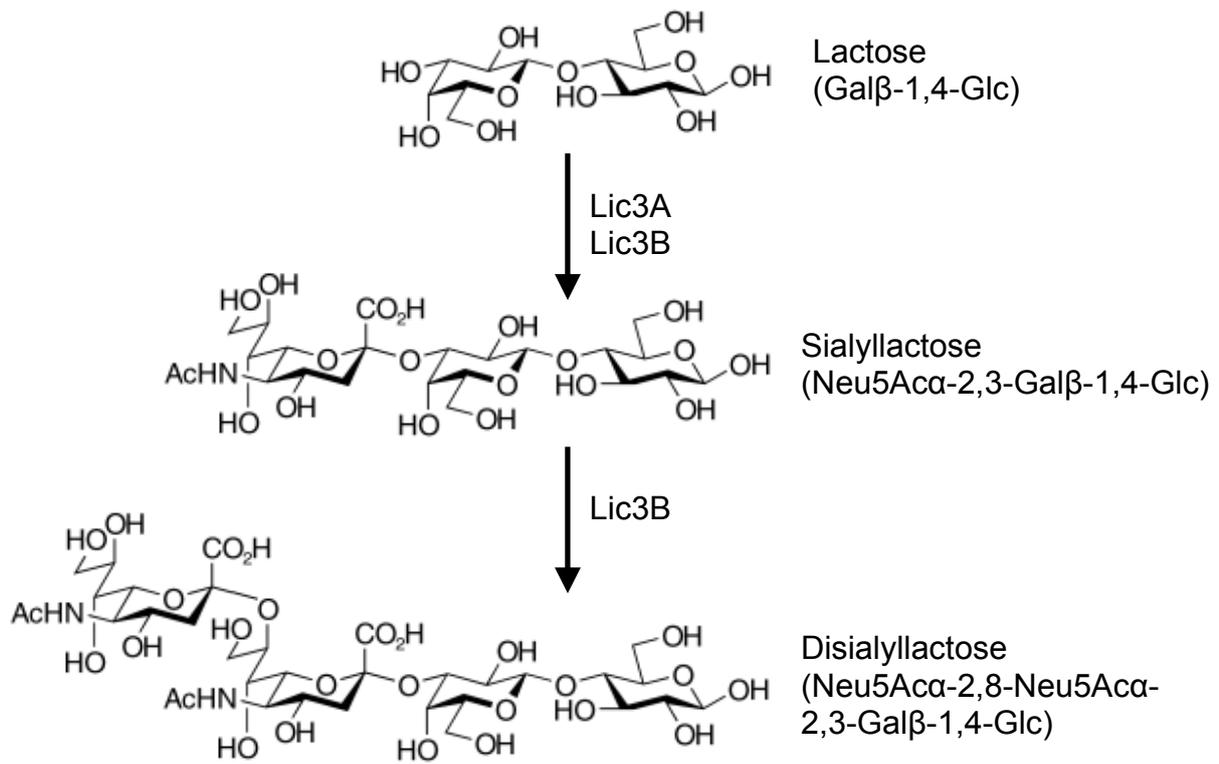
*H. influenzae* is a Gram-negative, non-motile, facultative anaerobic, obligate parasite of the human nasopharyngeal mucosa and potential pathogen, which can be found with a capsule or without (non-typable; NT) (Aubrey & Tang, 2003, Rao *et al.*, 1999, Vimr *et al.*, 2000). NT *H. influenzae* can cause conjunctivitis and upper respiratory tract infections such as otitis media and sinusitis. Typable *H. influenzae* can be split into six serogroups a-f, with most diseases caused by serogroup b, which include meningitis, septicaemia, septic arthritis, pneumonia and empyema (Aubrey & Tang, 2003).

*H. influenzae* cannot synthesise sialic acid *de novo* and so is reliant upon an exogenous supply of free sialic acid. Transport through the inner membrane is mediated by the TRAP transporter SiaPQM, where the membrane domains are fused into a single polypeptide.

Intracellular sialic acid can be used as a carbon, nitrogen and energy source via the *nan* genes, which are transcribed divergently from the *siaPQM* operon (Johnston *et al.*, 2007). It has been demonstrated that catabolism of sialic acid competes with LPS sialylation, which confers increased serum resistance, biofilm formation and pathogenicity (Vimr *et al.*, 2000). Two sialyltransferases, Lic3A and Lic3B, have been identified and characterised in *H. influenzae*. Lic3A is an  $\alpha$ -2,3-sialyltransferase responsible for adding sialic acid to lactose acceptors via a 2,3-linkage (Figure 1.16) (Hood *et al.*, 2001). Lic3B transfers sialic acid via 2,3-linkage to a lactose acceptor and via 2,8-linkage to sialyllactose acceptor (Figure 1.16) (Fox *et al.*, 2006). This addition of di-sialic acid to the LPS results in further increased serum resistance.

### **1.7 Aims of this investigation**

This investigation will focus on the ligand binding mechanism and transport function of the *H. influenzae* SBP for sialic acid, SiaP. The primary aim of this project is to investigate ligand binding and co-ordination by SiaP and to examine the function of SiaP in transport, both *in vitro* using SiaQM reconstituted into proteoliposomes and *in vivo* by expression in *E. coli* BW25113 lacking its native sialic acid transporter, NanT. Also, the transport and catabolism of several sialic acids will be examined using the expression of different sialic acid transporters and deletions of known sialometabolic genes. This project also aims to investigate the possibility of producing a SiaP variant that is compatible with the current technology used by our CASE sponsor, Authentix.



**Figure 1.16:** Sialyltransferase activities characterised in *H. influenzae*. Lic3A and Lic3B can both transfer sialic acid from CMP-sialic acid to a lactose-terminal LPS structure via an  $\alpha$ -2,3-linkage. Lic3B can transfer a second sialic acid to sialyllactose via an  $\alpha$ -2,8-linkage.

# Chapter Two

## Materials and Methods

## **2.1 Media and antibiotics**

### **2.1.1 Luria-Bertani broth**

Luria-Bertani (LB) broth was made up of 10 g/l tryptone (Formedium) 5 g/l powdered, dried yeast (Formedium) 10 g/l NaCl (Fisher Scientific).

### **2.1.2 M9 minimal medium**

M9 minimal medium salts (Neidhardt *et al.*, 1974) were made up from 6 g/l Na<sub>2</sub>HPO<sub>4</sub> (Fisher Scientific), 3 g/l KH<sub>2</sub>PO<sub>4</sub> (Fisher Scientific), 0.5 g/l NaCl<sub>2</sub>, 1 g/l NH<sub>4</sub>Cl (Melford). 1 M MgSO<sub>4</sub> (Fisher Scientific) and 25 mg/ml FeSO<sub>4</sub>.7H<sub>2</sub>O (AnalaR) were sterilised separately. The M9 minimal medium salts were supplemented with 2 mM MgSO<sub>4</sub> and 25 µg/ml FeSO<sub>4</sub>.7H<sub>2</sub>O. The sterilised carbon source of interest was then added to this.

### **2.1.3 Solid media**

To prepare a solid medium, 1% agar (Oxoid) was added to LB broth and 1% agarose (Melford) was added to double M9 minimal medium.

### **2.1.4 Antibiotics**

Antibiotic selection, where appropriate, was used at 30 µg/ml chloramphenicol (Sigma-Aldrich) in 80% ethanol (Fisher Scientific), 100 µg/ml ampicillin (Melford) and 50 µg/ml kanamycin (Sigma-Aldrich) each from 1000-fold stock, 0.22 µm filter sterilised.

## **2.2 Table of primers**

The primers used in this work are shown in Table 2.1.

## **2.3 Strains and plasmids**

Table 2.2 contains all of the bacterial strains used here, while Table 2.3 lists all of the plasmids used and constructed during this work.

Table 2.1 Primer list

Primer	Target	RE	Sequence
siaPfor	<i>siaP</i>	KpnI, NdeI	gcggtacctaagaaggagatatacatatgatgaaattgacaaaaac
siaPrev	<i>siaP</i>	XhoI	ccgctcgagttatggattgattgcttc
siaPhis6rev	<i>siaP-His<sub>6</sub></i>	XhoI	ccgctcgagttagagatggggtgatgatgtggattgattgcttcaatttg
R147Afor	R147A		aaacttgctgtgccaaatgcagcaacaaac
R147Arev	R147A	StuI	ggcacagcaagtttcaggcctttcatatctgc
R147Efor	R147E		aaacttgaagtgccaaatgcagcaacaaac
R147Erev	R147E	StuI	ggcacttcaagtttcaggcctttcatatctgc
R147Kfor	R147K		aaacttaaagtgccaaatgcagcaacaaac
R147Krev	R147K	StuI	ggcactttaagtttcaggcctttcatatctgc
R127Afor	R127A	SfuI	ggaactgccccaaagcacttcgaatcgtgc
R127Arev	R127A		gctttgggcagttccggtataagcttggg
R127Efor	R127E	SfuI	ggaactgaacaaagcacttcgaatcgtgc
R127Erev	R127E		gctttgttcagttccggtataagcttggg
R127Kfor	R127K	SfuI	ggaactaaacaaagcacttcgaatcgtgc
R127Krev	R127K		gctttgtttagttccggtataagcttggg
F170Afor	F170A	NcoI	atggcagcttctgaagtttatcttgcgttac
F170Arev	F170A		cttcagaagctgccatgggtggttgatgc
F170Wfor	F170W	NcoI	atggcatggtctgaagtttatcttgcgttac
F170Wrev	F170W		cttcagaccatgccatgggtggttgatgc
F170Yfor	F170Y	NcoI	atggcatattctgaagtttatcttgcgttac
F170Yrev	F170Y		cttcagaatatgccatgggtggttgatgc
N187Afor	N187A		tcaagaagccccgtagccgcggtgcaagc
N187Arev	N187A	SacII	acggggcttcttgaccatcgacggcattgg
N187Dfor	N187D		tcaagaagaccgtagccgcggtgcaagc
N187Drev	N187D	SacII	acgggtcttcttgaccatcgacggcattgg
F75Wfor	F75W	EagI	cagctgtggtaccctgaagcggccgtatattgccc
F75Wrev	F75W		agggtaccacagctggaagcgagcagattc
Y76Wfor	Y76W	EagI	ctgttttgccctgaagcggccgtatattgccc
Y76Wrev	Y76W		ttcaggccaaaacagctggaagcgagcac
Y161Wfor	Y161W		gctaaatgggttggtgcatcaccaacacc
Y161Wrev	Y161W	PspLI	accaaccatattagcgtacgctaagtttgttgc

A11Nfor	A11N	KpnI	gtatgaataatggtacctcatcaaataaagcgg
A11Nrev	A11N		gtaccattattcataccgaatttcaagtcataatcagc
A66Ifor	A66I		cctttatagaatctgctcgcttccagc
A66Irev	A66I	XbaI	gcagattctataaaggtaaagtctagagaaccg
A66Mfor	A66M		cctttatggaatctgctcgcttccagc
A66Mrev	A66M	XbaI	gcagattccataaaggtaaagtctagagaaccg
A151Vfor	A151V		ccaaatgtagcaacaaacttagcctatgc
A151Vrev	A151V	KpnI	gtttgttgctacatttggtagccgaagttttaagcc
S15Dfor	S15D	KpnI	gctggtaccgattcaaataaagcggc
S15Kfor	S15K	KpnI	gctggtaccaaataaataaagcggc
A195Drev	A195D	SacII	gcacttcatagaatTTTTgatcttgcaccgaggtaacg
A195Krev	A195K	SacII	gcacttcatagaatTTTTgttttgcaccgaggtaacg
Q72Efor	Q72E	Kpn2I	gctcgcttcgagctgttttatccggaagcgg
A151Krev	A151K	KpnI	ggctaagttgttttgcatttggtagccgaagttttaagcc
A152Krev	A152K	KpnI	ggctaagttgttgctttatggtagccgaagttttaagcc
N150Dfor	N150D		gtaccagatgcagcaacaaacttagcc
N150Drev	N150D	KpnI	gttgctgcatctggtagccgaagttttaagcc
KpnI siaP	<i>siaP</i>	KpnI	gcggtaccatgatgaaattgacaaaac
siaPKpnI	<i>siaP</i>	KpnI	gcggtacctggattgattgcttc
siaPup	<i>siaP</i>		gacttctttggcaaacatttctgccgc
siaQMdown	<i>siaQM</i>		cgctttaaccgtattgataatgattaacgc
siaQMup	<i>siaQM</i>		cccacgtaacaaataagagcttggcg
siaPdown	pWKS30		catgattacgccaagcgcgcaattaacc
siaQMrev	pWKS30		gcgcgcgtaatacgaactcactatagggcg
15midDOWN	<i>siaP</i>		cgtagcagcggtagcaagcacaataattc
15midUP	<i>siaQM</i>		ccaccgataattaataacggcggttaagattgcc

Table 2.2 Bacterial strains

<i>E. coli</i> strain	Genotype	Source
DH5 $\alpha$	K-12 F' $\phi$ 80 <i>lacZ</i> $\Delta$ M15 <i>recA1 endA1 gyrA26 thi-1 supE44 relA1 deoR</i> $\Delta$ ( <i>lacZYA-argF</i> )U169	Invitrogen
BL21 (DE3)	F' <i>ompT hsdSB</i> (rB-, mB-) <i>gal dcm</i> (DE3)	Novagen
MC1061	<i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7696 $\Delta$ <i>lacX74 galU galK</i> <sup>-</sup> <i>hsdR2</i> (r <sub>k</sub> - m <sub>k</sub> +) <sup>-</sup> <i>mcrB1 rpsL</i> (F <sup>-</sup> )	(Casadaban & Cohen, 1980)
BW25113	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514</i>	(Datsenko & Wanner, 2000)
BW25113 <i>nanT</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nanT</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>nanA</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nanA</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>nanE</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nanE</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>nanK</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nanK</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>nagA</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nagA</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>nagB</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 nagB</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 <i>ytfQ</i> ::kan	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514 ytfQ</i> ::kan	(Baba <i>et al.</i> , 2006)
BW25113 $\Delta$ <i>nanT</i>	$\Delta$ ( <i>araBAD</i> )567 $\Delta$ <i>lacZ4784</i> (::rrnB-3) $\lambda$ <sup>-</sup> <i>rph-1</i> $\Delta$ ( <i>rhaBAD</i> )568 <i>hsdR514</i> $\Delta$ <i>nanT</i>	

Table 2.3 Plasmid list

Plasmid	Description	Resistance
pET21b	High copy-number expression plasmid	Amp
pGTY3	pET21b with <i>siaP</i> between NdeI and BamHI	Amp
pAH16	pET21b with <i>siaP-His<sub>6</sub></i> between NdeI and XhoI	Amp
pAH35	pET21b with <i>siaP-His<sub>6</sub>:R147A</i> (introduced StuI RE site)	Amp
pAH36	pET21b with <i>siaP-His<sub>6</sub>:R147E</i> (introduced StuI RE site)	Amp
pAH37	pET21b with <i>siaP-His<sub>6</sub>:R147K</i> (introduced StuI RE site)	Amp
pAH41	pET21b with <i>siaP-His<sub>6</sub>:R127A</i> (introduced SfuI RE site)	Amp
pAH42	pET21b with <i>siaP-His<sub>6</sub>:R127E</i> (introduced SfuI RE site)	Amp
pAH57	pET21b with <i>siaP-His<sub>6</sub>:R127K</i> (introduced SfuI RE site)	Amp
pAH44	pET21b with <i>siaP-His<sub>6</sub>:F170A</i> (introduced NcoI RE site)	Amp
pAH45	pET21b with <i>siaP-His<sub>6</sub>:F170W</i> (introduced NcoI RE site)	Amp
pAH46	pET21b with <i>siaP-His<sub>6</sub>:F170Y</i> (introduced NcoI RE site)	Amp
pAH47	pET21b with <i>siaP-His<sub>6</sub>:N187A</i> (introduced SacII RE site)	Amp
pAH48	pET21b with <i>siaP-His<sub>6</sub>:N187D</i> (introduced SacII RE site)	Amp
pAH38	pET21b with <i>siaP-His<sub>6</sub>:F75W</i> (introduced EagI RE site)	Amp
pAH39	pET21b with <i>siaP-His<sub>6</sub>:Y76W</i> (introduced EagI RE site)	Amp
pAH40	pET21b with <i>siaP-His<sub>6</sub>:Y161W</i> (introduced PspLI RE site)	Amp
pAH63	pET21b with <i>siaP-His<sub>6</sub>:F170W;R147A</i> (with NcoI and StuI RE sites)	Amp
pAH64	pET21b with <i>siaP-His<sub>6</sub>:F170W;R147E</i> (with NcoI and StuI RE sites)	Amp
pAH65	pET21b with <i>siaP-His<sub>6</sub>:F170W;R147K</i> (with NcoI and StuI RE sites)	Amp
pAH66	pET21b with <i>siaP-His<sub>6</sub>:F170W;R127A</i> (with NcoI and SfuI RE sites)	Amp
pAH67	pET21b with <i>siaP-His<sub>6</sub>:F170W;R127E</i> (with NcoI and SfuI RE sites)	Amp
pAH68	pET21b with <i>siaP-His<sub>6</sub>:F170W;R127K</i> (with NcoI and SfuI RE sites)	Amp

pAH69	pET21b with <i>siaP-His<sub>6</sub></i> :F170W;N187A (with NcoI and SacII RE sites)	Amp
pAH70	pET21b with <i>siaP-His<sub>6</sub></i> :F170W;N187D (with NcoI and SacII RE sites)	Amp
pAH52	pET21b with <i>siaP-His<sub>6</sub></i> :A66I (introduced XbaI RE site)	Amp
pAH53	pET21b with <i>siaP-His<sub>6</sub></i> :A66M (introduced XbaI RE site)	Amp
pAH54	pET21b with <i>siaP-His<sub>6</sub></i> :A151V (introduced KpnI RE site)	Amp
pAH55	pET21b with <i>siaP-His<sub>6</sub></i> :A11N (introduced KpnI RE site)	Amp
pAH49	pET21b with <i>siaP-His<sub>6</sub></i> :S15D:A195K (with KpnI and SacII RE sites)	Amp
pAH50	pET21b with <i>siaP-His<sub>6</sub></i> :S15K:A195D (with KpnI and SacII RE sites)	Amp
pAH51	pET21b with <i>siaP-His<sub>6</sub></i> :Q72E:A151K (with Kpn2I and KpnI RE sites)	Amp
pAH58	pET21b with <i>siaP-His<sub>6</sub></i> :Q72E:A152K (with Kpn2I and KpnI RE sites)	Amp
pAH84	pET21b with <i>siaP-His<sub>6</sub></i> :S15D:A195D (with KpnI and SacII RE sites)	Amp
pAH85	pET21b with <i>siaP-His<sub>6</sub></i> :S15K:A195K (with KpnI and SacII RE sites)	Amp
pAH56	pET21b with <i>siaP-His<sub>6</sub></i> :N150D (introduced KpnI RE site)	Amp
pWKS30	Low copy number <i>in vivo</i> expression vector	Amp
pES7	pWKS30 with <i>siaPQM</i> between KpnI and BamHI RE sites	Amp
pAH15	pWKS30 with <i>siaP-His<sub>6</sub>-siaQM</i> between KpnI and BamHI with an intragenic XhoI site	Amp
pAH87	pWKS30 with <i>siaP-His<sub>6</sub>:N150D-siaQM</i> between KpnI and BamHI with an intragenic XhoI site	Amp
pET24b	High copy-number expression plasmid	Kan
pJPW4	pET24b with <i>siaP-His<sub>6</sub></i> between NdeI and XhoI	Kan
pAH88	pET24b with <i>siaP-His<sub>6</sub>:N150D</i> (introduced KpnI RE site)	Kan
pFLIPP <sub>i</sub> -260n	pRSET with BamHI- <i>cfp</i> -KpnI- <i>p<sub>i</sub>bp</i> -KpnI- <i>yfp</i> -HindIII	Amp
pFLIP-Neu5Ac	pRSET with BamHI- <i>cfp</i> -PinAI- <i>siaP</i> -KpnI- <i>yfp</i> -HindIII	Amp

## **2.4 General cloning techniques**

### **2.4.1 Agarose gel electrophoresis**

Separation of DNA fragments by electrophoresis was performed using 1% agarose gel in TBE buffer. This buffer was made up with 1.62 g/l Tris (Invitrogen), 2.75 g/l boric acid (Fisher Scientific) and 0.95 g/l EDTA (Fisher Scientific). The gel was submerged in TBE buffer with ethidium bromide in the lower reservoir and DNA separation was performed by applying a 75 volt potential difference for 45 minutes. Ethidium bromide-stained DNA was visualised on a transilluminator (Syngene Imaging System).

### **2.4.2 Plasmid preparation**

Plasmid DNA was prepared initially using a Qiagen kit, before switching to one provided by Machery-Nagel.

### **2.4.3 Polymerase Chain Reaction (PCR)**

PCR was performed in 50 µl reaction volumes with 0.1 x dilution of the polymerase-specific 10 x reaction buffer, 20 pmoles of forward and reverse oligonucleotide primer, 0.2 mM mixed nucleotides, 1 mM MgSO<sub>4</sub> and 1000 x dilution of the template DNA plasmid.

Following a 2 minute, 95 °C denaturation step, the target DNA was amplified by 25 cycles of 94 °C for 1 minute, 55 °C for 1 minute and 72 °C for 1 minute per kilobase of target. Following a final extension step of 5 minutes at 72 °C, the PCR product was removed from the thermal cycler (Techne) and frozen.

### **2.4.4 Site-directed mutagenesis PCR**

#### **2.4.4.1 Mutagenic primer design**

Oligonucleotide primers for site-directed mutagenesis were based on that of Zheng *et al.* (2004). This used 30-50 base primers with a 5' overlap to reduce the primer-primer annealing temperature. The overlap covered a fifteen base region centred on the target codon and each primer was extended in the 3' direction for at least 15 bases or until the expected melting temperature reached 70 °C. Using Webcutter2.0 (<http://users.unimi.it/~camelot/tools/cut2.html>), a silent restriction site was introduced

into one of the overhanging regions that would produce a unique restriction digestion pattern.

#### **2.4.4.2 Mutagenic PCR conditions**

The 50 µl PCR reaction mixture was made up as in Section 2.4.3. Following a 5 minute, 95 °C denaturation step, the whole target plasmid was amplified by 16 cycles of 94 °C for 1 minute, 55 °C for 1 minute and 68 °C for 3 minute per kilobase of target. Following a final extension step of 1 hour at 68 °C, the PCR product was held at 10 °C until it was removed from the thermal cycler and frozen. Before transformation in to competent *E. coli* DH5α, the PCR product was subjected to DpnI restriction enzyme digestion to destroy the template DNA.

### **2.4.5 Megaprimed mutagenic PCR**

#### **2.4.5.1 Megaprimer primer design**

To introduce two mutations at a distance from each other, a technique based on that of (Kirsch & Joly, 1998) was developed. Oligonucleotide primers were designed to amplify the region covering the two mutagenic targets with a silent restriction site introduced to each primer using Webcutter2.0 (<http://users.unimi.it/~camelot/tools/cut2.html>).

#### **2.4.5.2 Megaprimer PCR**

The 50 µl PCR reaction mixture was made up as in Section 2.4.3. Following a 2 minute, 95 °C denaturation step, the target region was amplified by 30 cycles of 94 °C for 1 minute, 55 °C for 1 minute and 68 °C for 2 minutes. Following a final extension step of 15 minutes at 68 °C, the PCR product was held at 10 °C until it was removed from the thermal cycler and frozen. The megaprimer was then purified by 1% agarose gel electrophoresis and extracted from the gel.

#### **2.4.5.3 Megaprimed mutagenic PCR**

Megaprimed mutagenic PCR was performed in 50 µl reaction volumes with 0.1 x dilution of the polymerase-specific 10 x reaction buffer, 360 ng of megaprimer, 0.2 mM mixed nucleotides, 2 mM MgSO<sub>4</sub> and 1000 x dilution of the template DNA plasmid.

Following a 5 minute, 95 °C denaturation step, the whole target plasmid was amplified by 10 cycles of 94 °C for 1 minute, 55 °C for 1 minute and 65 °C for 20 minutes, followed by 10 cycles of 94 °C for 1 minute, 50 °C for 1 minute and 65 °C for 20 minutes. Following a final extension step of 1 hour at 65 °C, the PCR product was removed from the thermal cycler and frozen. Before transformation into competent *E. coli* DH5 $\alpha$ , the PCR product was subjected to DpnI restriction enzyme digestion to destroy the template DNA.

## **2.4.6 Preparation of DNA fragments and cloning**

### **2.4.6.1 Restriction enzyme digestions**

Diagnostic digestions were performed in 20  $\mu$ l reaction volumes with 0.1 x dilution of the optimal 10 x reaction buffer, 1  $\mu$ l of restriction enzymes in total and 2  $\mu$ l plasmid preparation or 5  $\mu$ l PCR product. To prepare cut DNA from downstream use, a 50  $\mu$ l reaction volume with 0.1 x dilution of the optimal 10 x reaction buffer and 1  $\mu$ l of each restriction enzyme was used to digest 20  $\mu$ l plasmid DNA or 45  $\mu$ l PCR product.

### **2.4.6.2 Dephosphorylation reaction conditions**

Digested vector DNA was dephosphorylated in 50  $\mu$ l reaction volumes with 0.1 x dilution of the 10 x reaction buffer and 1  $\mu$ l T4 calf intestinal alkaline phosphatase (CIAP; Promega). This was incubated at 37 °C for 30 minutes before the addition of another 1  $\mu$ l CIAP and 30 minutes incubation.

### **2.4.6.3 Gel extraction**

DNA components were purified by 1% agarose gel electrophoresis and extracted from the gel using a Qiagen kit, before switching to one provided by Machery-Nagel.

### **2.4.6.4 PCR clean up**

To remove contaminating enzymes or buffers, DNA components were cleaned using a Qiagen kit, before switching to one provided by Machery-Nagel.

### **2.4.6.5 Ligation of DNA**

The insert and 50 ng dephosphorylated vector were mixed at a ratio of 5:1 in 15  $\mu$ l specific buffer and incubated overnight at room temperature with 1  $\mu$ l T4 DNA ligase (NEB). This was then transformed directly into competent *E. coli* DH5 $\alpha$ .

## **2.4.7 Transformations**

### **2.4.7.1 Chemically competent *E. coli* DH5 $\alpha$ stock**

*E. coli* DH5 $\alpha$  were inoculated into 250 ml LB and grown at 37 °C with shaking at 160 rpm until OD<sub>650</sub> 0.3–0.4. Growth was stopped by incubation on ice for 1 hour and the cells harvested by centrifugation at 6000 rpm for 10 minutes at 4 °C (Evolution RC; Sorvall). These were resuspended in 50 ml ice cold 50 mM CaCl<sub>2</sub> and incubated on ice for 1 hour. The competent cells were spun down again, resuspended in 8 ml 50 mM CaCl<sub>2</sub> 20% glycerol (Fisher Scientific) and kept at -80 °C in 100  $\mu$ l aliquots.

### **2.4.7.2 Small volumes of chemically competent cells**

The recipient strain of *E. coli* was inoculated into 20 ml LB, with selection where appropriate, grown at 37 °C with shaking at 250 rpm until OD<sub>650</sub> 0.3–0.4 and stopped by incubation on ice for 1 hour. 1.5 ml cells were harvested by centrifugation at 8000 rpm for 1 minute (Mikro20; Hettich Zentrifugen), resuspended in 1 ml ice cold 50 mM CaCl<sub>2</sub> and incubated on ice for 1 hour. The competent cells were spun down again, resuspended in 0.5 ml 50 mM CaCl<sub>2</sub> and split into 100  $\mu$ l aliquots for transformation.

### **2.4.7.3 Heat shock**

An aliquot of chemically competent *E. coli* was mixed with 1–5  $\mu$ l of plasmid preparation or 5–45  $\mu$ l PCR product and incubated on ice for 1 hour. The heat shock was performed at 42 °C for 1.5 minutes before returning the samples to ice. 900  $\mu$ l LB was added to the cell suspension and this was incubated at 37 °C for 1 hour before being spread on selective LB agar.

### **2.4.7.4 Freeze-thaw competency**

The recipient strain of *E. coli* was grown overnight at 37 °C and 500  $\mu$ l of this was inoculated into 5 ml LB before incubating at 37 °C for 1 hour. These were harvested by

centrifugation at 8000 rpm for 1 minute, resuspended in 1 ml ice cold 50 mM CaCl<sub>2</sub> and split into 100 µl aliquots. These were mixed with 2–10 µl of plasmid preparation, frozen in liquid nitrogen and thawed at 37 °C for 2 minutes. 900 µl LB was added to the cell suspension and this was incubated at 37 °C for 1 hour before being spread on selective LB agar.

## **2.5 Growth of bacteria**

### **2.5.1 Expression of periplasmic proteins from pET-based constructs**

*E. coli* BL21 (DE3) pLysS pET21b-*siaP*(variant) were grown in 5 ml LB ampicillin chloramphenicol for 5 hours or overnight at 37 °C with shaking at 250 rpm. These cultures were spun down at 8000 rpm, washed with M9 salts and resuspended in 1 ml M9 salts. This was inoculated into a starter culture of 50 ml M9 minimal medium containing 0.4% D-glucose and grown at 37 °C overnight with shaking at 180 rpm. The starter cultures were inoculated into 625 ml M9 minimal medium containing 0.4% D-glucose to an OD<sub>650</sub> 0.1, grown at 25 °C to an OD<sub>650</sub> 0.2 - 0.3 when they were induced by the addition of 1 mM IPTG (Melford) and incubated overnight at 25 °C.

### **2.5.2 Expression of cytoplasmic proteins from pRSET-based constructs**

*E. coli* BL21 (DE3) pLysS pAH80 (pRSET-*cfp-siaP-yfp*) was grown in 5 ml LB ampicillin chloramphenicol for 5 hours at 37 °C with shaking at 250 rpm. These cultures were spun down at 8000 rpm, washed with M9 salts and resuspended in 1 ml M9 salts. This was inoculated into two flasks of 50 ml M9 minimal medium 0.4% D-glucose to an OD<sub>650</sub> 0.1, grown at 25 °C to an OD<sub>650</sub> 0.2 - 0.3 when they were induced by the addition of 1 mM IPTG (Melford) and incubated for two days in the dark at 25 °C.

### **2.5.3 Expression of membrane proteins from pBAD-based constructs**

*E. coli* MC1061 pBADnQM were grown in 5 ml LB ampicillin for 5 hours at 37 °C with shaking at 250 rpm. These cultures were inoculated into a starter culture of 50 ml LB medium containing ampicillin and 0.5% glycerol and grown at 37 °C overnight with shaking at 180 rpm. 6.25 ml of overnight culture was inoculated into 625 ml LB 0.5% glycerol ampicillin, grown at 37 °C to an OD<sub>660</sub> 1 when they were induced by the

addition of 0.005% arabinose (Calbiochem). These were then incubated for a further two hours at 37 °C before being harvested by centrifugation at 5000 rpm for 15 minutes at 4 °C. The pelleted cells were resuspended in 50 mM potassium phosphate buffer 20% glycerol pH7.8 and kept at -80 °C.

## **2.6 Preparation of bacterial extracts**

### **2.6.1 Periplasmic fraction preparation**

After overnight growth, induced cells were pelleted at 6000 rpm for 20 minutes at 4 °C and resuspended in 25 ml 5 mM EDTA, 50 mM Tris, 0.5 M sucrose (Fisher Scientific) pH 8.0 and incubated with 12 mg lysozyme (chicken egg white; Sigma) at 30 °C for 2 hours. This was then centrifuged at 12000 rpm for 10 minutes at 4 °C and the supernatant collected.

### **2.6.2 Cytoplasmic protein recovery**

Induced cells were harvested by centrifugation at 4000 rpm for 20 minutes at 4 °C, resuspended in 35 ml 50 mM Tris/HCl pH 8.0 and sonicated on ice for a total of 10 minutes to rupture the cells. Insoluble protein and cell debris were removed by centrifugation at 12000 rpm for 10 minutes at 4 °C and the supernatant collected.

### **2.6.3 Preparation of cell membranes as total membrane vesicles**

MgCl<sub>2</sub> (Fisher Scientific) and DNase (Sigma-Aldrich) were added to the harvested *E. coli* cells to a final concentration of 1 mM and 100 µg/ml, respectively. The cells were ruptured by sonication on ice for a total of 10 minutes and then incubated on ice for a further 5 minutes before the addition of 5 mM EDTA pH 7.5. The lysate was centrifuged at 10000 rcf for 30 minutes at 4 °C and the supernatant spun again at 40000 rpm for 1 hour at 4 °C (L7 Ultracentrifuge; Beckman). The pellet was resuspended in 11 ml 50 mM potassium phosphate buffer 10% glycerol pH 7.8 and kept at -80 °C.

## **2.7 Protein purification techniques**

### **2.7.1 Hydrophobic Interaction Chromatography (HIC)**

The periplasmic fraction preparation was dialysed into 50 mM Tris/HCl, 1.5 M  $(\text{NH}_4)_2\text{SO}_4$  pH 8.0 and clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. A Tricorn 10/300 column packed with Source 15PHE (Amersham Biosciences) was connected to a BioLogic DuoFlow FPLC (BioRad) and washed with 50 mM Tris/HCl, 1.5 M  $(\text{NH}_4)_2\text{SO}_4$  pH 8.0. The clarified, dialysed periplasmic preparation was loaded onto the column at 0.5 ml/minute. The separation was performed by decreasing the  $(\text{NH}_4)_2\text{SO}_4$  concentration to zero over 3 hours with a flow rate of 1 ml/minute and fractions collected for 2 minutes (Fraction Collector model 2128; BioRad). Protein-containing fractions were visualised using SDS PAGE.

### **2.7.2 Size Exclusion Chromatography (SEC)**

Following HIC, the SiaP-containing fractions were dialysed into 50 mM Tris/HCl, 150 mM NaCl pH 8.0 and clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. A G75 sepharose column was connected to an AKTA UPC-900 FPLC (Amersham Pharmacia Biotech) and washed with 10 column volumes (CV) of 50 mM Tris/HCl, 150 mM NaCl pH 8.0. The clarified, dialysed SiaP-containing preparation was loaded onto the column in 5 ml aliquots at 0.45 ml/minute. The separation was performed over 5 hours, allowing 1.5 CV to be passed. The fraction volume collected was 2.5 ml minutes (Fraction Collector Frac-950; Amersham Pharmacia Biotech). Protein-containing fractions were visualised using SDS PAGE.

### **2.7.3 Nickel-affinity chromatography**

#### **2.7.3.1 FPLC**

The periplasmic fraction preparation was dialysed into 20 mM Tris/HCl, 300 mM NaCl pH 7.5 and clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. A 1 ml HisTrap HP column containing Ni-sepharose (GE Healthcare) was connected to a BioLogic DuoFlow FPLC (BioRad) and washed with 20 mM Tris/HCl, 300 mM NaCl, 12 mM imidazole pH 8.0. Imidazole was added to the clarified, dialysed periplasmic preparation to a concentration of 12 mM and this was loaded onto the column at 1 ml/minute. The column was then washed with 20 column volumes (CV) 20 mM Tris/HCl, 300 mM NaCl, 20 mM imidazole pH 7.5. His-tagged protein was eluted with

20 mM Tris/HCl, 300 mM NaCl, 400 mM imidazole pH 7.5 and 1 ml fractions were collected (Fraction Collector model 2128; BioRad). Protein-containing fractions were visualised using SDS PAGE.

#### **2.7.3.2 Peristaltic pump**

The periplasmic fraction preparation was dialysed into 20 mM Tris/HCl, 300 mM NaCl pH 7.5 and clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. A 1 ml HisTrap HP column containing Ni-sepharose (GE Healthcare) was connected to a P-1 peristaltic pump (Amersham Biosciences) and washed with 20 mM Tris/HCl, 300 mM NaCl, 12 mM imidazole pH 8.0. Imidazole was added to the clarified, dialysed periplasmic preparation to a concentration of 12 mM and this was loaded onto the column at a medium flow rate. The column was then washed with 20 column volumes (CV) 20 mM Tris/HCl, 300 mM NaCl, 20 mM imidazole pH 7.5. The pump was then filled with 20 mM Tris/HCl, 300 mM NaCl, 400 mM imidazole pH 7.5, reconnected to the column and eluted protein was collected as 0.5 CV, followed by 2 CV fractions. Protein-containing fractions were visualised using SDS PAGE.

#### **2.7.4 Nickel-Nitrilotriacetic acid (Ni-NTA) resin – FLIP**

##### **2.7.4.2 Purification of soluble protein**

The sample containing Histidine-tagged protein was dialysed into 20 mM Tris/HCl, 200 mM NaCl pH 7.5 (in the dark, where necessary) and clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. A High-imidazole buffer containing 20 mM Tris/HCl, 200 mM NaCl, 400 mM imidazole pH 7.5 was added to the clarified lysate to give a final imidazole concentration of 12 mM. This was mixed with 500 µl pre-equilibrated Ni-nitrilotriacetic acid (Ni-NTA) resin (Qiagen) per 25 ml of sample and mixed for 1 hour at 4 °C (in the dark, where necessary). Upto 0.75 ml of protein-bound resin was sedimented in a disposable column (Qiagen), washed with 10 column volumes (CV) of 20 mM Tris/HCl, 200 mM NaCl, 20 mM imidazole pH 7.5 and then the protein was eluted with 0.5, 2 and 2 CV of 20 mM Tris/HCl, 200 mM NaCl, 400 mM imidazole pH 7.5. The second fraction was at 4 °C (in the dark, where necessary).

#### **2.7.4.2 Purification of membrane protein**

Approximately 10 mg of SiaQM in *E. coli* total membrane vesicles was mixed in 7 ml 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 10 mM imidazole, 0.5% dodecyl-D-maltoside (DDM) pH 7.8 and incubated on ice for 30 minutes. This was centrifuged at 53000 rpm for 20 minutes at 4 °C, the supernatant mixed with 300 µl Ni-nitrilotriacetic acid (Ni-NTA) resin (Qiagen) and mixed for 1 hour at 4 °C. After the resin was sedimented in a disposable column (Qiagen), the column was washed with 10 column volumes (CV) of 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 40 mM imidazole pH7.8 containing 0.05% DDM, followed by 10 CV of the same buffer containing 0.15% decyl-D-maltopyranoside (DM). SiaQM was then eluted with 0.5, 2 and 2 CV of DM-containing buffer with 500 mM imidazole.

#### **2.7.5 Reconstitution of membrane protein into proteoliposomes**

200 µg of SiaQM was then mixed with 8 mg *E. coli* lipid in 50 mM potassium phosphate buffer 0.15% DM pH 7.0, incubated on ice for 10 minutes and reconstituted into proteoliposomes by rapid dilution in 65 ml 50 mM potassium phosphate buffer pH 7.0 on ice. The Proteoliposomes were then collected by centrifugation at 40000 rpm for 1 hour 30 minutes at 4 °C, resuspended in 0.5 ml 50 mM potassium phosphate buffer pH 7.0 and kept at -80 °C. As the requirements for reconstituted SiaQM increased, this was supplied by Judith Hawkshead.

### **2.8 Large volume fermentation for the production of SiaP-His<sub>6</sub>:A11N**

#### **2.8.1 Expression of protein from the pAH55 construct**

*E. coli* BL21 (DE3) pLysS pAH55 was grown in 20 ml LB ampicillin chloramphenicol for overnight at 37 °C with shaking at 250 rpm. This culture was spun down at 4000 rpm for 10 minutes at 4 °C and resuspended in 10 ml M9 salts. This was inoculated into a starter culture of 800 ml M9 minimal medium containing 0.6% D-glucose and grown at 37 °C overnight with shaking at 180 rpm. A 70 l fermenter with CBC10 CH-8604 control unit ( $\alpha$ -ALPHA-LAVAL CHEMAP AG) was set up with 50 l M9 minimal medium supplemented with 0.05% D-glucose, 0.5% glycerol, 0.2%  $\alpha$ -lactose (Sigma) and 0.01%

antifoam 204 (Sigma). The starter culture was inoculated directly into this and grown at 30 °C for 32 hours with O<sub>2</sub> saturation kept above 20% by a variable stirring rate.

### **2.8.2 Preparation of bacterial extracts**

Induced cells were pelleted at 5000 rpm for 15 minutes and resuspended in 250 ml 5 mM EDTA, 50 mM Tris, 0.5 M sucrose (Fisher Scientific) pH 8.0. Glycerol was added to a final concentration of 10% and the mixture was frozen and kept overnight at -20 °C. Following thawing, this was incubated with 0.5 mg/ml lysozyme at 30 °C for 4 hours and was then centrifuged at 14000 rpm for 30 minutes at 4 °C. The supernatant was collected and kept at 4 °C before purification of the soluble protein using Ni-NTA resin.

## **2.9 Polyacrylamide gel electrophoresis (PAGE)**

### **2.10.1 Native conditions**

#### **2.9.1.1 Buffers and gel**

The separating polyacrylamide gel was cast from 6% acrylamide (Geneflow), 375 mM Tris/HCl, 0.1% ammonium persulphate (APS; Sigma), 0.01% (TEMED; Sigma) pH 8.8, while the stacking gel was cast from 4.5% acrylamide, 125 mM Tris/HCl, 0.1% APS, 0.01% TEMED pH 6.8. The running buffer in the tank was made up of 3 g/l Tris, 14 g/l glycine (Fisher Scientific).

#### **2.9.1.2 Sample preparation**

Each protein was resuspended in sample buffer, which was made up of 188 mM Tris/HCl, 10% glycerol, 0.0005% bromophenol blue (Sigma-Aldrich) pH 8.8.

#### **2.9.1.3 Running conditions**

The migration of proteins under native conditions electrophoresis was performed using 6% polyacrylamide gel. Each sample was loaded so that 250 ng of protein was added to each lane. The gel was submerged in running buffer and a 60 volt potential difference was applied for 3.5 hours.

### **2.9.2 Denaturing conditions**

### **2.9.2.1 Buffers and gel**

The separating polyacrylamide gel was cast from 12% acrylamide, 375 mM Tris/HCl, 0.1% sodium dodecyl sulphate (SDS; Melford), 0.1% APS, 0.01% TEMED pH 8.8, while the stacking gel was cast from 4% acrylamide, 125 mM Tris/HCl, 0.1% SDS, 0.1% APS, 0.01% TEMED pH 6.8. The running buffer in the tank was made up of 3 g/l Tris, 14 g/l glycine and 1 g/l SDS.

### **2.9.2.2 Sample preparation**

The sample buffer was made up of 18 mM Tris/HCl, 20% glycerol, 0.2% SDS, 0.001% Brilliant Blue-R250 (Fisher Bioreagents) pH 7.2. Each protein was resuspended in this buffer and heated to 97.5 °C for 5 minutes. For whole cell lysates, 1 ml of culture was spun down at 13000 rpm for 5 minutes and resuspended in a volume of sample buffer equal to 50 x OD<sub>650</sub> of the culture. These were then heated to 97.5 °C for 5 minutes, clarified by centrifugation at 13000 rpm for 15 minutes and samples removed from the supernatant.

### **2.9.2.3 Running conditions**

The separation of proteins under denaturing conditions by electrophoresis was performed using 12% polyacrylamide gel. Once the gel was submerged in running buffer, the samples were loaded and a 200 volt potential difference was applied for 1 hour. When finished, the gels were rinsed with distilled water.

### **2.9.3 Staining/destaining**

Coomassie stain was made up of 2.5% Brilliant Blue-R250 in 10% acetic acid (Fisher Scientific), 45% methanol (Fisher Scientific). Polyacrylamide gels were submerged in this stain for 10 minutes with gentle agitation. The gels were destained in 10% acetic acid, 10% ethanol before an image was captured using the Syngene Imaging System.

### **2.9.4 Western blotting**

#### **2.9.4.1 Buffers**

The transfer buffer was made up of 200 mM glycine, 25 mM Tris/HCl, 10% methanol pH 8.3. TBST buffer was made up of 150 mM NaCl, 20 mM Tris/HCl, 0.1% Tween-20 (Sigma) pH 8.3. For the blocking solution, 5% milk powder (Marvel) was added to TBST. The primary hybridization buffer was made up of a 1:2000 dilution of mouse anti(tetrahistidine) antibodies (Qiagen) in TBST 1% milk powder. The secondary hybridization buffer was made up of a 1:2000 dilution of horseradish peroxidase-conjugated rabbit anti(mouse) antibodies (Zymed Invitrogen) in TBST 1% milk powder.

#### **2.9.4.2 Running conditions**

The polyacrylamide gel was removed from the PAGE tank and equilibrated in transfer buffer for 15 minutes. Once fitted into the transfer cassette and submerged in transfer buffer, the proteins were transferred to nitrocellulose membrane (Amersham Biosciences) by a 0.35 ampere current for 1 hour. Following this, the membranes were stained with Ponceau S solution (Sigma) to confirm the transfer of proteins.

#### **2.9.4.3 Sample visualisation**

Protein-bound nitrocellulose membranes were incubated in blocking solution for 30 minutes at room temperature with gentle agitation. Following four washing steps with TBST, the membrane was incubated with the primary hybridisation buffer over night at 4 °C. The next morning, the membrane was washed four times with TBST, incubated with the secondary hybridisation buffer for 1 hour at room temperature and immediately washed seven times with TBST. For visualisation, the membrane was mixed with 5 ml Supersignal West Pico chemiluminescent substrate (Thermo Scientific), dried and exposed to CL-Xposure photographic film (Thermo Scientific).

### **2.10 Fluorescence spectroscopy**

Fluorescence spectroscopy was performed in either 50 mM Tris/HCl pH 8.0 or 20 mM phosphate buffer of the specified pH value, made up from the sodium or potassium salts as indicated.

Purified protein was used at a concentration dependent upon its affinity for sialic acid. Where the  $K_d$  value was less than 2  $\mu\text{M}$ , the protein concentration was 0.05  $\mu\text{M}$ ; for  $K_d$

values above this but less than 10  $\mu\text{M}$ , the protein concentration used was 0.25  $\mu\text{M}$ ; when the  $K_d$  value was indeterminate, the protein concentration used was 1  $\mu\text{M}$ . Protein fluorescence experiments were performed in 3 ml quartz cuvettes (Starna) using a FluoroMax2 (Instruments SA, Inc.) with an LTD6 waterbath (Grant) or a FluoroMax4 (Horiba Jobin Yvon) with a NesLab RTE water bath (Thermo Scientific). Both the FluoroMax2 and the FluoroMax4 were controlled with the supplied software, DataMax-Std version 2.20 and FluorEssence version 3.0.0.19, respectively. The sample was excited at the specified wavelength with slit widths of 2 – 5 nm in order to give a signal intensity of 2 – 3  $\times 10^6$  units. Ligand was added at specified concentrations and times to produce spectra and time-course titrations as described. For titrations, the cumulative fluorescence change was plotted using SigmaPlot (version 10.0) and the  $K_d$  value was determined using a fit to a simple hyperbolic curve.

## **2.11 Isothermal Titration Calorimetry (ITC)**

### **2.11.1 ITC protocol**

For this analysis the ligand was dissolved in a volume of the dialysis buffer from the protein preparation. Both the protein and ligand were degassed at 2 °C below the experimental temperature before being loaded into the VP-ITC microcalorimeter (MicroCal), controlled by VPViewer2000 version 1.4.24 (MicroCal LLC). 10 – 14.6  $\mu\text{M}$  SiaP-His<sub>6</sub> was titrated with additions of 100 – 200  $\mu\text{M}$  sialic acid. Initially, the titration pattern was a first injection of 3  $\mu\text{l}$ , followed by 6  $\mu\text{l}$  injections of 14 seconds, separated by 180 seconds. To reduce running time, this pattern was modified to a first injection of 3  $\mu\text{l}$ , three 18  $\mu\text{l}$  injections separated by 240 seconds, followed by 6  $\mu\text{l}$  injections of 14 seconds, separated by 180 seconds. These were then analysed using Origin 7SR2 version 7.0383(B383) (OriginLab Corporation) by fitting to one set of binding sites.

### **2.11.2 ITC for SiaP-His<sub>6</sub>:F170W**

For SiaP-His<sub>6</sub>:F170W, 40  $\mu\text{M}$  protein was titrated with additions of 600  $\mu\text{M}$  sialic acid. The titration pattern was a first injection of 3  $\mu\text{l}$ , followed by 6  $\mu\text{l}$  injections of 14 seconds, separated by 180 seconds.

### **2.11.3 ITC for higher $K_d$ values**

For  $K_d$  values above 2  $\mu\text{M}$ , the protein concentration used was ten times the expected  $K_d$  value and the concentration of ligand used was ten times the protein concentration. The titration pattern was a first injection of 3  $\mu\text{l}$ , followed by 6  $\mu\text{l}$  injections of 14 seconds, separated by 180 seconds.

## **2.12 Circular Dichroism (CD)**

### **2.12.1 CD spectra determination**

CD spectra were determined using a J-810 spectropolarimeter (Jasco) controlled by the supplied software SpectraManager version 1.53.00 (Jasco); the temperature was maintained at 20.0  $^{\circ}\text{C}$  using a Peltier unit PFD-425S (Jasco). Protein was dialysed into 10 mM potassium phosphate buffer pH 8.0 and diluted to a concentration of 5  $\mu\text{M}$ . The spectra were determined in a 1 mm pathlength quartz cuvette (Starna) between 240 – 180 nm at 100 nm/minute with 1 nm pitch.

### **2.12.2 Thermal stability by CD**

The melting temperature was determined by measurement of the CD signal at 222 nm every 5  $^{\circ}\text{C}$  between 20 – 90  $^{\circ}\text{C}$ . The rate of temperature increase was set at 3  $^{\circ}\text{C}/\text{minute}$  and the sample was incubated at the new temperature for 5 minutes before each measurement.

## **2.13 *In vivo* growth assays**

### **2.13.1 Growth on solid medium**

For growth experiments on solid medium, the *E. coli* BW25113  $\Delta nanT$  strains carrying the plasmids of interest were grown overnight in 5 ml LB with selection and then streaked onto M9 minimal medium 1% agarose with 1 mg/ml sialic acid or 0.4% glucose as the sole carbon source. These were then incubated at 37  $^{\circ}\text{C}$  for two days.

### **2.13.2 Liquid culture**

The *E. coli* BW25113  $\Delta nanT$  strains carrying the plasmids of interest were grown overnight in 5 ml LB with selection. These cells were harvested by centrifugation at 8000 rpm for 1 minute, washed M9 salts by resuspension and centrifugation and then 10  $\mu\text{l}$  was

inoculated into 3 ml M9 minimal medium with 1 mg/ml sialic acid as the sole carbon source. These were then incubated at 37 °C overnight.

### **2.13.3 Prototype incubated plate shaker – buffers etc and pre-growth**

#### **2.13.3.1 24-well plate set up**

Cultures of the *E. coli* strains of interest were grown to mid-log phase and then washed and resuspended in M9 salts. These were then inoculated to an OD<sub>650</sub> 0.01 in 700 µl M9 minimal medium in the wells of a sterile 24-well glass-bottom plate. The wells were closed by an air-permeable sterile lid, which was taped onto the top of the plate.

#### **2.13.3.2 The prototype incubated plate shaker**

Culture growth was monitored using a prototype incubated plate shaker (EnzyScreen). The 24-well plate was clamped into the shaker cabinet and incubated at 35 °C with shaking at 250 rpm. The shaking was halted every 30 minutes for about 1 minute while the density of the culture was measured. A flat-bed scanner captured an image of the base of the plate and the increasing whiteness of the growing cultures was converted to a value (G value) by the associated software (ImageAnalysisGIU version 1.0.0.0).

#### **2.13.3.3 Standard curve**

The G value was correlated to an OD<sub>650</sub> value (OD<sub>650</sub><sup>\*</sup>) by measuring the G value of non-growing cultures of known optical density (Figure 2.1). Below a G value of 57, the relationship is linear; above this value,  $y$  is approximately equal to  $104/(1 + OD_{650}(-(x - 0.289)/0.095))$ . Using this, all G values were converted to their apparent OD<sub>650</sub> values (OD<sub>650</sub><sup>\*</sup>).

## **2.14 ELISA-based assays for the detection of SiaP-His<sub>6</sub>**

### **2.14.1 Target substrates**

BSA-Neu5Ac was prepared using a method based on that used by Telmer and Shilton (2003) to produce BSA coupled to Amylose and Maltoheptaose. 50 mg Neu5Ac was dissolved in 50 ml ddH<sub>2</sub>O and activated by the addition of 50 mg NaIO<sub>4</sub> (s) followed by incubating on ice for 30 minutes. 20 ml 0.05 mg/ml BSA 200 mM Na<sub>2</sub>CO<sub>3</sub> pH 9.1 was added to the activated Neu5Ac and stirred at room temperature for six hours. This was then dialysed four times against 5 l ddH<sub>2</sub>O for 48 hours at 4 °C, freeze dried and stored at 4 °C.

### **2.14.2 96-well plate assay**

5 nanomoles (nmoles) and 0.5 nmoles of sialic acid, Neu5Ac-BSA, colominic acid and porcine mucin (submaxillary gland; Sigma) were added to columns of 8 wells on a 96-well MaxiSorp plate (Nunc) and left open at 37 °C to dry. The plate was washed with 140 mM NaCl, 2.7 mM KCl, 0.1% Tween-20, 8.1 mM sodium phosphate, 1.5 mM potassium phosphate pH 7.3 (TPBS). The remaining well surface was blocked by incubating with 50 µl TPBS 2% BSA at 4 °C for two days. The plate was washed again in TPBS. 50 µl 20 µM SiaP-His<sub>6</sub> was added across each 12-well row in a 1/3 serial dilution to the lowest concentration of 27 nM in the 7<sup>th</sup> row. These were incubated at room temperature for 3.5 hours, washed with TPBS and incubated with 1/2000 mouse anti(His<sub>4</sub>)antibodies overnight at 4 °C. Unbound antibodies were washed off with TPBS and the wells were incubated with 1/2000 HRP-conjugated rabbit anti(mouse)antibodies at room temperature for 1.5 hours. 50 µl Supersignal (Pierce) was added to each well and the resultant luminescence was observed at 425 nm using a PolarStar Optima plate reader (BMG Labtech).

## **2.15 <sup>14</sup>C-radiolabelled sialic acid-based assays**

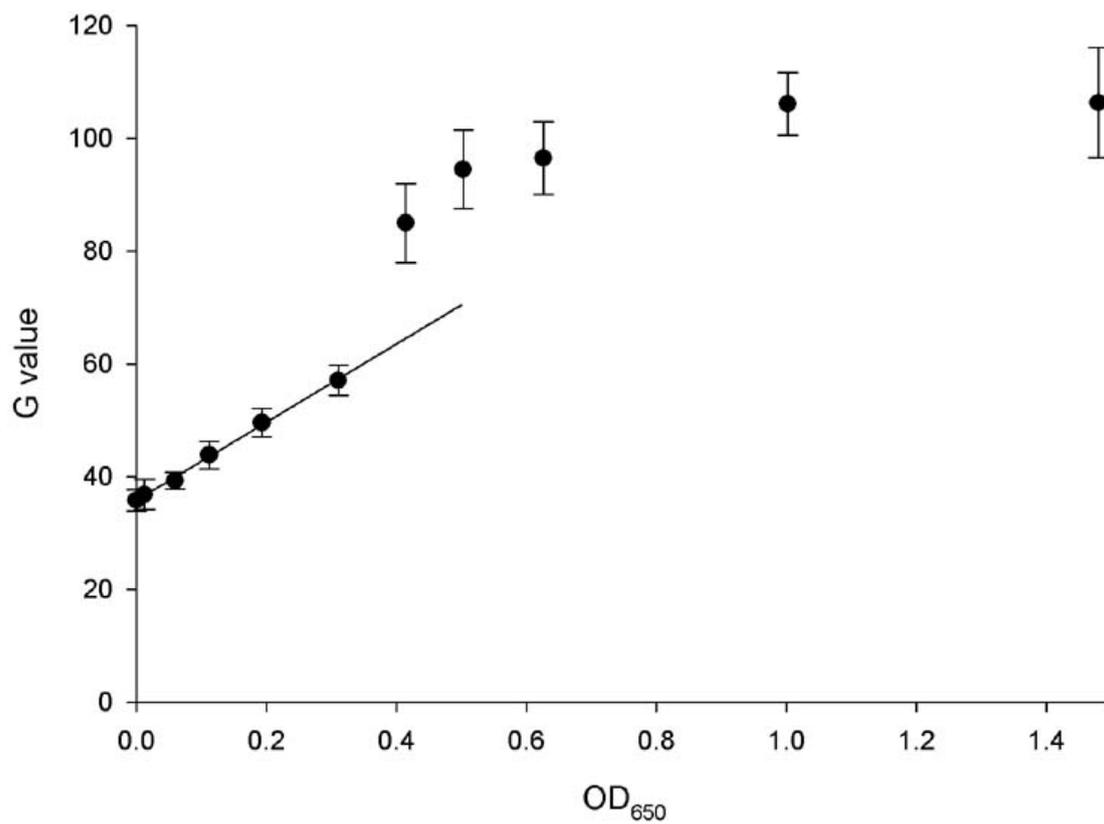
### **2.15.1 Filter binding assay**

Protein at 2.5 µM was incubated with 2.5 µM <sup>14</sup>C-labelled sialic acid (<sup>14</sup>C-Neu5Ac; Sigma-Aldrich) in 50 µl 50 mM Tris/HCl pH 8.0 for 10 minutes on ice. The protein was precipitated by incubation on ice for 20 minutes with 1 ml saturated (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. Precipitated protein was trapped on 0.22 µm nitrocellulose filter (Sigma-Aldrich), washed with 2 ml 50 mM Tris/HCl pH 8.0 and added to 3 ml scintillation fluid (Perkin Elmer) for measuring.

### **2.15.2 In vitro <sup>14</sup>C-sialic acid transport assay**

#### **2.15.2.1 Preparation of SiaQM-containing proteoliposomes**

Proteoliposomes containing 200 µg SiaQM were centrifuged at 53000 rpm for 20 minutes at 4 °C (TLA100 rotor, TL-100 Ultracentrifuge; Beckman) then resuspended in 0.5 ml 100 mM potassium acetate, 2 mM MgSO<sub>4</sub>, 20 mM potassium phosphate pH 7.0 and



**Figure 2.1:** Standard curve for the relationship between OD<sub>650</sub> and G value for *E. coli* BW25113 in M9 minimal medium. Below a G value of 57, the relationship is linear; above this value, y is approximately equal to  $104/(1 + OD_{650} \cdot (x - 0.289)/0.095)$ .

extruded 11 times through a 400 nm polycarbonate filter (Avestin) to make a homogenous suspension of proteoliposomes containing this Na<sup>+</sup>-free buffer. These extruded proteoliposomes were centrifuged at 53000 rpm for 20 minutes at 15 °C and resuspended in 50 µl of the same buffer.

#### **2.15.2.2 Buffers for gradients**

The reaction buffer for the uptake assay was made up of 100 mM sodium acetate, 2 mM MgSO<sub>4</sub>, 20 mM sodium PIPES pH 7.0, 1 µM valinomycin. This combination of components on the outside and inside of the proteoliposomes set up a sodium-inside gradient.

#### **2.15.2.3 Standard assay**

For each uptake assay, 5 µM protein and 5 µM <sup>14</sup>C-Neu5Ac were added to 300 µl of the reaction buffer and incubated at the reaction temperature of 30 °C for 1 minute. 1.15 µM SiaQM in proteoliposomes (6 µl addition) was added to start the reaction and 50 µl samples were taken twenty seconds later and every forty seconds following, up to 180 seconds. As each of these was removed, they were mixed with 50 µl 100 mM sodium acetate, 2 mM MgSO<sub>4</sub>, 20 mM sodium PIPES, 1 mM unlabelled sialic acid pH 7.0. After 10 seconds, these were added to a 0.22 µm nitrocellulose filter, washed with 2 ml 50 mM potassium phosphate buffer pH 7.0 and added to 3 ml scintillation fluid for measuring.

#### **2.15.2.4 Altered <sup>14</sup>C-sialic acid concentration**

Doubling the volume of <sup>14</sup>C-Neu5Ac added to the reaction buffer increased the sialic acid concentration to 10 µM. To reach a reaction concentration of 30 µM sialic acid, 5 µM <sup>14</sup>C-Neu5Ac was mixed with 25 µM unlabelled sialic acid.

#### **2.15.2.5 Competition assay**

This assay used 8.12 µM <sup>14</sup>C-Neu5Ac and 11.88 µM unlabelled Neu5Ac to reach a reaction concentration of 20 µM sialic acid. The concentration of valinomycin was doubled to 2 µM and 5 µM SiaP-His<sub>6</sub> was included in the reaction buffer. For the competition assay, 5 µM SiaP-His<sub>6</sub>:N150D was added, giving a maximum total protein

concentration of 10  $\mu\text{M}$ . The reaction was started by the addition of 2.3  $\mu\text{M}$  SiaQM in proteoliposomes (12  $\mu\text{l}$  addition) and 50  $\mu\text{l}$  samples were taken 30 seconds later and every 60 seconds following, up to 270 seconds.

## **2.16 Guanidine hydrochloride (GnHCl) denaturation for protein recycling**

### **2.16.1 Buffers**

The protein solution was made up to 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 10 mM imidazole pH 7.5 from higher concentration stocks. Guanidine hydrochloride (GnHCl), was added to 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 20 mM imidazole pH 7.5 to give wash buffers containing 2 M, 1.5 M, 1 M and 0.5 M GnHCl. The elution buffer was made up of 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 500 mM imidazole pH 7.5.

### **2.16.2 In-column denaturation and refolding**

The protein in imidazole-containing buffer was clarified by centrifugation at 12000 rpm for 10 minutes at 4 °C. This was mixed with 500  $\mu\text{l}$  pre-equilibrated Ni-NTA resin and mixed for 1 hour. After the resin was sedimented in a disposable column, the bound protein was denatured with 30 column volumes (CV) of buffer containing 2 M GnHCl, followed by 4 CV of each of the solutions containing decreasing GnHCl. The column was washed with 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 20 mM imidazole pH 7.5 and then the protein was eluted with 0.5, 2 and 2 CV of 50 mM potassium phosphate buffer, 200 mM NaCl, 20% glycerol, 500 mM imidazole pH 7.5.

## Chapter Three

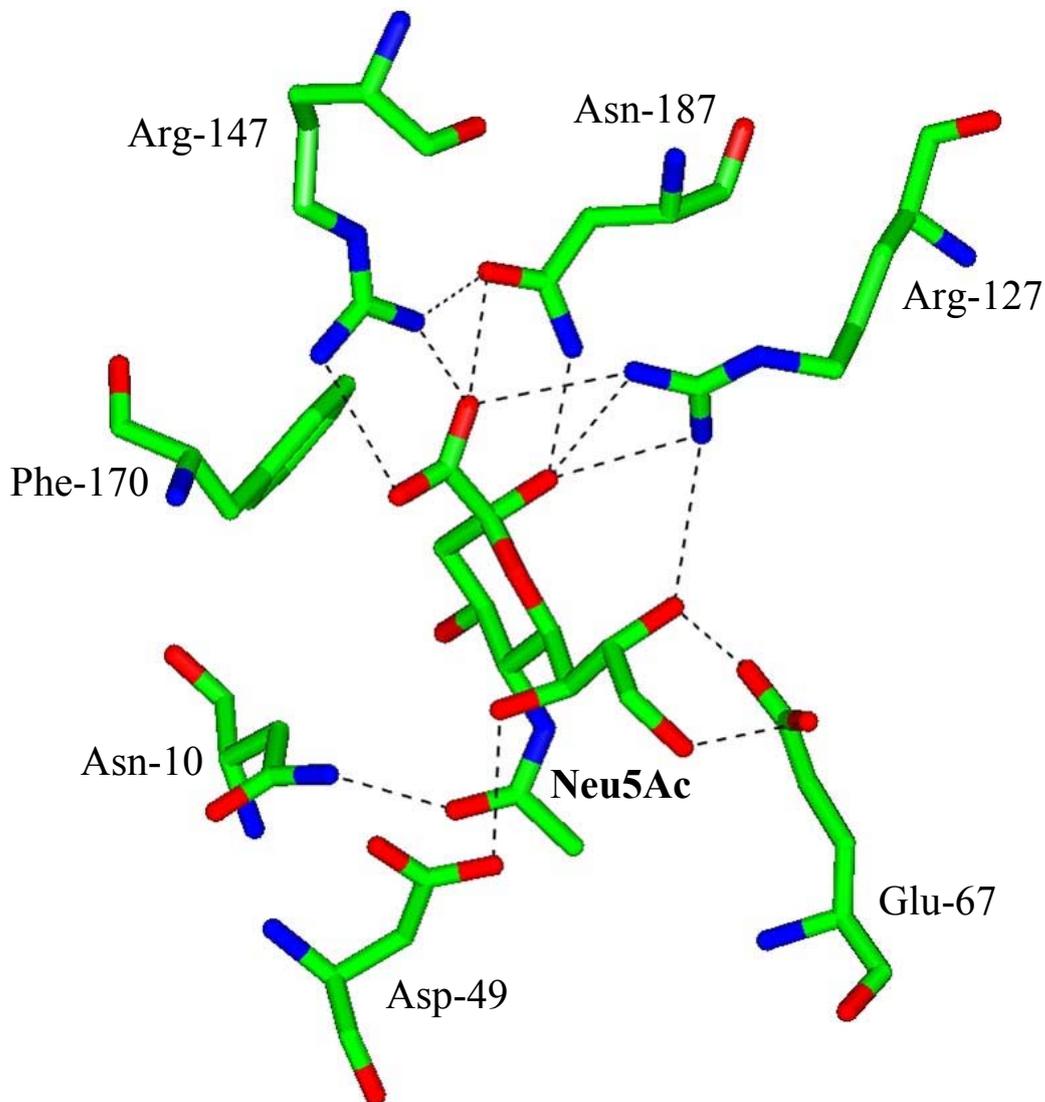
Examination of the contribution of individual amino acid residues to high-affinity sialic acid binding by SiaP

### **3.1 Strategy for site-directed mutagenesis of the *siaP* gene to alter sialic acid-binding properties of SiaP**

From the known structures of sialic acid binding proteins, such as bacterial sialidases and eukaryotic siglecs, the most common method used to coordinate monomeric sialic acid is bivalent salt bridging with an arginine residue (Angata *et al.*, 2004, Morschhauser *et al.*, 1990, van der Merwe *et al.*, 1996, Vinson *et al.*, 1996). As mentioned previously, the DctP-like TRAP SBPs invariably use their conserved binding site arginine to co-ordinate carboxylate-containing ligands and it was expected that the corresponding arginine (Arg-147) would be a residue involved in the interaction with sialic acid in SiaP. Due to the potential importance of sialic acid carboxylate stabilization, not only Arg-147, but also Arg-127 and Asn-187 were chosen as targets for mutation (Figure 3.1). The arginine residues were mutated to three different amino acids: the first was a conservative change to lysine; the second a charge swap to glutamate; and the third a truncation of the side chain using alanine (Table 3.1). The asparagine was replaced by alanine or aspartate. In the crystal structures of SiaP in the closed and open conformations, Phenylalanine-170 shows an additional movement and rotation that projects it into the closed binding cleft in close proximity with the ligand (Muller *et al.*, 2006, Johnston *et al.*, 2008). This suggests that this residue could be acting as a lid for the binding site and so it was truncated to alanine, given a polar side chain as tyrosine or increased in size as tryptophan.

### **3.2 Introduction of Arginine-147 mutants into native SiaP**

A method to create the selected mutations in *siaP* was needed and we achieved rapid and extremely efficient production of site-directed mutants using a technique based on that of Zheng *et al.* (2004). Briefly, this uses 30–50 base primers with a 5' overlap of about half their length to reduce the primer-primer annealing temperature and a silent mutation to add a restriction endonuclease cleavage site to allow for screening of potential mutants (Zheng *et al.*, 2004). The long overhang allows the primers to anneal across the opposite end of the plasmid PCR product, so introducing PCR-like properties to the reaction. This improves the yield by allowing amplification of the entire plasmid and the high melting temperature of the overhang region removes the need for a plasmid



**Figure 3.1:** Representation of the interactions between Neu5Ac and the binding site of SiaP, based on 3B50.pdb (Johnston et al., 2008). Amino acid residues and the sialic acid ligand are shown as atom coloured cylinders; hydrogen bonds are shown as dashed black lines; Phenylalanine-65 and Alanine-151 are not shown for clarity.

**Table 3.1:** Target residues in SiaP and their chosen replacements.

Target residue	Amino acid change
Arginine-127	Alanine
	Glutamate
	Lysine
Penylalanine-170	Alanine
	Tryptophan
	Tyrosine
Arginine-147	Alanine
	Glutamate
	Lysine
Asparagine-187	Alanine
	Aspartate

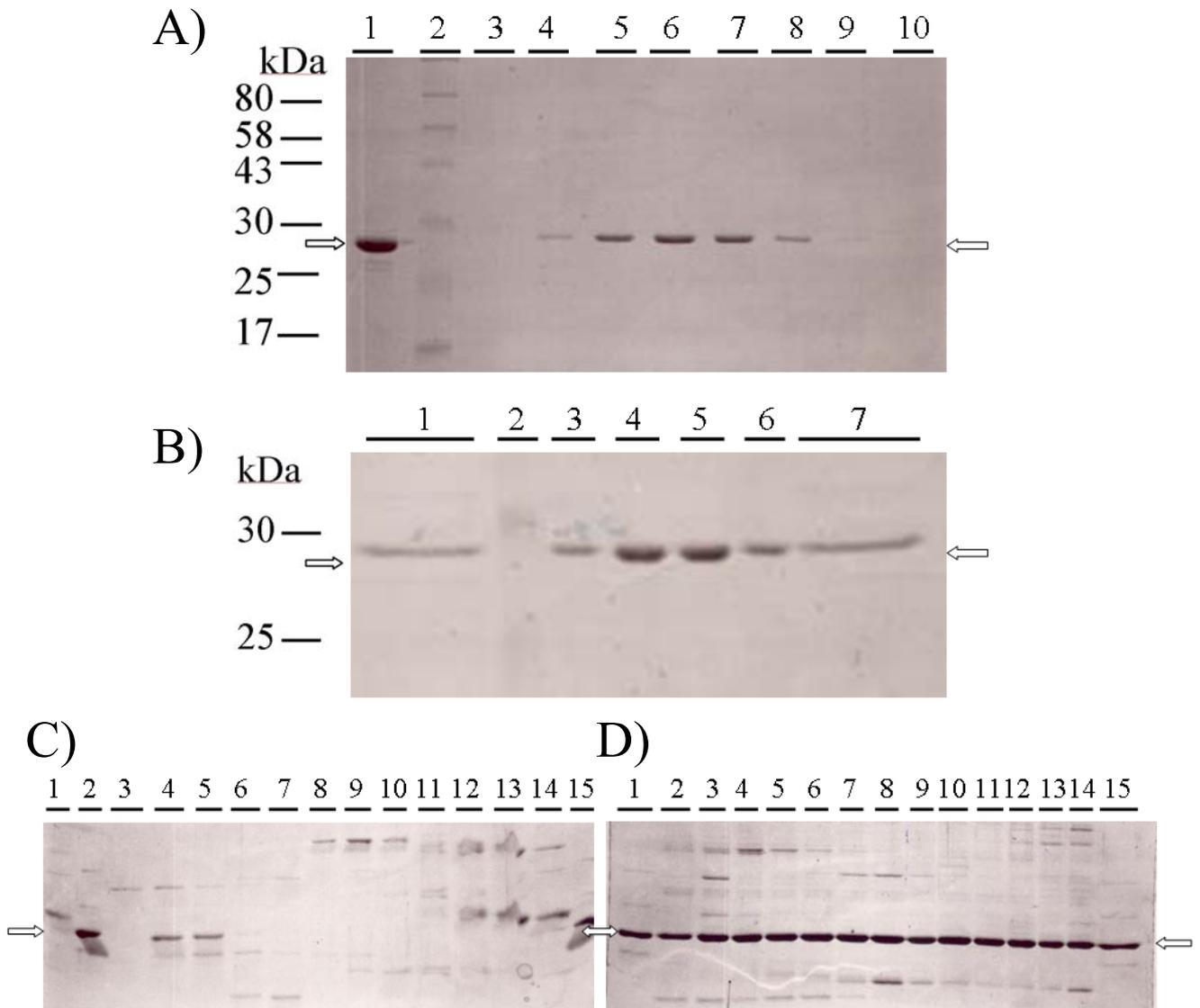
ligation step before transformation. The population of transformants bearing the template plasmid was reduced still further by digestion of the PCR product with DpnI restriction endonuclease.

Originally, all mutagenesis of *siaP* was performed in the plasmid pES8 (Table 2.3), to allow the mutated *siaP* gene to be shuttled into two downstream vectors. Once confirmed by sequencing, the mutated genes in the pES8 vector were inserted into pET21b (between NdeI and XhoI) for over-expression and *in vitro* analysis, or into pES9, along with a strong ribosome binding site, upstream of *siaQM* (between KpnI and XhoI) for *in vivo* complementation analysis. During this project, it was decided to focus on the analysis of purified protein *in vitro*, relieving the need to shuttle the mutated *siaP* genes between vectors, and so mutagenesis was performed directly on the gene in the larger pAH16 construct (6.4 kb in total). For each variant, the DNA sequence of the new gene was determined and the protein expression was tested using standard methods.

For large scale production, native SiaP and its variants were produced at 25 °C in *E. coli* BL21 (DE3) pLysS containing the plasmid bearing the *siaP* gene, or the chosen variant, in M9 minimal medium with 0.4% glucose and appropriate antibiotic selection. Expression was induced at OD<sub>650</sub> 0.2–0.4 with 1 mM IPTG and the cells were left growing overnight.

SiaP contains an N-terminal signal peptide for its export to the periplasm and, when transported, this peptide is cleaved to give processed SiaP. This processed SiaP can then be easily separated from the unprocessed SiaP by preparing the periplasmic fraction. Processed SiaP was then purified from the periplasmic fraction by hydrophobic interaction chromatography (HIC) followed by size exclusion chromatography (SEC) (Figure 3.2a).

SiaP:R147A, SiaP:R147E and SiaP:R147K accumulated to a similar level as the native protein. However, only SiaP:R147E purified in an identical manner to the native SiaP using HIC (Figure 3.2ab). Purification of SiaP:R147A and SiaP:R147K was attempted twice, but these mutants did not separate on the HIC column with the same profile as the native SiaP, rather they eluted at the highest concentrations of ammonium sulphate



**Figure 3.2:** SDS PAGE gels showing the purification of native SiaP, SiaP:R147E and SiaP-R147K. **A)** Fractions of SiaP eluted by SEC. Lane 1: sample from HIC column loaded onto SEC column; lane 2: pre-stained MW marker; lanes 3–9 fractions covering the SEC elution peak. **B)** Fractions of SiaP:R147E eluted by SEC. Lanes 1 and 7: nativeSiaP as size marker; lane 2: pre-stained MW marker; lanes 3–6 fractions covering the SEC elution peak. **C)** Fractions from SiaP:R147A-containing periplasmic preparation eluted by HIC. These fractions cover the expected elution peak. Lanes 1 and 15: SiaP as size marker, indicated by arrows; lanes 2–13: fractions 21–32. **D)** The final fractions from SiaP:R147A-containing periplasmic preparation eluted by HIC. Lanes 1 and 15: SiaP as size marker, indicated by arrows; lanes 2–13 fractions 78–90. These show the elution of SiaP:R147E over a large number of the final fractions.

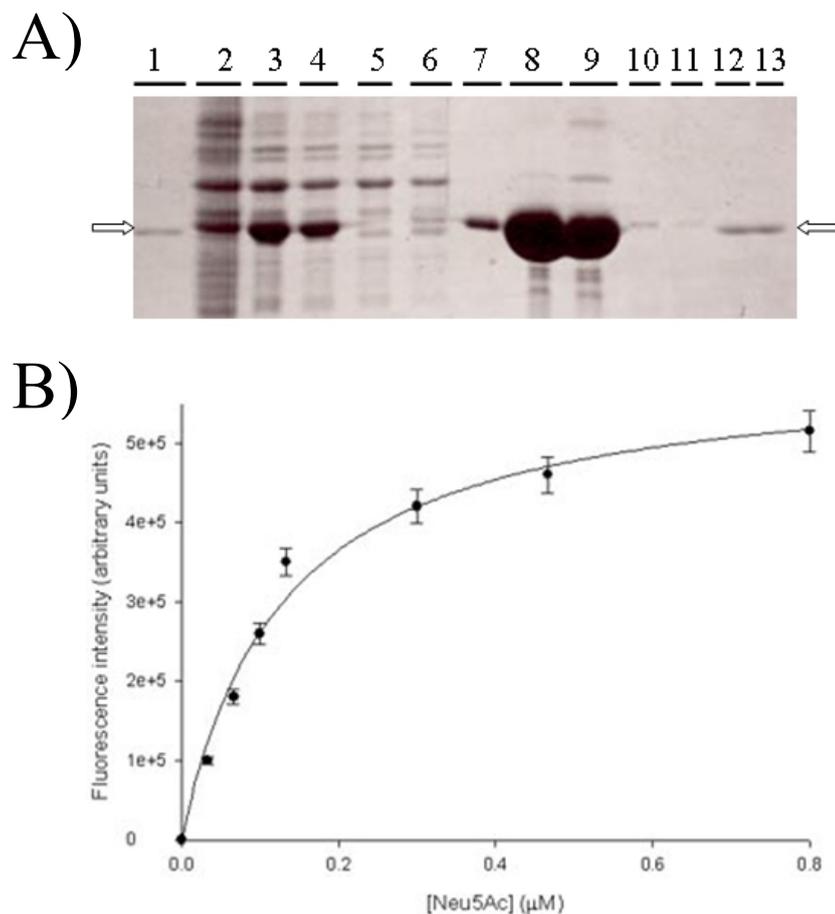
(Figure 3.2cd). This could be due to the amino acid substitutions causing very slight disruption to the conformation of the protein. As can be seen, basing a standard purification procedure for the SiaP mutants using HIC would be unreliable due to its sensitivity to changes in SiaP.

### 3.3 SiaP-His<sub>6</sub> functions as SiaP *in vitro* and *in vivo*

Due to the inability of HIC to reliably separate mutant SiaP proteins, a C-terminal hexahistidine-tag (His<sub>6</sub>-tag) was added to *siaP* and the resultant SiaP-His<sub>6</sub> was then purified and characterised for sialic acid binding *in vitro* and *in vivo*. The C-terminus was chosen as the position for the His<sub>6</sub>-tag since this would not require the removal of the signal peptide, which could have an effect on the folding of the protein and would result in a cytoplasmic SiaP variant.

The *siaP-His<sub>6</sub>* gene was engineered using a 51 base reverse primer and the original forward primer (SiaPfor, Table 2.1) and ligated into the pBlueScriptII vector (making pAH14). This was then cloned into both the pES9-based *in vivo* and pET21b- based *in vitro* vectors (to make pAH15 and pAH16, respectively). Expression and periplasmic fractionation of SiaP-His<sub>6</sub> was performed exactly the same as for native SiaP (Figure 3.3a). His<sub>6</sub>-tagged SiaP was separated from the periplasmic preparation by Ni-chromatography using the standard protocol.

All fractions were visualised by SDS-PAGE for the presence of the tagged protein (Figure 3.3a) and the second elution fraction was sufficiently clean to be analysed. Following dialysis into the required buffer, this was then kept at 4 °C. As with native SiaP, the affinity of SiaP-His<sub>6</sub> for sialic acid was determined by tyrosine fluorescence titration. This protein undergoes a fluorescence signal increase of just 5% on ligand binding (Severi *et al.*, 2005), making titration of this change difficult. Fluorescence titration of SiaP-His<sub>6</sub> was achieved under the same conditions as native SiaP, giving a  $K_d$  value of  $0.14 \pm 0.04 \mu\text{M}$  in 50 mM Tris/HCl pH 8.0 at 37 °C (Figure 3.3b), which is identical to the native protein (Severi *et al.*, 2005).



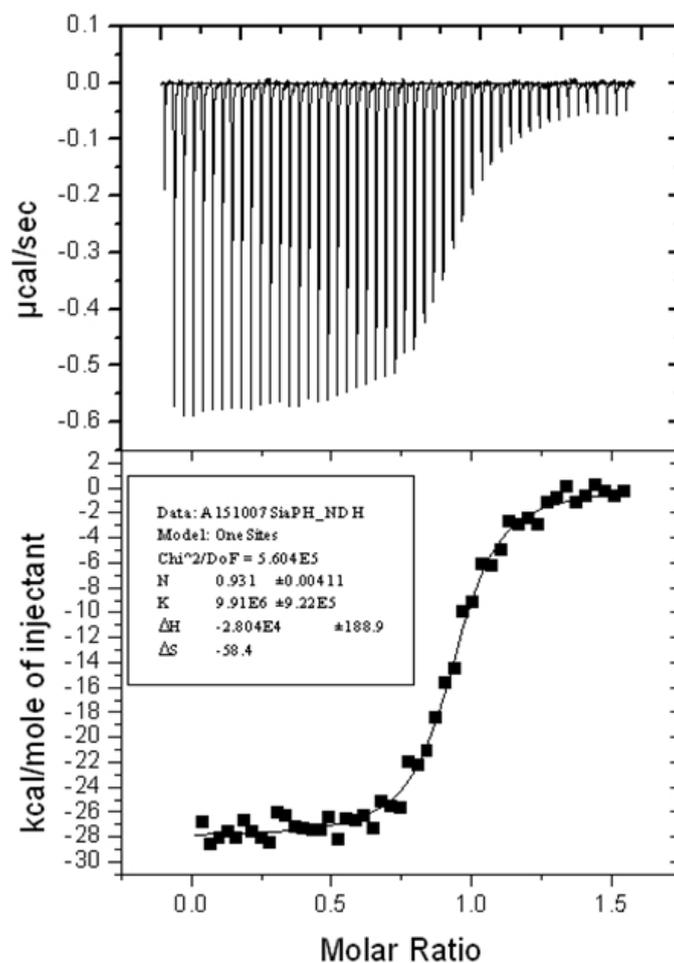
**Figure 3.3:** Purification and *in vitro* ligand binding analysis of SiaP-His<sub>6</sub>. **A)** Fractions from the Ni-affinity purification of SiaP-His<sub>6</sub>. Lanes 1 and 12: 150 ng SiaP as size marker; lane 2: spheroplasts from induced cells; lane 3: periplasmic preparation; lane 4: periplasmic preparation dialysed into Ni-column buffer; lane 5: Ni-column flow through; lane 6: Ni-column washing step; lanes 7–11: Ni-column elution fractions 1–5. These show the elution of SiaP-His<sub>6</sub> over two 1 ml fractions. **B)** Titration of the tyrosine fluorescence signal of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>. This was repeated in triplicate and gave the  $K_d$  value as  $0.17 \pm 0.01 \mu\text{M}$ .

This  $K_d$  value was confirmed in triplicate using Isothermal Titration Calorimetry (ITC). 10  $\mu\text{M}$ , 13.2  $\mu\text{M}$  and 14.6  $\mu\text{M}$  SiaP-His<sub>6</sub> was titrated with 6  $\mu\text{l}$  additions of 150  $\mu\text{M}$ , 200  $\mu\text{M}$  and 100  $\mu\text{M}$  sialic acid in 50 mM Tris/HCl pH 8.0 at 37 °C (Figure 3.4). These gave a  $K_d$  value for sialic acid of  $0.11 \pm 0.02 \mu\text{M}$  and the thermodynamics of binding as  $\Delta H^\circ = -116 \text{ kJ mol}^{-1}$ ;  $\Delta S^\circ = -240 \text{ kJ mol}^{-1} \text{ K}^{-1}$ ;  $\Delta G^\circ = -41 \text{ kJ mol}^{-1}$  and so binding is very favourable and enthalpically driven. The thermodynamics of ligand binding will be discussed in greater detail in Section 4.1.4.

Since the presence of the His<sub>6</sub>-tag had no detectable effect on ligand binding *in vitro*, the effect of the His<sub>6</sub> tag on transport was monitored *in vivo* using a system developed by Dr. Emmanuele Severi. In this, the *siaPQM* genes are expressed from a single operon in the low copy-number plasmid, pWKS30, by IPTG induction in a strain of *E. coli* lacking its native sialic acid transporter, *E. coli* BW25113  $\Delta nanT$ .

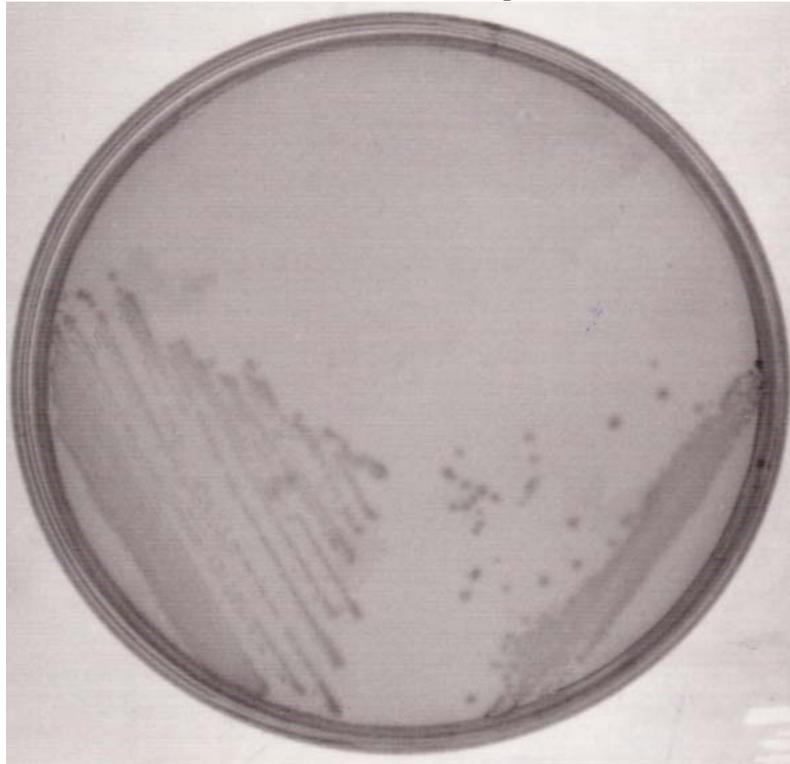
The *siaP* and *siaP-His<sub>6</sub>* genes were introduced upstream of *siaQM* to make pES7 (Severi *et al.*) and pAH15, respectively. Transformants bearing each of these were streaked on M9 minimal medium 1% agarose supplemented with 1 mg/ml Neu5Ac as the sole carbon source. These showed that growth occurred with expression of the whole transporter containing either SiaP or SiaP-His<sub>6</sub> (Figure 3.5).

The above results would indicate that the addition of the hexahistidine tag to SiaP has no detrimental effect on either ligand binding *in vitro* or ligand-dependent interaction with SiaQM *in vivo* and so SiaP-His<sub>6</sub> is an acceptable variant to work with.



**Figure 3.4:** ITC analysis of 10  $\mu\text{M}$  SiaP-His<sub>6</sub> in 50 mM Tris/HCl pH 8.0 at 37 °C. 150  $\mu\text{M}$  Neu5Ac was injected into the cell in 6  $\mu\text{l}$  aliquots. Here, for SiaP-His<sub>6</sub>  $K = 9.91 \times 10^6 \text{ M}^{-1}$  and so the  $K_d$  is 0.10  $\mu\text{M}$ .

*E. coli* BW25113  $\Delta nanT$  pWKS30



*E. coli* BW25113  $\Delta nanT$   
pWKS30-*siaP*-*His*<sub>6</sub>-*siaQM*

*E. coli* BW25113  $\Delta nanT$   
pWKS30-*siaPQM*

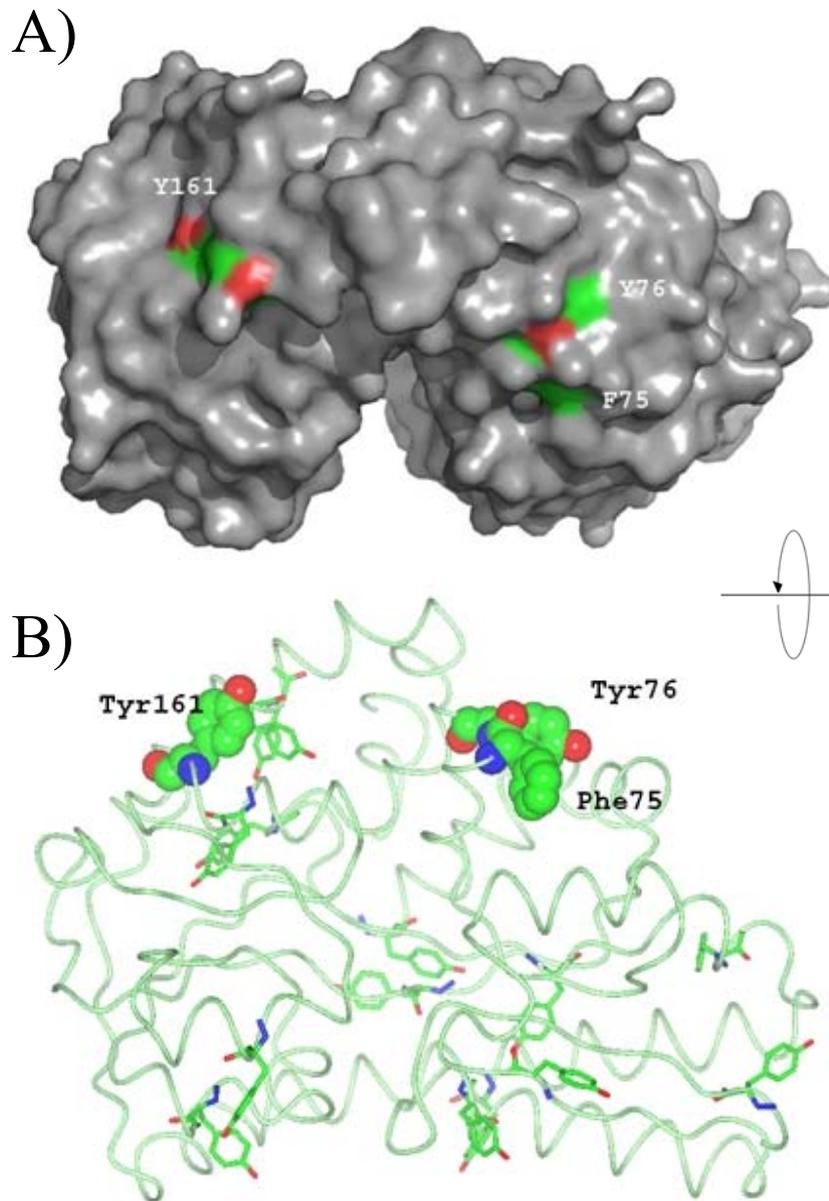
**Figure 3.5:** Growth of *E. coli* BW25113  $\Delta nanT$  strains containing the empty vector, pWKS30, expressing *siaPQM* or *siaP*-*His*<sub>6</sub>-*QM* on M9 minimal medium supplemented with 1 mg/ml Neu5Ac as the sole carbon source.

### **3.4 The introduction of tryptophan residues can improve the fluorescence signal change on ligand binding**

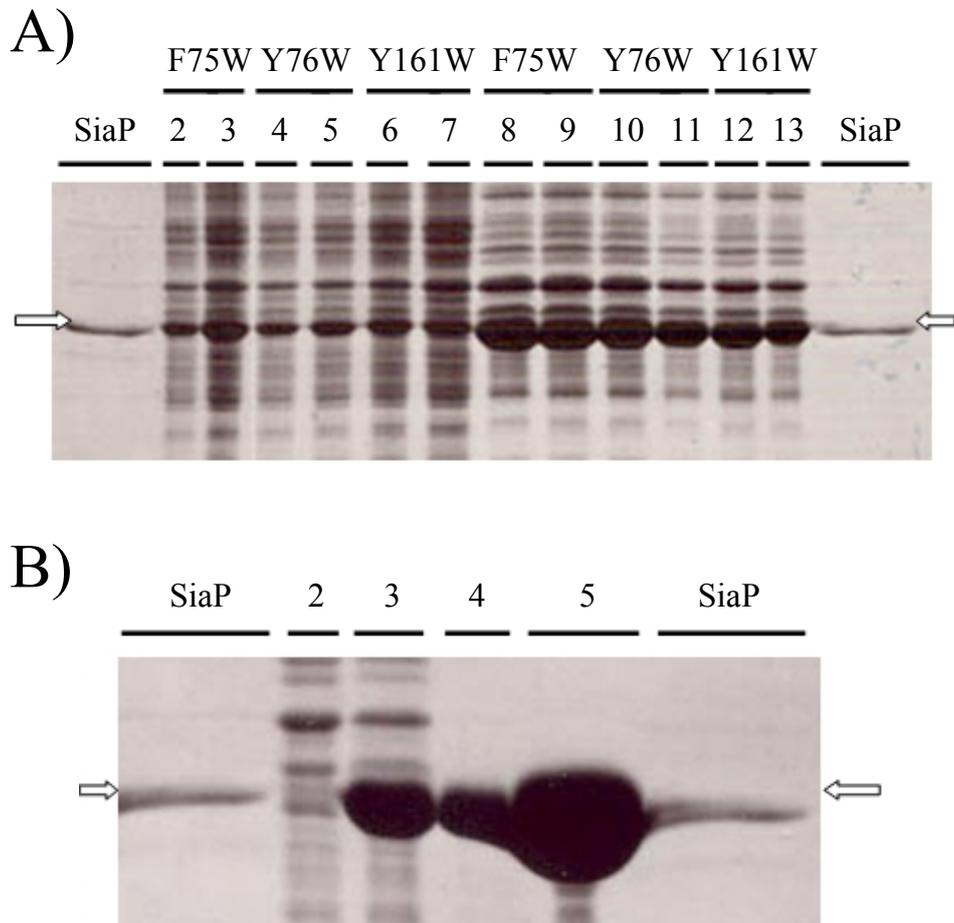
Due to the poor fluorescence signal change of SiaP on ligand binding, we attempted to create a variant of SiaP with a stronger fluorescence signal. Single tryptophan residues were introduced into three different positions in *siaP-His<sub>6</sub>*. The first two positions, Tyr-76 and Tyr-161, are close to the surface of the protein, near to the environment of the cleft and so were suspected of causing the fluorescence change in the protein upon ligand binding (Figure 3.6a). The third position, Phe-75, came from an alignment with VC1779, the homologous sialic acid binding protein from *Vibrio cholerae*, which has a very strong fluorescence signal change on ligand binding. Phe-75 corresponds to one of the four tryptophan residues in VC1779 and is surface exposed near the domain interface (Figure 3.6a).

These three mutant proteins were produced and purified identically to SiaP-His<sub>6</sub> (Figure 3.7). The tryptophan fluorescence spectra of these proteins were compared to SiaP-His<sub>6</sub> and no tryptophan emission could be seen on excitation for tryptophan residues at 297 nm (Figure 3.8). However, when excited for tyrosine fluorescence at 281 nm, emission peaks centred on 340 nm could be seen for the tryptophan insertion mutants. It is possible that these mutations have introduced intramolecular fluorescence resonance energy transfer (FRET) from the tyrosine residues surrounding the introduced tryptophan. FRET occurs when the absorbed energy from one fluorophore is absorbed by another fluorophore without emitting the energy packet as a photon, but resonating between the energy levels of the donating and accepting fluorophores. This is supported by the increase in intensity of the 340 nm peak in these mutants from Tyr-76 to Phe-75 to Tyr-161 corresponds to the increasing number of tyrosine residues in close proximity to the introduced tryptophan residue (Figure 3.6b).

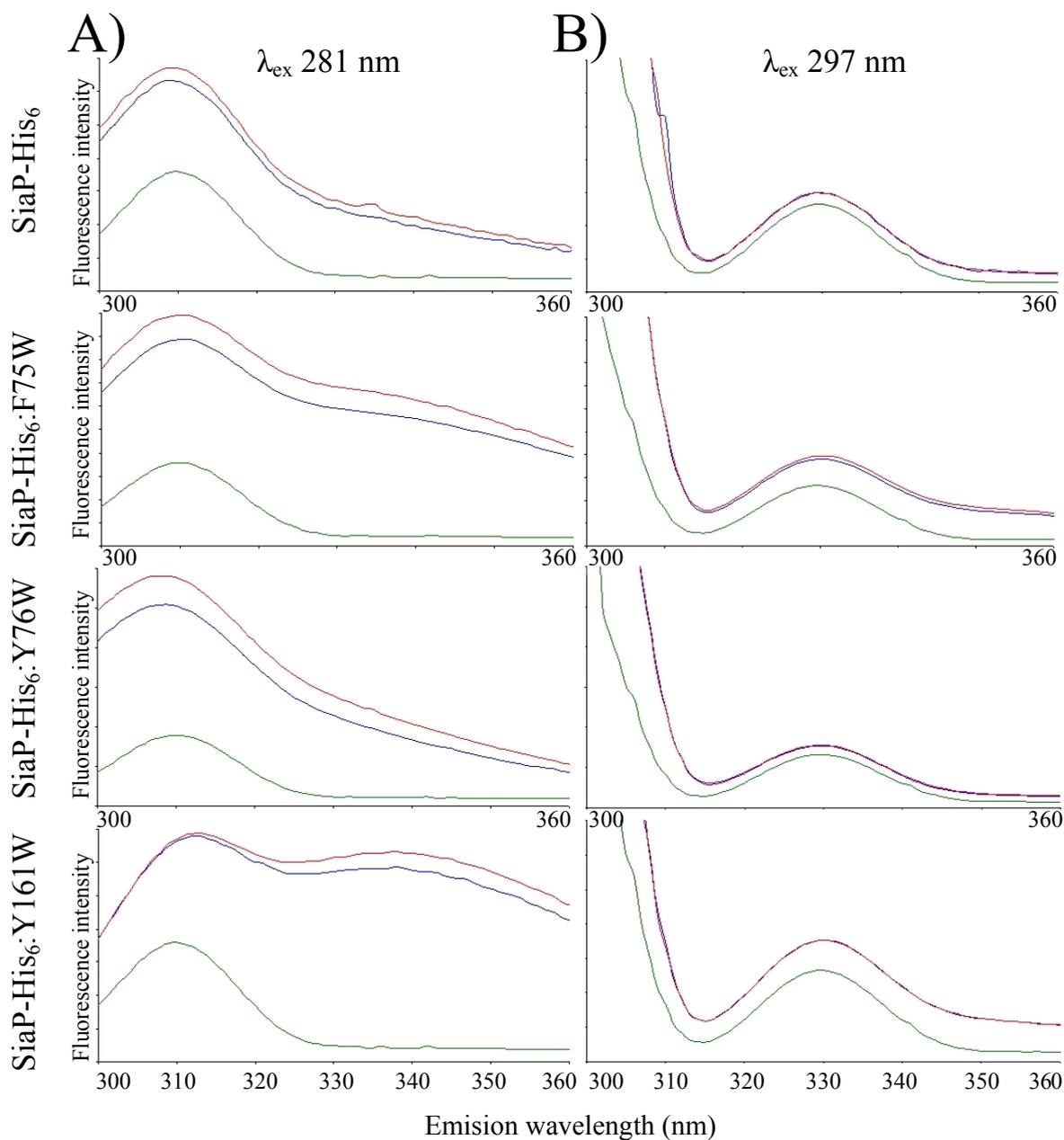
The fluorescence signal change of these mutants on ligand binding is not a large increase, of only 12 %, 13 % and 5 % for Phe-75, Tyr-76 and Tyr-161, respectively. The signal of the SiaP-His<sub>6</sub>:Y161W mutant was titrated and gave a  $K_d$  value of  $0.17 \pm 0.01$   $\mu$ M.



**Figure 3.6:** **A)** Surface representation of SiaP with targets for tryptophan mutants highlighted (PyMol). Tyr-75 and Tyr-161 are suspected of causing the fluorescence change in SiaP; Phe-75 corresponds to the suspect Trp residue from *Vibrio cholerae* VC1779, the sialic acid-binding PBP. **B)** Worm representation of SiaP with targets for tryptophan mutants as spheres and tyrosine residues as cylinders (CCP4MG).



**Figure 3.7:** **A)** Whole cell lysates from production of SiaP-His<sub>6</sub>:F75W, SiaP-His<sub>6</sub>:Y76W and SiaP-His<sub>6</sub>:Y161W. Lanes 1 and 14: SiaP as size marker, indicated by arrows; lanes 2 and 3: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH38 expressing SiaP-His<sub>6</sub>:F75W; lanes 4 and 5: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH39 expressing SiaP-His<sub>6</sub>:Y76W; lanes 6 and 7: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH40 expressing SiaP-His<sub>6</sub>:Y161W; lanes 8 and 9: Periplasmic fractions from induced *E. coli* BL21 (DE3) pLysS pAH38 expressing SiaP-His<sub>6</sub>:F75W; lanes 10 and 11: Periplasmic fractions from induced *E. coli* BL21 (DE3) pLysS pAH39 expressing SiaP-His<sub>6</sub>:Y76W; lanes 12 and 13: Periplasmic fractions from induced *E. coli* BL21 (DE3) pLysS pAH40 expressing SiaP-His<sub>6</sub>:Y161W. **B)** Ni-affinity purification of SiaP-His<sub>6</sub>:F75W. Lanes 1 and 6: SiaP as size marker, indicated by arrows; lane 2: Ni-column flow through; lane 3: Ni-column washing step; lanes 4–5: Ni-column elution fractions 1–4. These show the greatest elution of SiaP-His<sub>6</sub>:F75W in fraction 2. SiaP-His<sub>6</sub>:Y76W and SiaP-His<sub>6</sub>:Y161W elute in a similar manner.



**Figure 3.8:** Fluorescence emission spectra (300 nm – 360 nm) of SiaP-His<sub>6</sub>:F75W, SiaP-His<sub>6</sub>:Y76W and SiaP-His<sub>6</sub>:Y161W with and without a saturating concentration of sialic acid when excited at (A) 281 nm and (B) 297 nm. Protein fluorescence emission spectra are shown in blue, protein plus saturating sialic acid in red and buffer without protein in green. The tryptophan emission ( $\lambda_{\text{ex}}$  297 nm) of all proteins is insensitive to the addition of sialic acid, while the appearance of a peak around 340 nm (when excited at 281 nm) corresponds to tryptophan emission.

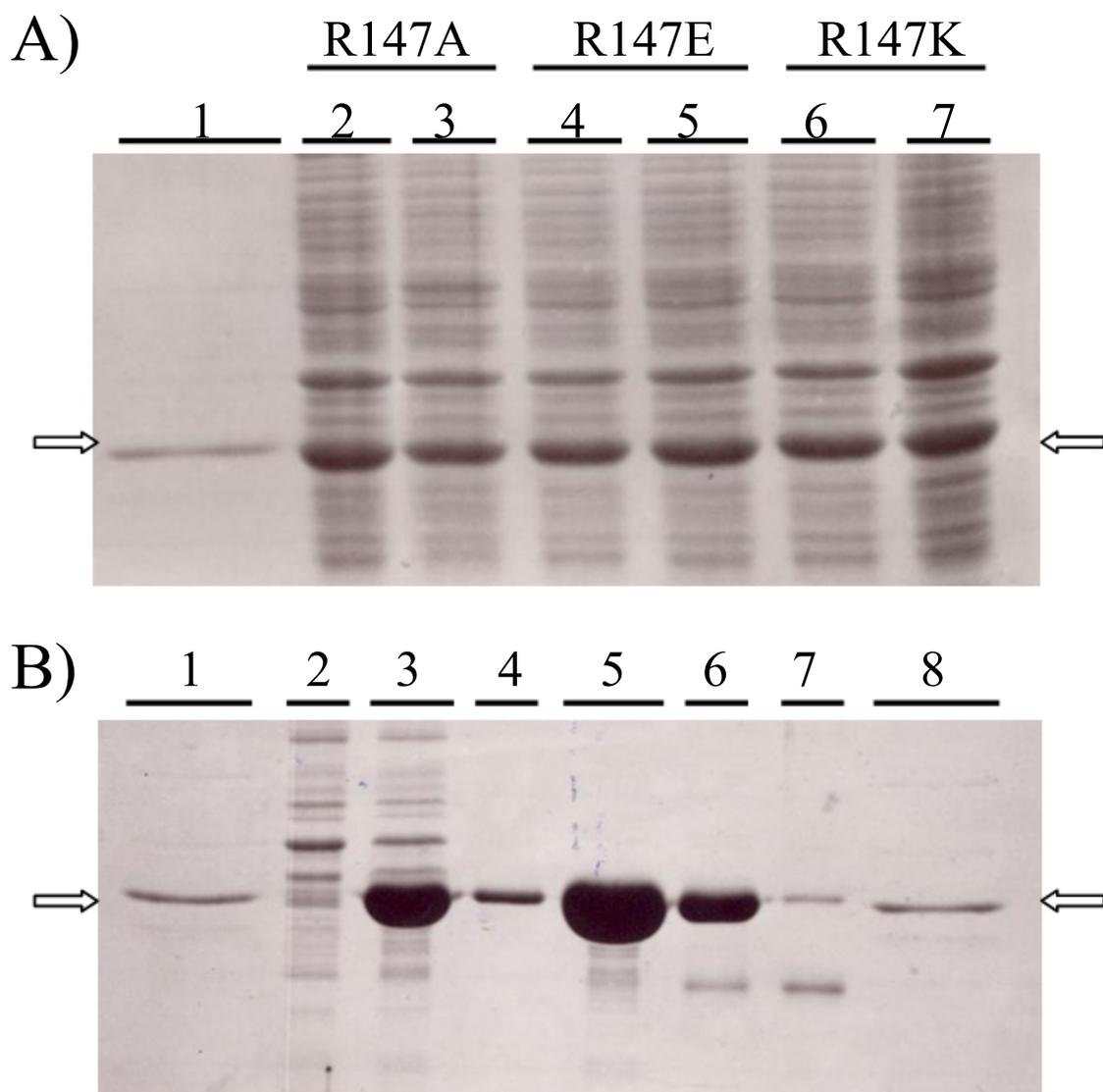
Unfortunately, this signal change was not a significant improvement over SiaP-His<sub>6</sub>, and so these mutations were not used any further for functional work.

### **3.5 Mutations of residues co-ordinating the ligand carboxylate group disrupt ligand binding**

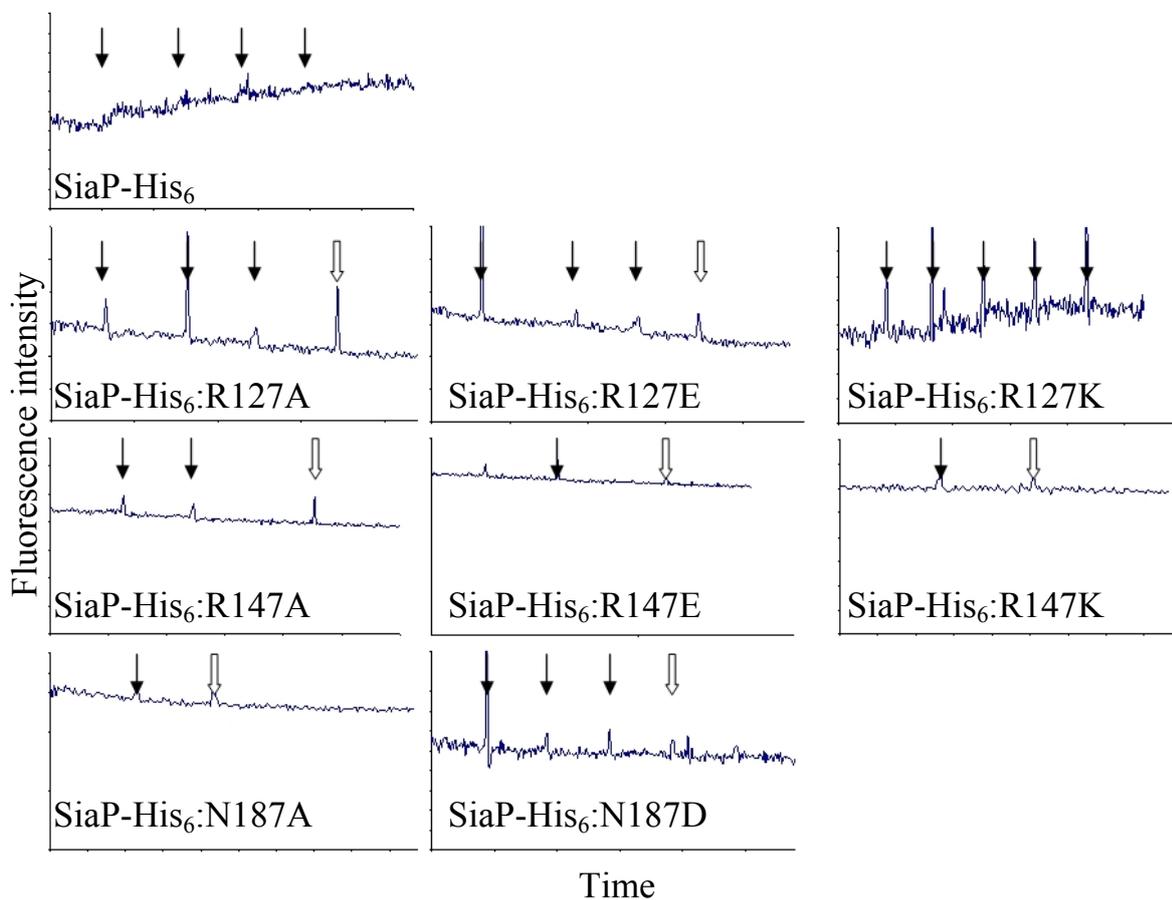
All mutations of the ligand carboxylate-coordinating residues (Table 3.1) were introduced into the *siaP-His<sub>6</sub>* allele in the pAH16 construct for *in vitro* analysis and showed similar expression in *E. coli* BL21 (DE3) pLysS (Figure 3.9a). The proteins were then purified in a similar manner to SiaP-His<sub>6</sub> (Figure 3.9b shows R147A as an example) so that their binding affinity for sialic acid could be determined.

Using the same conditions as for the determination of ligand binding affinity by native SiaP and SiaP-His<sub>6</sub>, each protein was monitored for its ligand-induced fluorescence signal change. For all of the mutants, no signal change could be seen on the addition of up to 3 mM sialic acid. The protein concentration chosen previously (0.05 μM) was a compromise between achieving measurable signal intensity and maintaining pseudo-first order conditions, where the protein concentration was significantly below the  $K_d$  value. Since the  $K_d$  value for these mutants was clearly higher than the native SiaP, the protein concentration in the fluorimeter cell could be increased, thereby increasing the signal intensity. The fluorescence signal of 1 μM protein was monitored on the addition of 3 mM Neu5Ac (90 μl addition of 100 mM Neu5Ac 50 mM Tris/HCl pH 8.0, supplemented with 1 μM protein to avoid dilution effects on fluorescence signal), and only SiaP-His<sub>6</sub>:R127K demonstrated a detectable fluorescence signal change (Figure 3.10; Table 3.2).

The fluorescence signal change of SiaP-His<sub>6</sub>:R127K was titrated and gave a  $K_d$  value of  $0.98 \pm 0.19$  mM. Only this most conservative of mutations showed any ligand interaction at all, indicating that the other mutants have severely disrupted sialic acid binding. For all of these proteins, it is worth re-iterating that the very small fluorescence signal changes on ligand binding made analysis of these proteins problematic.



**Figure 3.9:** SDS PAGE gels showing the production of SiaP-His<sub>6</sub>-R147A, SiaP-His<sub>6</sub>-R147E and SiaP-His<sub>6</sub>-R147K. **A)** Whole cell lysates from *E. coli* BL21 (DE3) pLysS expressing SiaP-His<sub>6</sub>-R147A, SiaP-His<sub>6</sub>-R147E and SiaP-His<sub>6</sub>-R147K. Lane 1: SiaP as size marker, indicated by arrows; lanes 2 and 3: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH35 expressing SiaP-His<sub>6</sub>-R147A; lanes 4 and 5: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH36 expressing SiaP-His<sub>6</sub>-R147E; lanes 6 and 7: Whole cell lysates from induced *E. coli* BL21 (DE3) pLysS pAH37 expressing SiaP-His<sub>6</sub>-R147K. **B)** Ni-affinity purification of SiaP-His<sub>6</sub>-R147A. Lanes 1 and 8: SiaP as size marker, indicated by arrows; lane 2: Ni-column flow through; lane 3: Ni-column wash step; lanes 4–7: Ni-column elution fraction 1–4.



**Figure 3.10:** Fluorescence titration of SiaP-His<sub>6</sub> with mutations as indicated. Excitation wavelength 281 nm, emission wavelength 310 nm. Each black arrow shows an addition of sialic acid where the last addition (block arrow) gives a final concentration of 3 mM sialic acid. A titration of the fluorescence signal is only apparent for SiaP-His<sub>6</sub>:R127K, which gave a  $K_d$  of  $0.98 \pm 0.19$  mM.

**Table 3.2:** Sialic acid-binding affinities of the ligand-binding mutant proteins. Indeterminate  $K_d$  values arise when no titration can be seen on the addition of 3 mM sialic acid to 0.5  $\mu$ M protein.

Protein	$K_d$
SiaP-His <sub>6</sub>	$0.11 \pm 0.02 \mu\text{M}$
SiaP-His <sub>6</sub> :R147K	> 3 mM
SiaP-His <sub>6</sub> :R147A	> 3 mM
SiaP-His <sub>6</sub> :R147E	> 3 mM
SiaP-His <sub>6</sub> :R127K	$0.98 \pm 0.19 \text{ mM}$
SiaP-His <sub>6</sub> :R127E	> 3 mM
SiaP-His <sub>6</sub> :R127A	> 3 mM
SiaP-His <sub>6</sub> :N187A	> 3 mM
SiaP-His <sub>6</sub> :N187D	> 3 mM

The failure to detect any fluorescence change from the remaining mutants could have been due to the protein being unfolded or inactive. The folded state of a protein can be determined by its circular dichroism (CD) spectrum. This uses circularly polarised light to interrogate the secondary structure content of a protein in solution and its spectra between 200 – 250 nm reveals its folded state and  $\alpha/\beta$ /coil content.

The CD spectra of the three Arg-147 mutants were obtained and overlaid with that of SiaP-His<sub>6</sub>. As can be seen in Figure 3.11, the spectra of the three mutants were very similar to SiaP-His<sub>6</sub>, indicating that the secondary structure of these proteins are similar and so would seem to be folded in the same manner.

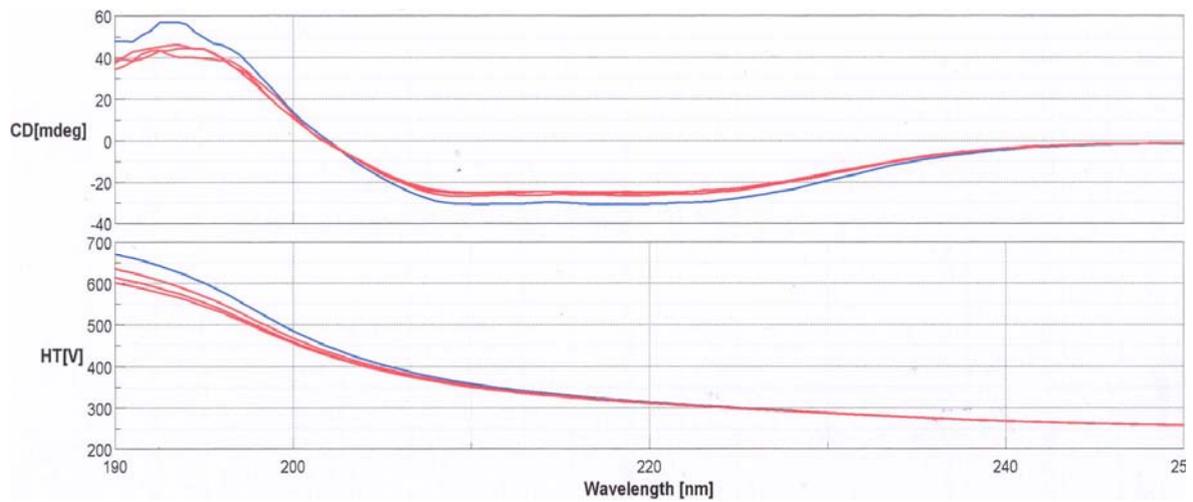
This similarity of their spectra suggests that the mutant proteins have folded correctly. However, the CD spectra cannot reveal very slight changes in structure and so these mutants might have very small malformations that render them non-functional.

### **3.6 An aromatic residue is required at position 170 for high-affinity ligand binding**

As mentioned earlier, Phenylalanine-170 was suspected of acting as a lid for the binding site. As in Table 3.1, the role of this residue was investigated by mutation to alanine, tryptophan and tyrosine.

Using the same mutagenic approach outlined above, these three mutations were introduced into the *siaP-His<sub>6</sub>* allele in the pAH16 construct, expressed and purified following the same protocol as for SiaP-His<sub>6</sub> and their affinities for sialic acid were determined by fluorescence signal titration.

The fluorescence signal changes of SiaP-His<sub>6</sub>:F170Y (an increase) and SiaP-His<sub>6</sub>:F170W (a quench) were larger and clearer than that of SiaP-His<sub>6</sub>, particularly the tryptophan substitution mutant. SiaP-His<sub>6</sub>:F170W had a  $K_d$  value for sialic acid of  $1.21 \pm 0.03 \mu\text{M}$  (a



**Figure 3.11:** Overlaid CD spectra of 0.1 mg/ml SiaP-His<sub>6</sub> (blue) and the SiaP-His<sub>6</sub>;R147 mutants (red) in 20 mM sodium phosphate buffer pH 8.0 at 37 °C. This also shows the potential applied to detector, which remains below 700 V.

10-fold decrease), while that of SiaP-His<sub>6</sub>:F170Y was similar at  $1.13 \pm 0.05 \mu\text{M}$  (Figure 3.12). There was little change in the signal from SiaP-His<sub>6</sub>:F170A, but its significantly decreased affinity of  $340 \pm 70 \mu\text{M}$  suggests that an aromatic residue is important in this position (Figure 3.12).

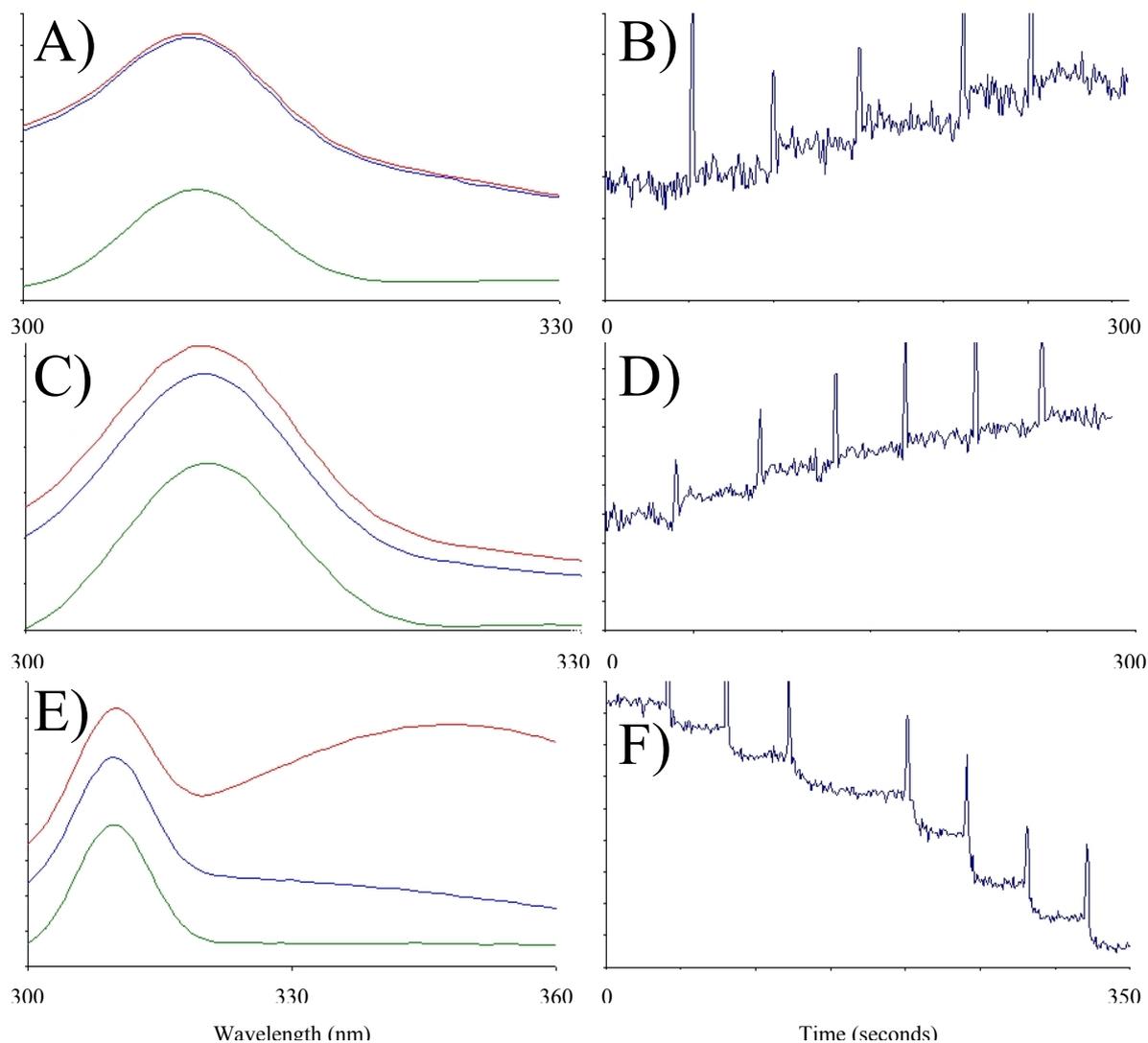
### **3.7 Co-ordination of the ligand carboxylate group is important for high-affinity ligand binding**

In the previous section, SiaP-His<sub>6</sub>:F170W was found to show a strong fluorescence quench on ligand binding. This was examined further as a possible ligand binding-reporter variant of SiaP.

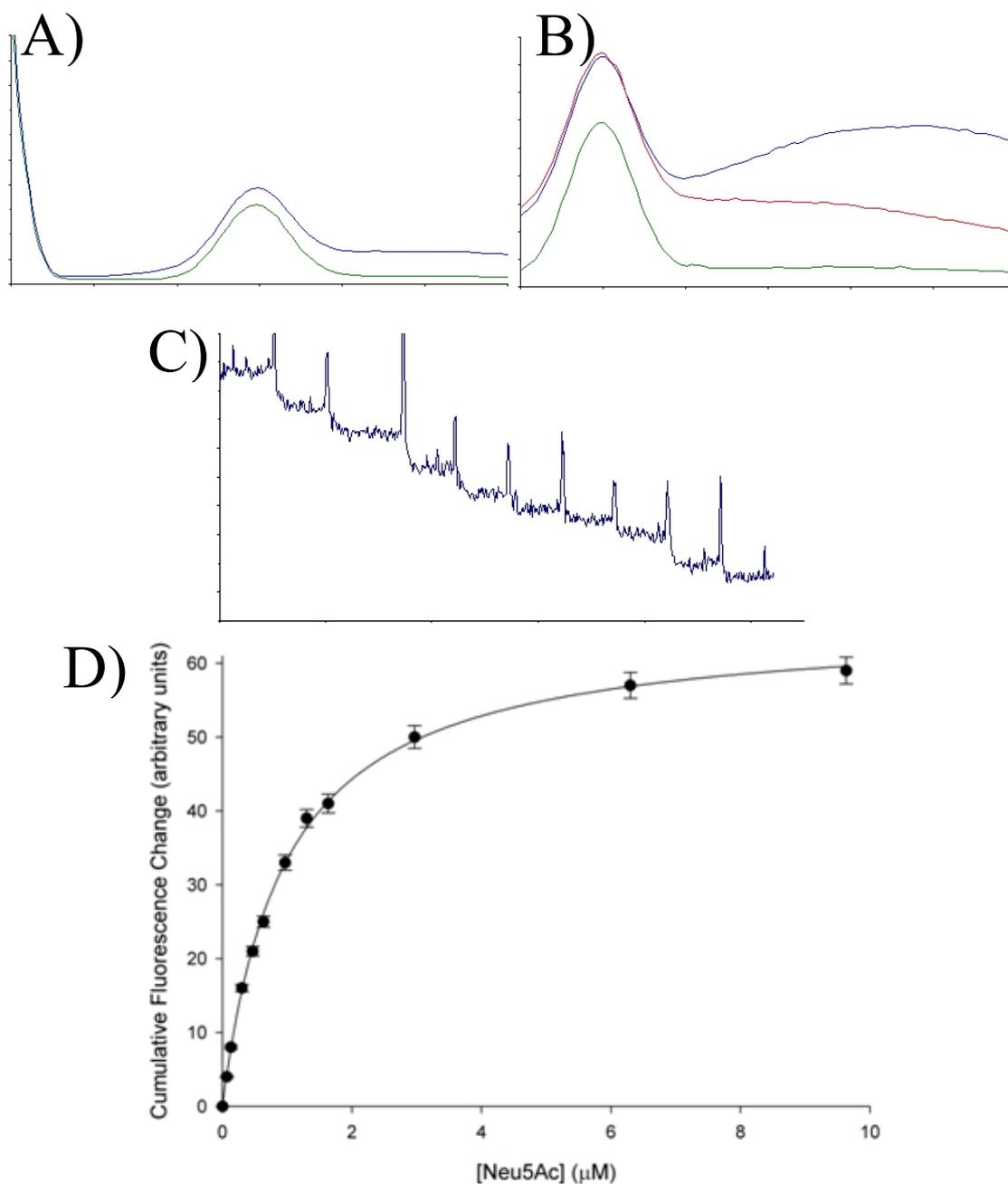
As with previous attempts at tryptophan introduction, this protein showed no tryptophan fluorescence on excitation at 297 nm (Figure 3.13a). However, as expected, when excited at 281 nm for tyrosine residues, it showed a strong emission peak at a longer wavelength, this time centred around 340-350 nm (Figure 3.13b). This shows a 50 % signal quench on ligand binding and was easily titrated to give a  $K_d$  value for sialic acid of  $0.96 \pm 0.03 \mu\text{M}$  (Figure 3.13cd). This was also measured in triplicate by ITC, giving a  $K_d$  value for sialic acid of  $1.21 \pm 0.03 \mu\text{M}$ .

To take advantage of this reporter variant of SiaP, the Arg-127, Arg-147 and Asn-187 mutations (Table 3.1) were introduced in to the SiaP-His<sub>6</sub>:F170W background and purified identically to SiaP-His<sub>6</sub>:F170W.

The mutants in the reporter variant were analysed by fluorescence signal titration identically to those in the SiaP-His<sub>6</sub>:F170W protein. Up to 5 mM sialic acid was added to 1  $\mu\text{M}$  protein (150  $\mu\text{l}$  addition of 100 mM Neu5Ac 50 mM Tris/HCl pH 8.0, supplemented with 1  $\mu\text{M}$  protein to avoid dilution effects on fluorescence signal) in an attempt to detect any fluorescence signal change. As expected, SiaP-His<sub>6</sub>:F170W;R127K showed a fluorescence signal change on the addition of ligand, which was titrated and gave a  $K_d$  value of  $1.5 \pm 0.3 \text{ mM}$ , the highest measured  $K_d$  value of any mutant in this



**Figure 3.12:** Fluorescence titration of 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170A, 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170Y and 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W. **A)** Fluorescence emission spectra of 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170A ( $\lambda_{\text{ex}}$  281 nm). **B)** Example titration of 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170A ( $\lambda_{\text{ex}}$  281 nm,  $\lambda_{\text{em}}$  340 nm). **C)** Fluorescence emission spectra of 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170Y ( $\lambda_{\text{ex}}$  281 nm). **D)** Example titration of 0.25  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170Y ( $\lambda_{\text{ex}}$  281 nm,  $\lambda_{\text{em}}$  340 nm). **E)** Fluorescence emission spectra of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W ( $\lambda_{\text{ex}}$  281 nm). **F)** Example titration of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W ( $\lambda_{\text{ex}}$  281 nm,  $\lambda_{\text{em}}$  340 nm). In emission spectra, protein spectra are shown in blue, protein plus saturating sialic acid in red and buffer without protein in green.



**Figure 3.13:** Fluorescence spectra and titration of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W. **A)** Tryptophan fluorescence emission spectra ( $\lambda_{\text{ex}}$  297 nm) of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W in blue and buffer without protein in green. **B)** Fluorescence emission spectra ( $\lambda_{\text{ex}}$  281 nm) of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W in blue, protein plus saturating sialic acid in red and buffer without protein in green. **C)** Example titration of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W. **D)** Fluorescence titration of 0.05  $\mu\text{M}$  SiaP-His<sub>6</sub>:F170W giving a  $K_d$  value of  $1.21 \pm 0.03 \mu\text{M}$ .

project (Table 3.3). The only other mutant protein in the reporter background to give a signal change was SiaP-His<sub>6</sub>:F170W;N187A, which gave a  $K_d$  value of  $244 \pm 8 \mu\text{M}$ .

The titration of SiaP-His<sub>6</sub>:F170W;N187A could have been due to a beneficial interaction between these two close positions ( $< 4 \text{ \AA}$ ), or the reporter making the signal change more obvious, underlining the need for investigation of negative results by fluorescence methods.

### **3.8 No ligand binding can be detected for the Arg-147 mutants *in vitro***

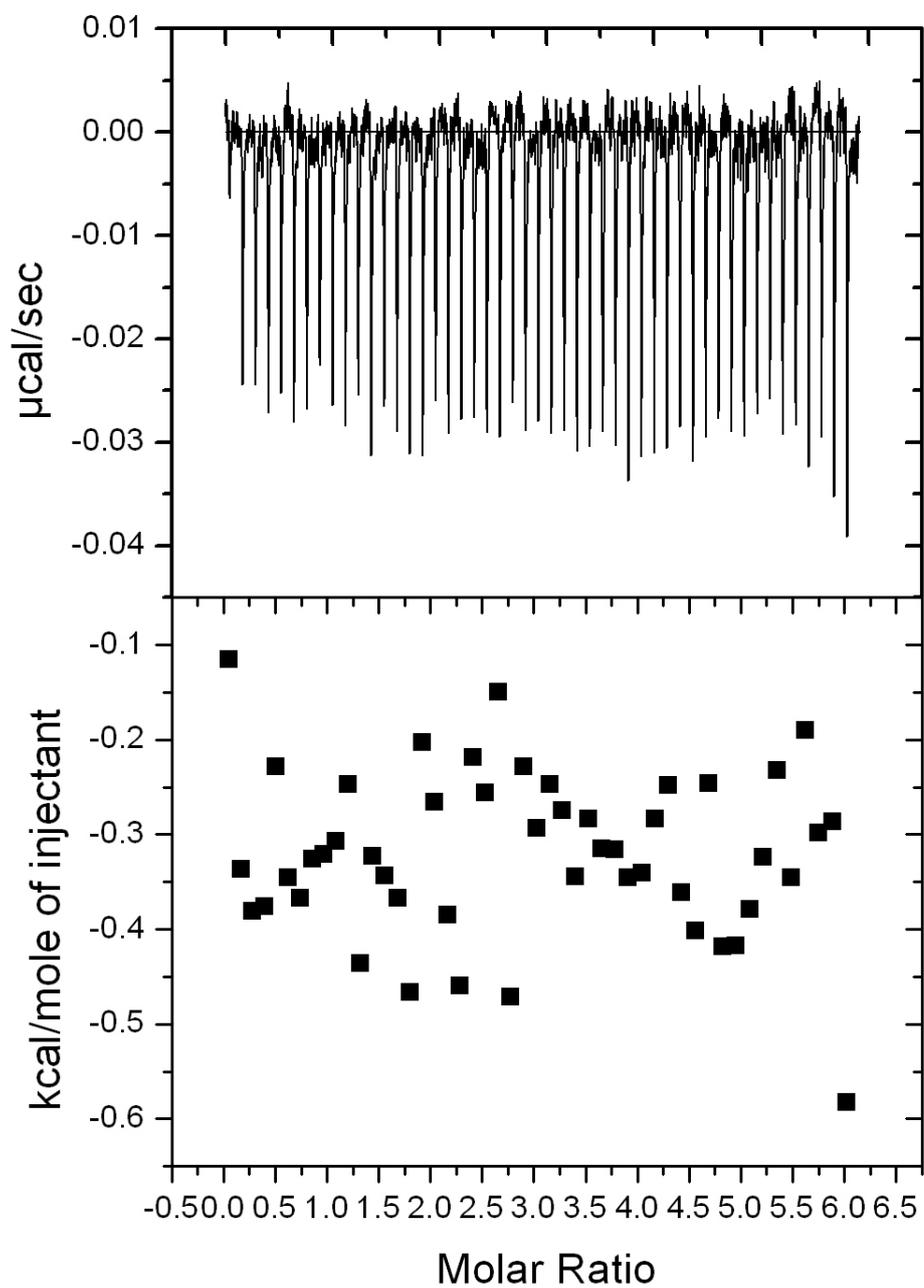
To confirm that an interaction between these mutants and sialic acid was not occurring with a similar affinity to the native SiaP but was undetectable by measuring fluorescence change, ITC was performed twice on the most conservative Arg-147 mutant, SiaP-His<sub>6</sub>:R147K.

In both of these experiments (6  $\mu\text{M}$  protein with 200  $\mu\text{M}$  Neu5Ac; 13  $\mu\text{M}$  protein with 600  $\mu\text{M}$  Neu5Ac), no apparent interaction between the two partners could be seen under these conditions (Figure 3.13).

Another non-fluorimetric method of determining if binding occurs is to use filter binding of precipitated protein following incubation with <sup>14</sup>C-labeled sialic acid. Measuring the intensity of radiation recovered can be compared to the amount of protein originally used and a percentage of this bound with labeled sialic acid can be determined. When incubated with 5  $\mu\text{M}$  <sup>14</sup>C-Neu5Ac,  $35 \pm 10\%$  of SiaP-His<sub>6</sub> and  $40 \pm 5\%$  of SiaP-His<sub>6</sub>:F170W was recovered with labeled sialic acid. For SiaP-His<sub>6</sub>, this could be reduced to  $0 \pm 5\%$  on the addition of 1 mM non-labeled sialic acid. For the SiaP-His<sub>6</sub>:Arg-147 mutants, no interaction (less than 1% recovery) could be detected using 5  $\mu\text{M}$  labeled sialic acid. Using these non-fluorimetric assays for binding, the results indicate that the affinities of these mutants of SiaP are significantly lower than that of the native SiaP.

**Table 3.3:** Sialic acid-binding affinities of the ligand-binding mutants in the reporter variant. Indeterminate  $K_d$  values arise when no titration can be seen using the highest concentration of protein and sialic acid possible.

Protein	$K_d$
SiaP-His <sub>6</sub>	0.11 ± 0.02 μM
SiaP-His <sub>6</sub> :F170W	1.21 ± 0.03 μM
SiaP-His <sub>6</sub> :F170W;R127A	> 5 mM
SiaP-His <sub>6</sub> :F170W;R127E	> 5 mM
SiaP-His <sub>6</sub> :F170W;R127K	1.5 ± 0.3 mM
SiaP-His <sub>6</sub> :F170W;R147A	> 5 mM
SiaP-His <sub>6</sub> :F170W;R147E	> 5 mM
SiaP-His <sub>6</sub> :F170W;R147K	> 5 mM
SiaP-His <sub>6</sub> :F170W;N187A	244 ± 8 μM
SiaP-His <sub>6</sub> :F170W;N187D	> 5 mM



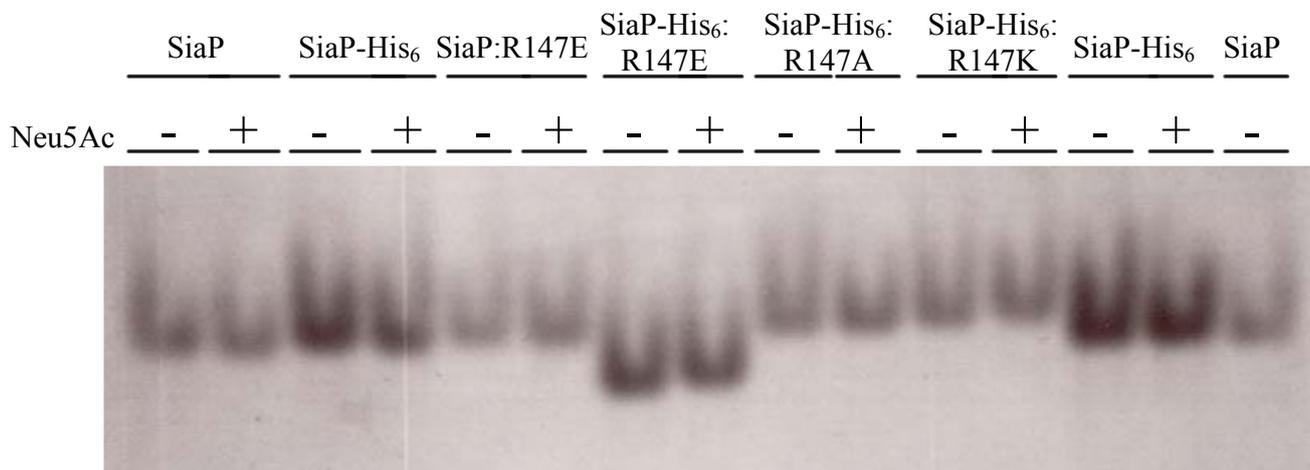
**Figure 3.14:** ITC analysis of 13  $\mu\text{M}$  SiaP-His<sub>6</sub>:R127K in 50 mM Tris/HCl pH 8.0 at 37 °C. 600  $\mu\text{M}$  Neu5Ac was injected into the cell in 6  $\mu\text{l}$  aliquots.

In the investigation of the molybdate binding protein ModA, a ligand-dependent mobility shift assay in native polyacrylamide gel was developed (Rech *et al.*, 1996). In an attempt to show an interaction with sialic acid, this was performed with some of the mutants of SiaP. Native, His<sub>6</sub>-tagged and mutant proteins were separated on 6% polyacrylamide native gel with and without pre-bound sialic acid. Despite several attempts, no clear ligand-dependent differences could be seen between any sample pairs (Figure 3.15). However, interesting size differences between the mutants were seen on these gels. The hexahistidine tag did not appear to cause a change in migration between SiaP and SiaP-His<sub>6</sub>, nor did the introduction of the R147E mutation, but the combination of the R147E mutation with the His<sub>6</sub>-tag caused an increase in mobility. Also, SiaP-His<sub>6</sub>:R147A and SiaP-His<sub>6</sub>:R147K mutants exhibit a small retardation compared to SiaP and SiaP-His<sub>6</sub>. This could be due to a slight decrease in size or overall charge. Unfortunately, this method does not appear to work reliably for SiaP, since the native protein showed no ligand-dependent mobility shift.

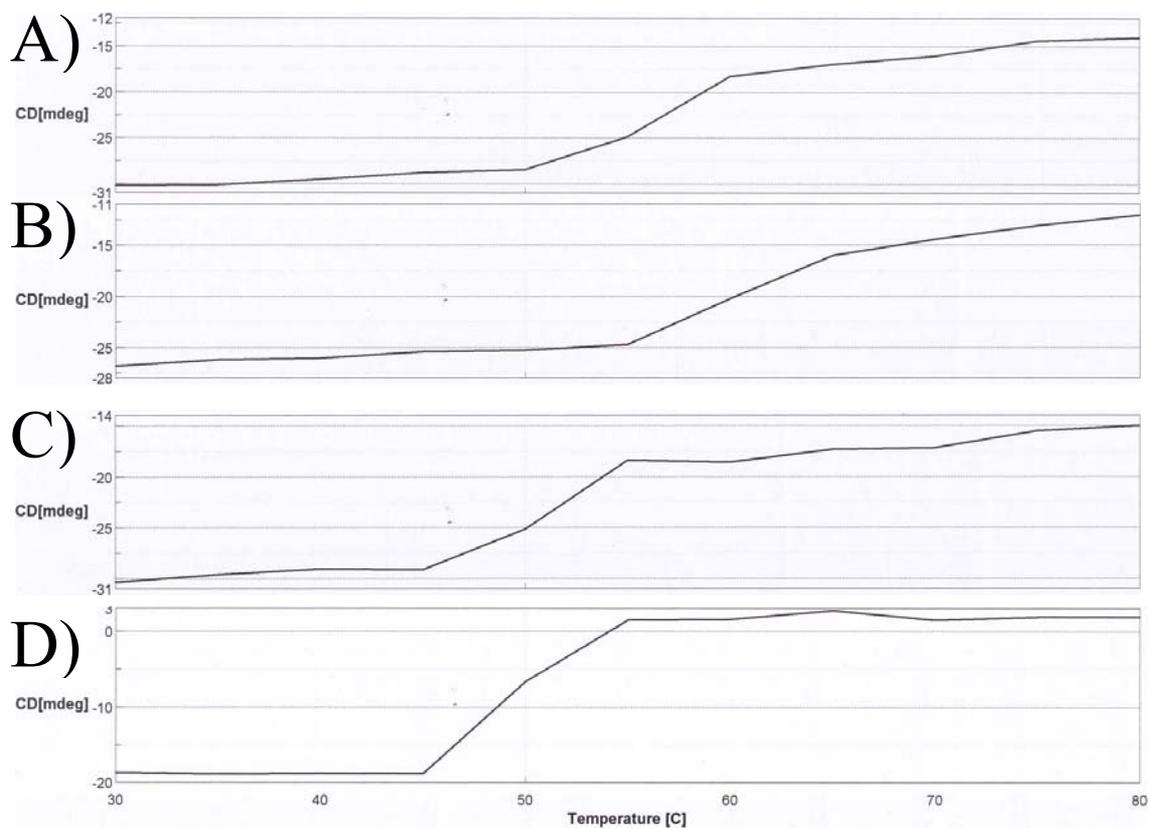
To determine if any detectable domain closure occurred, the difference in melting temperatures between the ligand-bound and unbound proteins could be compared. This change in melting temperature is a common behaviour of proteins which undergo a large domain closure on ligand binding, such as MBP (Gould *et al.*, 2009).

The thermal denaturation of SiaP-His<sub>6</sub> with and without sialic acid was monitored by CD at 222 nm as the temperature was increased (Figure 3.16ab). In a similar behaviour to MBP, the  $T_m$  of SiaP-His<sub>6</sub> increases from 50–60 °C up to 55–65 °C when saturated with sialic acid. The  $T_m$  of un-liganded SiaP-His<sub>6</sub>:R147K is 45–55 °C, indicating that this mutation has a destabilising effect on the protein; on the addition of 5 mM sialic acid, the  $T_m$  remains at 45–50 °C (Figure 3.16cd).

This suggests that no significant domain closure occurs with this concentration of sialic acid, which may be due to the  $K_d$  being above 5 mM or that this mutation has disturbed the mechanism of domain closure.



**Figure 3.15:** Native PAGE gel of 280 ng of each protein as indicated, with and without pre-incubation with sialic acid. No clear differences in mobility can be seen on the addition of sialic acid. However, the R147E point mutations appear to cause changes in the migration in the SiaP-His<sub>6</sub> background.



**Figure 3.16:** CD signal at 222 nm following the thermal denaturation of 0.2 mg/ml protein in 20 mM sodium phosphate buffer pH 8.0. **A)** SiaP-His<sub>6</sub> (T<sub>m</sub> 50-60 °C). **B)** SiaP-His<sub>6</sub> with 10 μM sialic acid (T<sub>m</sub> 55-65 °C). **C)** SiaP-His<sub>6</sub>:R147K (T<sub>m</sub> 45-55 °C). **D)** SiaP-His<sub>6</sub>:R147K with 5 mM sialic acid (T<sub>m</sub> 45-55 °C).

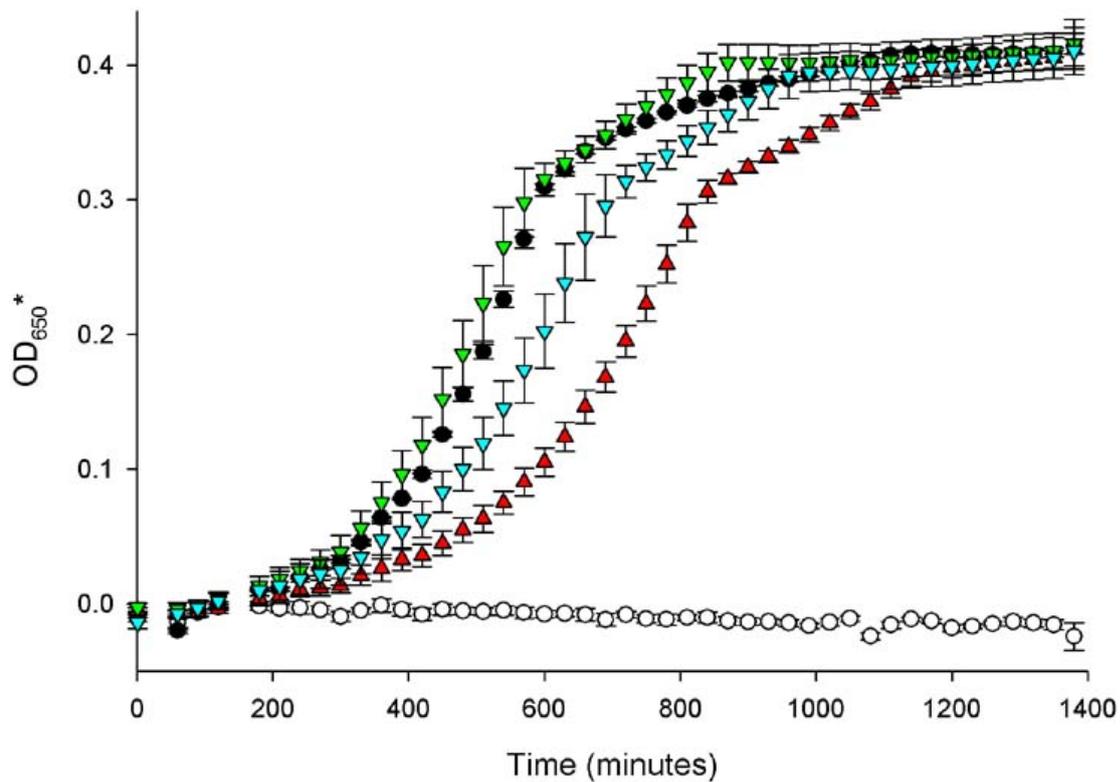
To investigate this hypothesis, the protein crystal structures of the three R147 mutants were determined in the presence of sialic acid by Marcus Fischer. These structures reveal that all of these proteins adopt the closed conformation with sialic acid inside the binding site, having overall structures identical to the native SiaP (Marcus Fischer, personal communication). This stands against the idea that the conserved arginine-carboxylate salt bridges of the DctP-like family are critical for ligand binding, but merely necessary for high affinity interaction.

### **3.9 The Arg-147 residue is not essential for sialic acid transport *in vivo***

The CD spectra of the mutant proteins suggested that they were correctly folded in solution and the protein crystal structures that they could adopt a correctly closed conformation. However, these mutants showed no interaction with sialic acid *in vitro* upto a concentration of 5 mM. The Arg-147 mutations were introduced into the *in vivo* system to see if they abolished function.

The *siaP:R147* mutant alleles were introduced upstream of *siaQM* (pES15, pES16 and pES17) and transformed into *E. coli* BW25113  $\Delta nanT$  to examine growth *in vivo*. Four technical replicates of each of these strains were then grown in 700  $\mu$ l Neu5Ac-supplemented M9 minimal medium in 24-well plates in a prototype incubated plate shaker (EnzyScreen). Cell density was measured twice every hour and converted to an apparent OD<sub>650</sub> (OD<sub>650</sub>\*) value using the standard curve (Figure 2.1). These values were averaged and their standard deviation determined.

*In vitro*, these mutants show no interaction with sialic acid up to 5 mM ligand. However, it can be seen in Figure 3.17 that strains bearing the SiaP:R147 mutations can maintain growth on 3.2 mM sialic acid, but show mutation-dependent phenotypes. This behaviour was also seen with 1.6 mM sialic acid. When attempted using 0.65 mM sialic acid growth could be seen to occur, but was not measurable using the incubated plate shaker. From the increases in the doubling time ( $T_{GEN}$ ), the growth defect caused by the R147A mutation is more severe than that of the R147E, while the conservative SiaP-R147K



**Figure 3.17:** Growth on 3.2 mM sialic acid of *E. coli* BW25113  $\Delta nanT$  strains expressing the *siaPQM* mutant alleles in M9 minimal medium supplemented with sialic acid. The mutations in *siaP* are denoted by: R147A as red up triangles, R147E as blue down triangles and R147K as green down triangles; native *siaPQM* is shown as black circles. Growth was measured by the opacity of the medium and converted to  $OD_{650}^*$  using a standard curve. Error bars are standard deviations of four repeats.

**Table 3.4:** Growth rates of *E. coli* BW25113  $\Delta nanT$  strains expressing *siaPQM* with mutations in *siaP* as indicated.

Transporter	3.2 mM sialic acid		1.6 mM sialic acid	
	Growth rate ( $\text{min}^{-1}$ )	$T_{\text{GEN}}$ (min)	Growth rate ( $\text{min}^{-1}$ )	$T_{\text{GEN}}$ (min)
SiaP–SiaQM	$8.59 \pm 0.19 \times 10^{-3}$	$116 \pm 3$	$6.70 \pm 0.23 \times 10^{-3}$	$149 \pm 5$
SiaP:R147K–SiaQM	$8.19 \pm 0.44 \times 10^{-3}$	$122 \pm 6$	$6.36 \pm 0.42 \times 10^{-3}$	$157 \pm 9$
SiaP:R147E–SiaQM	$6.35 \pm 0.36 \times 10^{-3}$	$157 \pm 8$	$5.04 \pm 0.25 \times 10^{-3}$	$198 \pm 9$
SiaP:R147A–SiaQM	$5.54 \pm 0.15 \times 10^{-3}$	$181 \pm 5$	$4.36 \pm 0.65 \times 10^{-3}$	$229 \pm 30$

mutant is very similar to that of the native SiaP (Table 3.4). Halving the sialic acid concentration causes an increase in  $T_{\text{GEN}}$  of about 27% in all cases. The increases in  $T_{\text{GEN}}$  caused by the mutations in *siaP* are consistent between the two carbon source concentrations; these cause a 5%, 34% and 55% decrease in growth rate for the lysine, glutamate and alanine mutations, respectively.

The growth demonstrates that the Arg-147 residue is not essential for transport and could be being supported by two methods. The first of these is based upon the properties of an equilibrium mixture. With a ligand concentration below the  $K_d$  value of the binding protein, a small population of transient SBP-ligand complexes would be found and as the ligand concentration: $K_d$  ratio decreased, this population would decrease. It could be that this small population is sufficient to maintain growth.

In the second method, unliganded binding protein, probably in a transient closed conformation, could interact with SiaQM and stabilise a transport-competent intermediate complex. If true, this method would suggest that these mutants were fulfilling a more critical role as a scaffold or catalyst for transport.

### 3.8 Summary

This chapter has focussed on the interactions between SiaP and this ligand, particularly the ligand carboxylate group. Once the more easily purified hexahistidine-tagged protein was produced, the investigation of the Arg-127, Arg-147 and Asn-187 residues proceeded more effectively. It was found that any mutations made to these three residues abolished high affinity binding *in vitro* (Table 3.2). These mutations were investigated further using the F170W reporter variant. No ligand binding could be detected for any mutation of Arg-147 and less conservative changes to Arg-127, while the titration of SiaP-His<sub>6</sub>:R127K was much clearer (Table 3.3). In the native background, the N187A mutant gave no detectable ligand binding, while in the reporter background ( $K_d$   $1.21 \pm 0.03 \mu\text{M}$ ), this gave a  $K_d$  value of  $244 \pm 8 \mu\text{M}$ .

The Arg-147 mutations were introduced to the *in vivo* growth system, based on *E. coli* BW25113  $\Delta nanT$  carrying *H. influenzae* *siaPQM* on a low copy-number plasmid. When grown on sialic acid as the sole carbon, nitrogen and energy source, it was found that these Arg-147 mutants allowed growth (Figure 3.17). It is likely that this growth is based on the interaction of the SiaP mutants with sialic acid, since a mutation-dependent phenotype is observed.

## Chapter Four

Investigation and manipulation of the  
ligand binding properties of SiaP for  
application as a biosensor

#### **4 Investigation and manipulation of the SiaP ligand binding properties for application as a sialic acid biosensor**

As a company involved in the detection of counterfeit, adulterated and smuggled products, Authentix are interested in SiaP for its sialic acid detection possibilities. Their objective is to be able to label clients' products with the food safety-certified sialic acid, which can then be detected further down the supply chain in order to confirm the authenticity of the product. Their aim is to take possession of a variant of SiaP that is compatible with a Lateral Flow Device (LFD), which can report detection of sialic acid in the nM range in 40% ethanol, can survive drying and temperature variations and be able to function when attached to a surface.

The current binding protein-based diagnostic device from Authentix uses a sugar-binding protein that recognises a polysaccharide with low affinity and a disaccharide with high affinity. SiaP is a candidate for adaptation to this diagnostic system since it is also reported to bind sialyllactose with a  $K_d$  value of 18  $\mu$ M (Severi *et al.*, 2005). The affinity of SiaP for sialic acid is a critical component of its exploitation. The  $K_d$  value needs to be decreased to at least 30 nM (~10 parts per billion), which is about a four-fold decrease on the currently reported value.

This chapter covers the investigation of the biophysical properties of SiaP that are related to the interests of Authentix and their ultimate impact upon the final product.

##### **4.1 Investigation of the biophysical properties of SiaP-His<sub>6</sub>**

The effect of the reaction buffer and conditions were investigated, since these could have a large impact upon the ligand binding affinity. The effect of organic solvent tolerance was investigated along with those of pH, salt and temperature, since a potential market for labelling would be products containing the solvents such as ethanol. These were all initially studied using the F170W reporter variant of SiaP due to its improved signal for ligand binding.

#### 4.1.1 Ethanol decreases ligand binding and promotes $\alpha$ -helix formation

The addition of organic solvents, such as alcohols, is known to affect the structure and function of proteins in solution. Alcohols disrupt hydrophobic interactions and favour polar interactions (Buhrman *et al.*, 2003, Deshpande *et al.*, 2005) such as amide–amide hydrogen bonds, which cause the formation of regions of high  $\alpha$ -helical content (Knubovets *et al.*, 1999). The effect of ethanol on the affinity of sialic acid binding by SiaP was investigated.

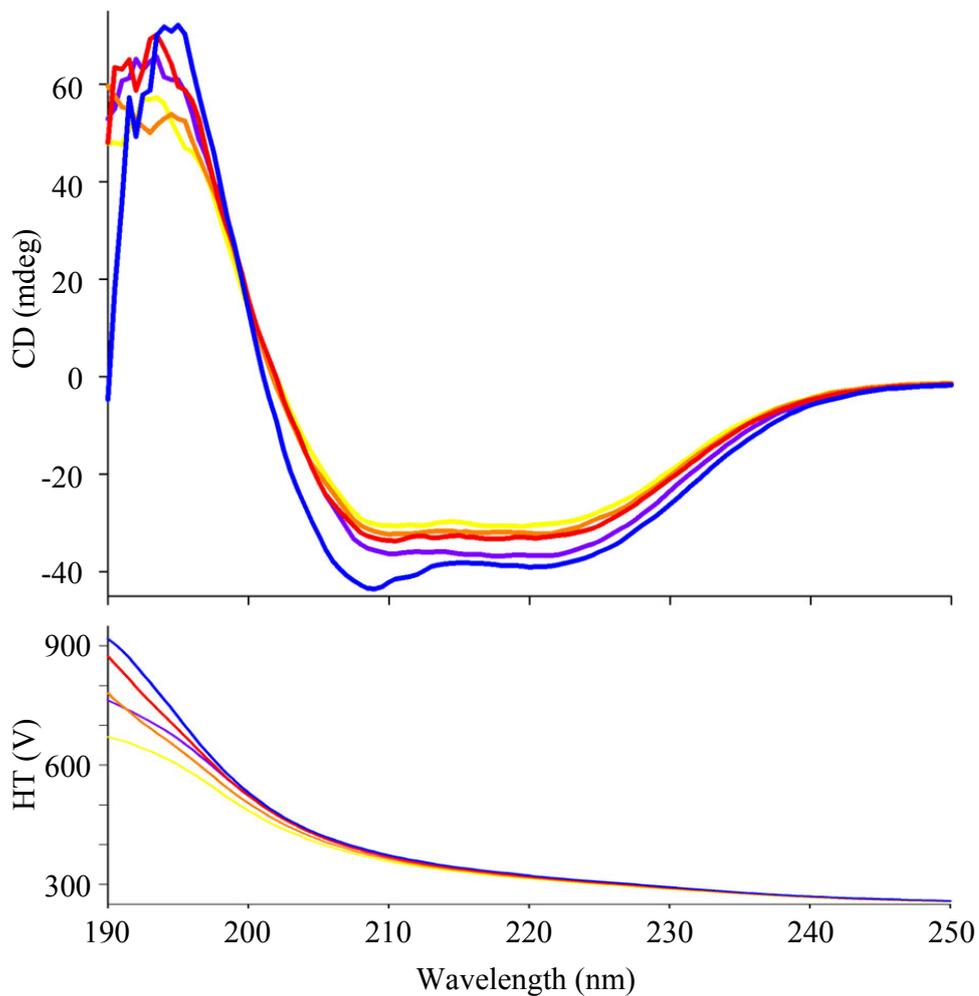
The ligand binding affinity of SiaP for sialic acid in the presence of ethanol was determined by fluorescence titration of the SiaP-His<sub>6</sub>:F170W reporter variant. Under the conditions used, the addition of 50 % ethanol gave a large fluorescence signal, preventing titration of the protein fluorescence emission. However, ligand binding to 0.05  $\mu$ M protein in 50 mM Tris/HCl pH 8.0 at 37 °C with up to 10 % ethanol could be determined (Table 4.1). The addition of 1% ethanol does not appear to have an effect on the  $K_d$  value of SiaP-His<sub>6</sub>:F170W for sialic acid. However, above this concentration, the affinity decreases by about 65% to  $1.9 \pm 0.2 \mu$ M with 5% ethanol and  $1.8 \pm 0.1 \mu$ M with 10 % ethanol.

To investigate this ethanol effect on the structure of SiaP-His<sub>6</sub>, the CD spectrum of the protein was determined in the presence of increasing concentrations of ethanol. The spectra of 0.2 mg/ml SiaP-His<sub>6</sub> was determined between 190 – 250 nm in the presence of 1 %, 5 %, 10 % and 50 % ethanol in 20 mM potassium phosphate buffer pH 8.0 at 37 °C (Figure 4.1).

As can be seen in Figure 4.1, the addition of ethanol causes a shift away from the spectrum of SiaP-His<sub>6</sub> with a lower intensity signal. This shift between 0% (red) and 1% ethanol (turquoise) is slight, while 5% (light green), 10% (dark green) and 50% ethanol (blue) have similar spectra with lower ellipticity. The shorter wavelengths (195 – 210 nm) appear to be more affected by the presence of ethanol.

**Table 4.1:** Sialic acid-binding affinities of SiaP-His<sub>6</sub>:F170W in the presence of an increasing concentration of ethanol.

Ethanol concentration	$K_d$	K	Relative affinity
0	$1.21 \pm 0.03 \mu\text{M}$	$8.26 \pm 0.20 \times 10^5 \text{ M}$	$100 \pm 2 \%$
1 %	$1.12 \pm 0.14 \mu\text{M}$	$8.93 \pm 0.99 \times 10^5 \text{ M}$	$108 \pm 11 \%$
5 %	$1.91 \pm 0.23 \mu\text{M}$	$5.34 \pm 0.66 \times 10^5 \text{ M}$	$65 \pm 12 \%$
10 %	$1.82 \pm 0.09 \mu\text{M}$	$5.49 \pm 0.26 \times 10^5 \text{ M}$	$66 \pm 5 \%$



**Figure 4.1:** CD spectra of 0.2 mg/ml SiaP-His<sub>6</sub> (yellow) in the presence of 1% (orange), 5% (red), 10% (purple) and 50% (blue) ethanol. The lower section shows the potential applied to the detector, which indicates decreased reliability of the spectrum at values greater than 800 V.

From the appearance of the spectra, it would appear that the addition of ethanol causes a slight change in the fold of the protein, indicating some sensitivity to high concentrations of ethanol. The secondary structural content of the protein at each concentration of ethanol was determined using the programme CDNN (Bohm *et al.*, 1992). A CD spectrum contains information on the overall secondary structure, which this programme can deconvolute to give the relative amounts of  $\alpha$ -helix,  $\beta$ -sheet,  $\beta$ -turn and random coil. As can be seen in Table 4.2, the secondary structural profiles below 20% ethanol are very similar. In 50 % ethanol, the secondary structural content is significantly altered, with a large increase in  $\alpha$ -helical content and decrease in random coil. This behaviour is similar to that of other proteins in alcohol, where this increase in  $\alpha$ -helical content occurs with a loss of tertiary structure, eventually resulting in a molten globule-like state (Knubovets *et al.*, 1999).

That ethanol causes a decrease in the ligand affinity of SiaP would be a problem for the application of this protein to a LFD. However, the affinity does not appear to decrease severely between 5 and 10 % ethanol and so, in the finished product, it may be necessary to dilute the testing sample so that the concentration of ethanol is within this range. The decrease in affinity seen here is likely due to the ethanol causing structural changes to the protein.

#### **4.1.2 Ligand binding affinity is affected by ionic strength and the presence of sodium ions**

It is clear that the composition and conditions of the reaction buffer will have a great effect upon the ligand binding affinity, for example, the pH can have a strong effect on the enthalpy of association of a complex (Chervenak & Toone, 1995). To investigate this, the pH-dependence of the sialic acid binding affinity was first determined.

**Table 4.2:** Relative content of secondary structural elements of SiaP-His<sub>6</sub> determined by CD spectra. The percentage of each element is shown as a percentage at each concentration of ethanol.

Secondary structural elements	Ethanol Concentration				
	0 %	1 %	10 %	20 %	50 %
$\alpha$ -helical	52 %	61 %	54 %	59 %	72 %
$\beta$ -sheet	7 %	5 %	7 %	7 %	4 %
$\beta$ -turn	14 %	13 %	14 %	13 %	11 %
Remainder (random coil)	27 %	21 %	25 %	21 %	13 %

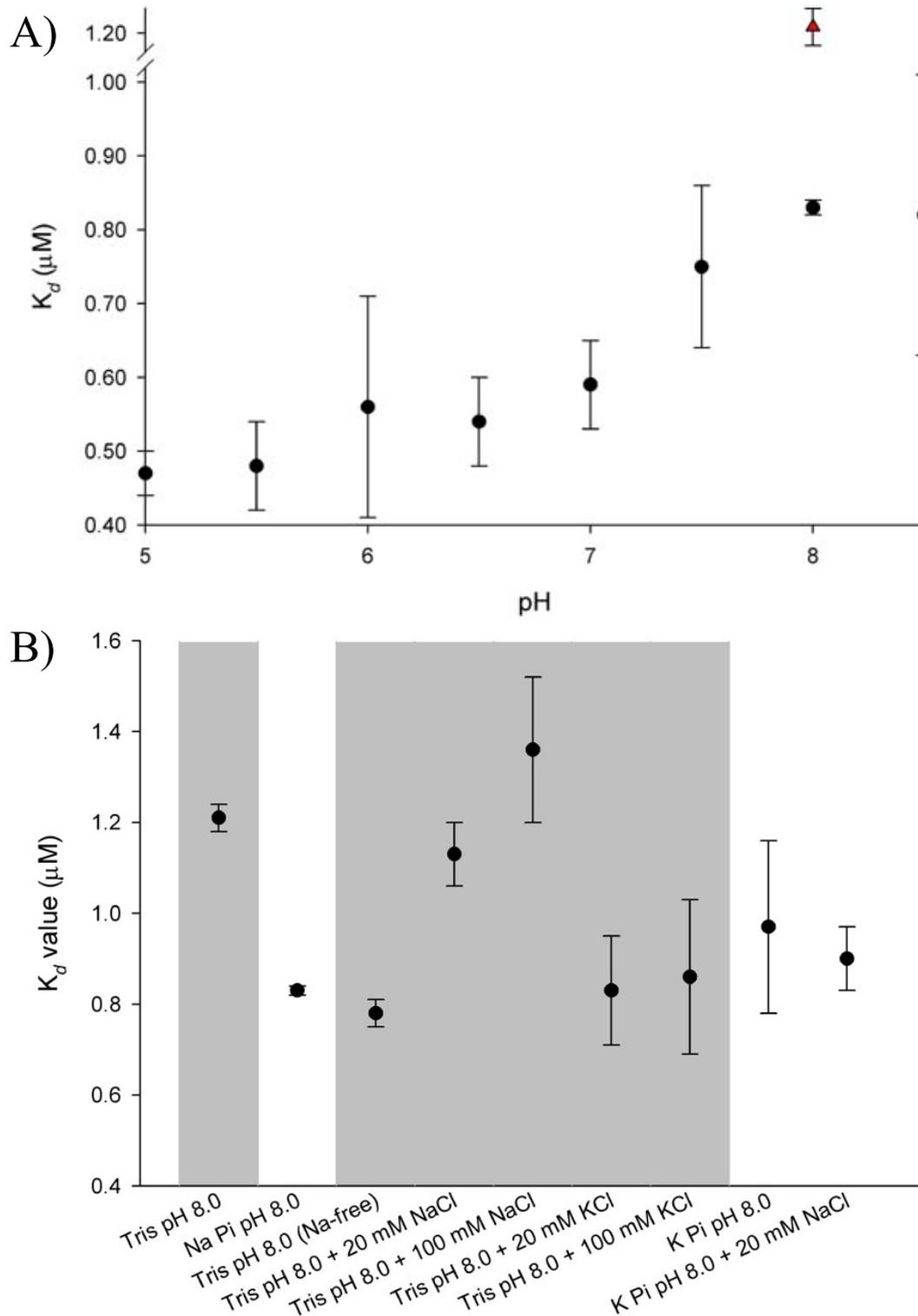
This was monitored using the SiaP-His<sub>6</sub>:F170W reporter variant. The  $K_d$  for sialic acid was determined by fluorescence titration between pH 5 – 9 in 20 mM sodium phosphate buffer at 37 °C.

As can be seen in Figure 4.2a, there is a clear decrease in the  $K_d$  for sialic acid when the pH drops to 7.5. The data also show an increase in affinity from  $1.21 \pm 0.03 \mu\text{M}$  in 50 mM Tris buffer to  $0.83 \pm 0.01 \mu\text{M}$  in the sodium phosphate buffer at pH 8.0.

This pH effect is not expected to be due to altering the protonation state of the critical binding site residues, since their  $\text{pK}_a$  values are not in this range, nor that of sialic acid, which has a  $\text{pK}_a$  value of 2.60 (Dawson, 1986). The only residue with an expected  $\text{pK}_a$  value in this range and near the binding cleft of SiaP is Aspartate-213, the side chain of which is predicted to have a  $\text{pK}_a$  of approximately 6.6 by the programme H<sup>++</sup> (Gordon *et al.*, 2005, Anandakrishnan & Onufriev, 2008). Asp-213 is located in the shell of residues around the ligand binding residues and near to the  $\beta$ -sheets of the hinge. However, it is more likely that the changing pH has acted through the sensitivity of SiaP to the ionic strength of the buffer.

The pH effect on the  $K_d$  of SiaP-His<sub>6</sub>:F170W was determined in 20 mM NaP<sub>i</sub> buffer, in contrast to all previous measurements in 50 mM Tris/HCl pH 8.0. It was noted in Figure 4.2a that this caused an increase in affinity from  $1.21 \pm 0.03 \mu\text{M}$  in Tris/HCl buffer to  $0.83 \pm 0.01 \mu\text{M}$  in sodium phosphate buffer at pH 8.0. Consequently, the effects of buffer of varying ionic strength would have to be examined.

The effects of the buffer were investigated initially by adding NaCl and KCl to 50 mM Na<sup>+</sup>-free Tris/HCl pH 8.0 and then to Na<sup>+</sup>-free 20 mM potassium phosphate buffer pH 8.0 (Figure 4.2b).



**Figure 4.2:** The effect of pH and ionic strength on the affinity of SiaP-His<sub>6</sub>:F170W for sialic acid, determined by fluorescence titration of 0.05 μM protein at 37 °C in triplicate. **A)** The pH-dependence of sialic acid affinity. This was determined in 20 mM sodium phosphate buffer. For comparison, the sialic acid affinity of SiaP-His<sub>6</sub>:F170W in 50 mM Tris/HCl pH 8.0 is shown in red. **B)** Sialic acid-binding affinities of SiaP-His<sub>6</sub>:F170W under different buffer conditions. These were determined in 50 mM Tris/HCl (grey background) or 20 mM phosphate buffer (white background).

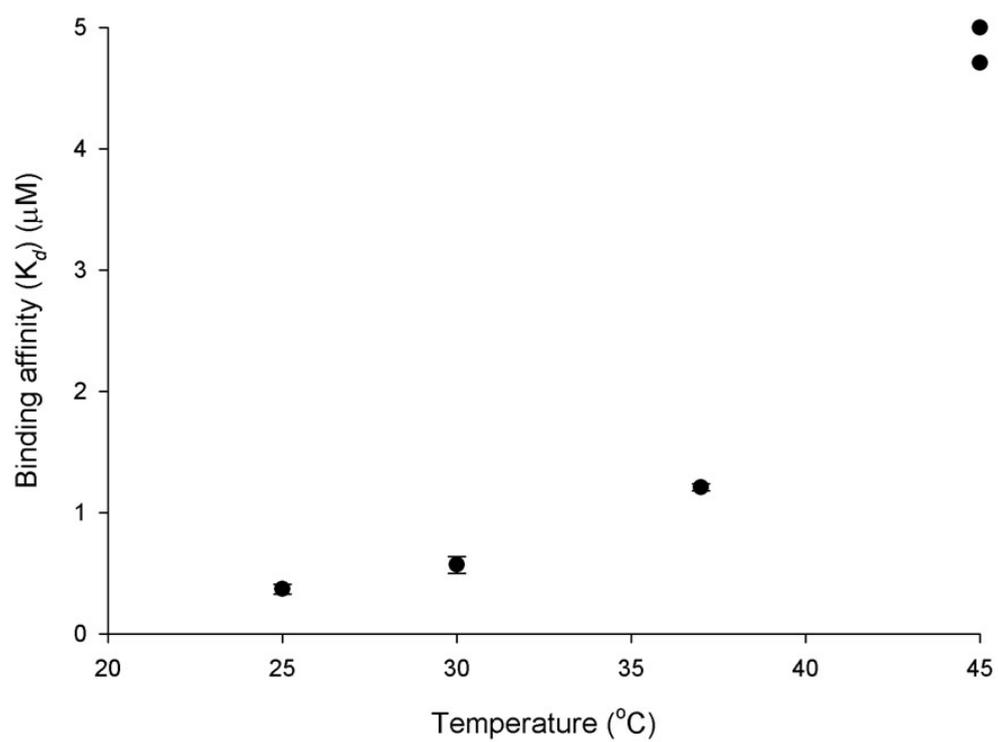
The affinity of SiaP-His<sub>6</sub>:F170W in Na<sup>+</sup>-free Tris/HCl pH 8.0 is  $0.78 \pm 0.03 \mu\text{M}$  compared to  $1.21 \pm 0.01 \mu\text{M}$  in Tris buffer made in glass beakers. The addition of NaCl to Na<sup>+</sup>-free Tris/HCl buffer decreases the affinity to the expected range of  $1.36 \pm 0.16 \mu\text{M}$ . This effect is not seen with the addition of KCl to Na<sup>+</sup>-free Tris/HCl. The  $K_d$  values in Na<sup>+</sup>-free Tris/HCl and sodium phosphate buffers are very similar and so the affinity was determined in potassium phosphate buffer. As can be seen in Figure 4.2b, the affinity here is not decreased, but does not appear to decrease on the addition of NaCl.

These results show that in Tris-buffer Na<sup>+</sup> ions and not K<sup>+</sup> ions have a detrimental effect on the  $K_d$ . However, in sodium phosphate buffer, the affinity for the ligand is higher than in Tris-buffer, indicating that phosphate buffer would have to be used to attain higher affinity detection of sialic acid. Surprisingly, the  $K_d$  value in potassium phosphate buffer is mid way between sodium phosphate and Tris buffer, but is insensitive to Na<sup>+</sup> ions. From these and due to the difficulties of avoiding even slight contamination with Na<sup>+</sup> ions, the highest sialic acid-binding affinity could be determined in 20 mM sodium phosphate buffer pH 6.0.

#### **4.1.3 Decreasing temperature increases ligand binding affinity**

Due to the nature of the ligand binding mechanism, it would be expected that temperature would have an effect on the affinity of this interaction. Since the SiaP-LFD would be used at ambient room temperature, the effect of temperature on ligand affinity was of interest. Due to the ease of its analysis, the SiaP-His<sub>6</sub>:F170W reporter variant was used to monitor the affinity for sialic acid at different temperatures.

The  $K_d$  values were determined in triplicate using fluorescence titration at 25, 30, 37 and 45 °C. This range was chosen to remain below the melting temperature of SiaP-His<sub>6</sub> (50–60 °C) and above room temperature. The  $K_d$  values approach a minimum value around 0.3  $\mu\text{M}$ , increasing as they approach the melting temperature of the protein (Figure 4.3).



**Figure 4.3:** The effect of temperature on the affinity of SiaP-His<sub>6</sub>:F170W for sialic acid. This was determined by fluorescence titration of 0.05 μM protein in 50 mM Tris/HCl pH 8.0 in triplicate at each temperature point, except for 45 °C, which is an average of two results (shown).

This minimum  $K_d$  value around 25 °C is serendipitous, since the expected operating temperature of an LFD would be room temperature, outside or in a warehouse. However, this data refers to the SiaP-His<sub>6</sub>:F170W reporter variant, not the native form.

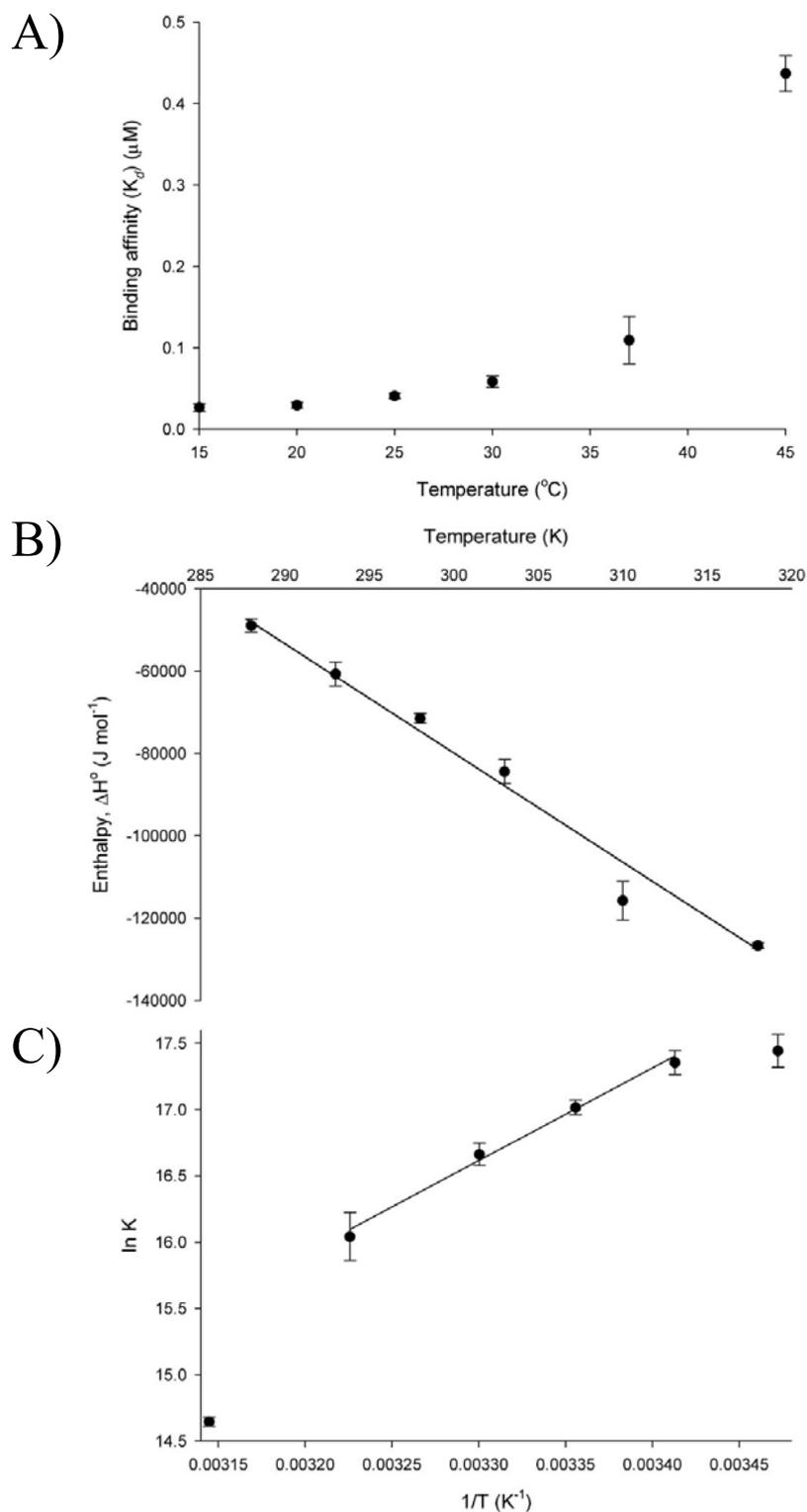
#### 4.1.4 Sialic acid-binding by SiaP is enthalpically driven and releases water

Since the  $K_d$  value of SiaP-His<sub>6</sub>:F170W for sialic acid is ten times that of SiaP-His<sub>6</sub> at 37 °C, it was assumed that this difference would be equivalent at each temperature point, giving initial protein and ligand concentrations that could be used to perform isothermal titration calorimetry (ITC) on SiaP-His<sub>6</sub>.

ITC was performed in triplicate on various concentrations of SiaP-His<sub>6</sub> in 50 mM Tris/HCl pH 8.0 at 15, 20, 25, 30, 37 and 45 °C. The  $K_d$  values at each temperature are shown in Figure 4.4a and the thermodynamic data in Table 4.3. Figure 4.4a shows that, as temperature decreases, the  $K_d$  for sialic acid approaches a lowest value around 30 nM.

The ITC also gives a wealth of thermodynamic data (Table 4.3). At 37 °C, the binding enthalpy,  $\Delta H^\circ$ , is -116 kJ mol<sup>-1</sup>, which is a very large value that is favourable for the binding event to occur. This indicates that a large number of bonding interactions are made upon ligand binding. Conversely, the binding entropy,  $-T\Delta S^\circ$ , is 74 kJ mol<sup>-1</sup> which is also very large, but unfavourable, indicating that there is a large loss in degrees of freedom upon ligand binding. The Gibbs free energy,  $\Delta G^\circ$ , is -42 kJ mol<sup>-1</sup>, indicating that the binding event is favourable, driven by the enthalpic change.

Plotting the binding enthalpy with changing temperature reveals the heat capacity of the system,  $C_p$ , to be -2.75 kJ mol<sup>-1</sup> K<sup>-1</sup> (Figure 4.4b). A negative heat capacity corresponds to polar solvation, such as the releasing of water from a lattice (Chervenak & Toone, 1995), which is in line with the mechanism of ligand binding and domain closure. That is, the dehydration of sialic acid and the ejection of numerous water molecules from the open binding cleft.



**Figure 4.4:** The thermodynamics of sialic acid binding by SiaP-His<sub>6</sub>. **A)** The effect of temperature on the affinity of SiaP-His<sub>6</sub> for sialic acid. This was determined by ITC using varying concentrations of protein in 50 mM Tris/HCl pH 8.0 in triplicate at each temperature point. **B)** Graph of the effect of temperature on binding enthalpy. The gradient of this line gives the heat capacity of the system. **C)** The van't Hoff plot of the natural log of the binding affinity,  $K$ , against the inverse of the temperature. The gradient and intercept of the short linear range between 20 °C and 37 °C were determined from the line shown.

**Table 4.3:** Thermodynamic data from the ITC analysis of SiaP-His<sub>6</sub>.

Temperature (°C)	K <sub>d</sub> (nM)	ΔH (J/mol)	ΔS (J/mol/deg)	ΔH-TΔS (J/mol)
15	26.7 ± 0.33	-48952 ± 1629	-24.9 ± 5.7	-41779 ± 1244
20	29.1 ± 0.26	-60737 ± 2908	-63.0 ± 10.2	-42267 ± 2195
25	40.8 ± 0.23	-71518 ± 1170	-98.3 ± 3.6	-42218 ± 911
30	58.2 ± 0.49	-84419 ± 2959	-139.9 ± 10.1	-42034 ± 2228
37	109.1 ± 2.1	-115785 ± 4727	-239.9 ± 14.0	-41422 ± 3690
45	437.0 ± 1.5	-126663 ± 635	-276.3 ± 2.4	-38805 ± 456

The enthalpy of a reaction can also be determined from the equilibrium constant at multiple temperatures where the reaction is a simple bi-molecular association, as is expected for SiaP (Severi *et al.*, 2005). However, this assumes that  $\Delta C_p$  is independent of temperature. The van't Hoff plot of  $\ln(K_{eq})$  against  $1/T$  allows  $\Delta H_{vH}$  and  $\Delta S_{vH}$  to be determined from the gradient and  $y$ -axis intercept over narrow temperature ranges (Figure 4.4c). Over the approximately linear range between 20 °C and 37 °C,  $\Delta H_{vH} = -1762 \text{ J mol}^{-1}$ , and  $\Delta S_{vH} = -5.2 \text{ J mol}^{-1} \text{ K}^{-1}$ , which are different to the experimentally determined values of  $\Delta H^\circ$  and  $\Delta S^\circ$ . This indicates that SiaP-His<sub>6</sub>-sialic acid binding is not a simple one step process, such as  $A + B \leftrightarrow A \cdot B$ , but one with multiple steps before forming the  $A \cdot B$  complex.

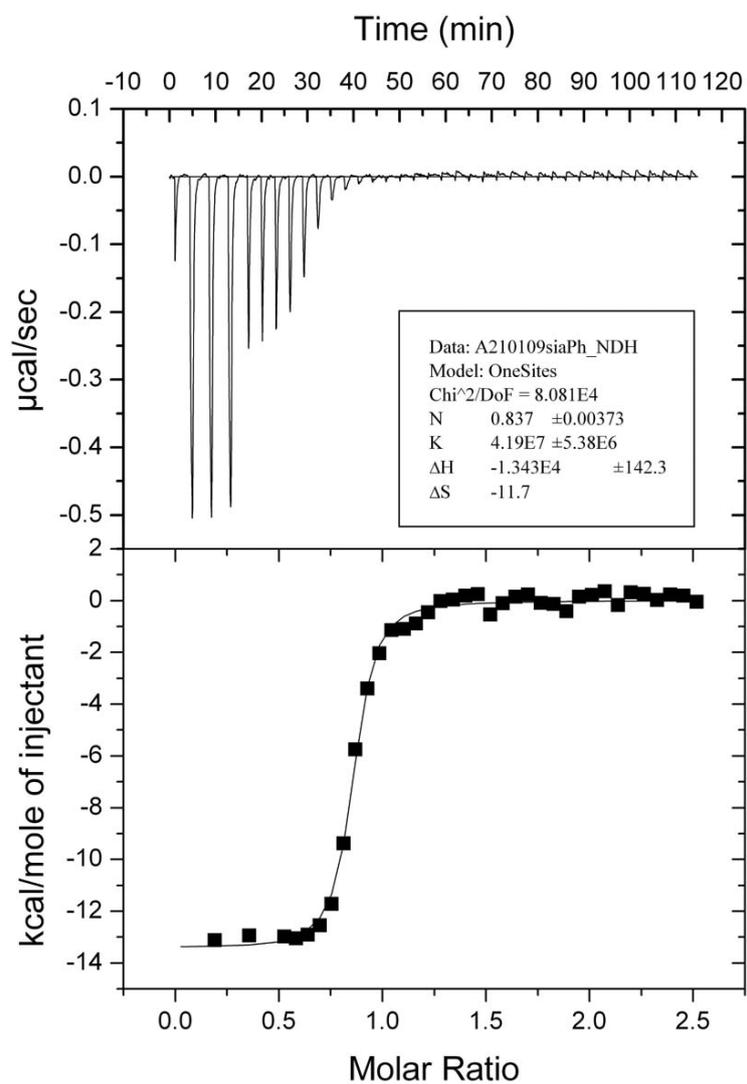
This contradicts the previous finding that SiaP binds sialic acid as a simple, bi-molecular association (Severi *et al.*, 2005). However, it has previously been shown that DctP does not follow this simple reaction, but exists in equilibrium between closed-unliganded, open-unliganded, open-liganded and closed-liganded forms (Walmsley *et al.*, 1992). This contradiction cannot be resolved at this point and requires further investigation.

#### **4.1.5 Optimised conditions for high affinity binding of Neu5Ac by SiaP**

From the findings in this section, the temperature, pH and buffer composition effects were combined to give the optimal conditions for the highest sialic acid-binding affinity of SiaP-His<sub>6</sub>.

The SiaP-His<sub>6</sub>-sialic acid interaction was examined by ITC in triplicate at 15 °C in 20 mM sodium phosphate buffer pH 6.0 using 10  $\mu\text{M}$  SiaP-His<sub>6</sub> titrated by the injection of 150  $\mu\text{M}$  sialic acid (Figure 4.5).

From these, the  $K_d$  value was determined as  $22 \pm 2 \text{ nM}$ , or  $6.8 \pm 1.2$  parts per billion (ppb) (Figure 4.5). This is the highest affinity for sialic acid measured in this project and is inside the range required by Authentix.



**Figure 4.5:** ITC titration of SiaP-His<sub>6</sub> in 20 mM sodium phosphate buffer pH 6.0 at 15 °C. The first injection of 2 µl was followed by three of 18 µl and then a series of 6 µl injections.

In the 50 mM Tris/HCl buffer at 15 °C,  $\Delta H^\circ$  is  $-49.0 \text{ kJ mol}^{-1}$  and  $-T\Delta S^\circ$  is  $7.2 \text{ kJ mol}^{-1}$ , where enthalpy is large and favourable and entropy is unfavourable. The Gibbs free energy,  $\Delta G^\circ$  ( $\Delta H - T\Delta S$ ), is  $-41.8 \text{ kJ mol}^{-1}$  so that binding is favourable and driven by the enthalpic change. In 20 mM sodium phosphate buffer,  $\Delta H^\circ$  is  $-53.5 \text{ kJ mol}^{-1}$ ,  $-T\Delta S^\circ$  is  $11.3 \text{ kJ mol}^{-1}$  and  $\Delta G^\circ$  is  $-42.2 \text{ kJ mol}^{-1}$ . From these, there is little change in the  $\Delta G^\circ$  between both buffers; however, switching from Tris/HCl to sodium phosphate buffer causes a change in  $\Delta H^\circ$  of  $-4.55 \text{ kJ mol}^{-1}$  and in  $-T\Delta S^\circ$  of  $4.1 \text{ kJ mol}^{-1}$ . This change in enthalpy suggests that, with sodium phosphate, unliganded SiaP-His<sub>6</sub> forms more interactions when in the open conformation or that fewer interactions are made on ligand binding. The change in entropy indicates that the system becomes more ordered upon ligand binding in sodium phosphate than Tris/HCl buffer. However, it is difficult to be concrete about these changes since they are a combination of many effects.

## **4.2 Detection of sialic acid**

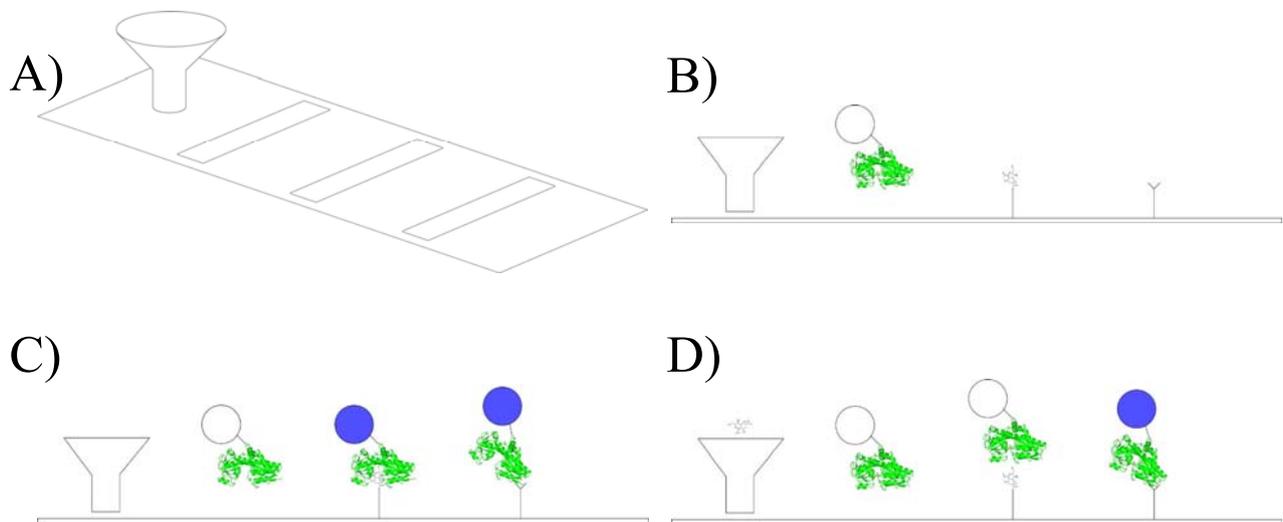
### **4.2.1 The Lateral Flow Device**

Currently, Authentix uses a substrate binding protein-based assay to detect the addition of the ligand in a labelled product. The current SBP-based Lateral Flow Device (LFD) uses the high affinity ligand to inhibit binding of the SBP to a lower affinity glyco-conjugated ligand, which is attached to the surface of the LFD (Figure 4.6). The presence or absence of the SBP on the conjugated ligand strip shows the presence of the high affinity ligand label in the sample.

### **4.2.2 Conjugated sialic acid as a basis for a Lateral Flow Device**

Since SiaP had previously been shown to bind the sialoconjugate sialyllactose with a  $K_d$  of  $18 \mu\text{M}$  (Severi *et al.*, 2005), this was taken to mean that SiaP might be suitable for a complementary LFD. Several sialic acid-conjugated ligands were tested for binding by SiaP-His<sub>6</sub>.

Neu5Ac, colominic acid (poly Neu5Ac) and porcine mucin (a highly sialylated protein) were chosen as potential substrates for SiaP. Also, Neu5Ac was conjugated to BSA using NaIO<sub>4</sub>-activation of a ring hydroxyl to form a covalent interaction with the BSA



**Figure 4.6:** The principle of a lateral flow device using SiaP. **A)** Three dimensional representation of the LFD. The small funnel represents the loading site of the sample; the sample is then drawn past three strips by capillary action. **B)** As the sample is drawn along the device by capillary action, it first encounters dried SiaP bound to microbeads. These are then washed on to the first strip that contains surface-bound sialic acid. The second strip contains an antibody to SiaP and so acts as a control for the LFD. **C)** With an unlabelled sample, that is, one which contains no sialic acid, the unliganded SiaP will bind to the first strip and give a signal. The excess SiaP will then continue on to the third strip and give a second signal, serving as a positive control. **D)** If the sample contains sialic acid, it will be bound by SiaP, which is then unable to interact with the surface-associated sialic acid and will not give the first signal. The SiaP will then continue on to the final strip, where it will give a final signal, indicating that the LFD has worked.

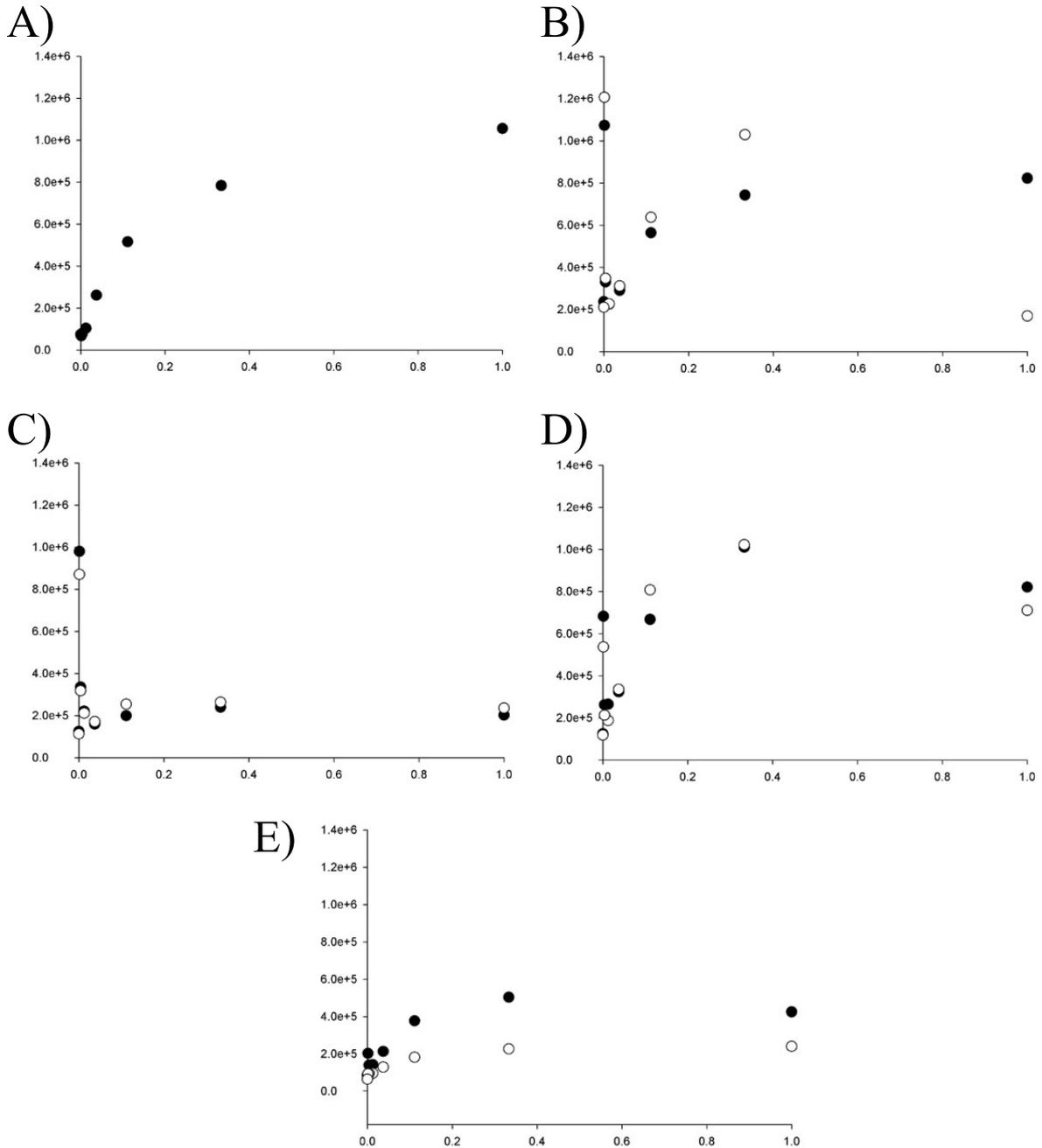
terminus (Telmer & Shilton, 2003). Serial dilutions of these potential substrates were dried onto the surface of a clear 96-well plate and washed with PBS 0.1 % Tween-20 (TPBS). The remaining surface of the wells was blocked with TPBS 4 % BSA and then washed with TPBS again. A serial dilution of SiaP-His<sub>6</sub> was incubated in each well, followed by Mouse anti(tetrahistidine) antibodies and HRP-conjugated Rabbit anti(mouse) antibodies with TPBS washing between each step. These were then measured by the addition of 50 µl SuperSignal and the luminescence monitored at 425 nm.

As can be seen in Figure 4.7, there appears to be SiaP-His<sub>6</sub> present with all substrates including the blocking control (BSA). All but BSA-Neu5Ac show some increase in signal with increase concentration of SiaP-His<sub>6</sub>. In fact, the clearest titration evident is for the BSA blocking control and all other substrates seems to cause a decrease from this maximal intensity.

This interaction of SiaP-His<sub>6</sub> with the blocking control could have been due to contamination of the crude BSA with some sialylated protein or an unfortunate physical interaction between SiaP-His<sub>6</sub> and BSA. At this point, it was decided that it would be more efficient to screen multiple glycans as targets for SiaP-His<sub>6</sub> using a microarray approach.

#### **4.2.3 SiaP is specific for monomeric sialic acid**

The Consortium for Functional Glycomics provides a service whereby glycan binding proteins, such as lectins, are screened for their interaction partners (CFG, 2010). This meant that many more sialoglycans could be screened for an interaction with SiaP-His<sub>6</sub>. This is performed on a GlycoArray (mammalian printed array version 3.2), which is a chip that is spotted with four hundred and six different glycans, allowing for the



**Figure 4.7:** The presence of SiaP-His<sub>6</sub> in wells containing potential substrates detected by antibodies against the hexahistidine tag (units of intensity are arbitrary and have been normalized between each sample). Filled circles indicate 5 nanomoles substrate per well and empty circles indicate 0.5 nanomoles substrate per well. The concentration of SiaP-His<sub>6</sub> used (μM) is shown on the x-axis. **A)** Interaction with blocking BSA. **B)** Interaction with dried-down sialic acid. **C)** Interaction with sialic acid conjugated to BSA. **D)** Interaction with colominic acid (polysialic acid). **E)** Interaction with porcine mucin.

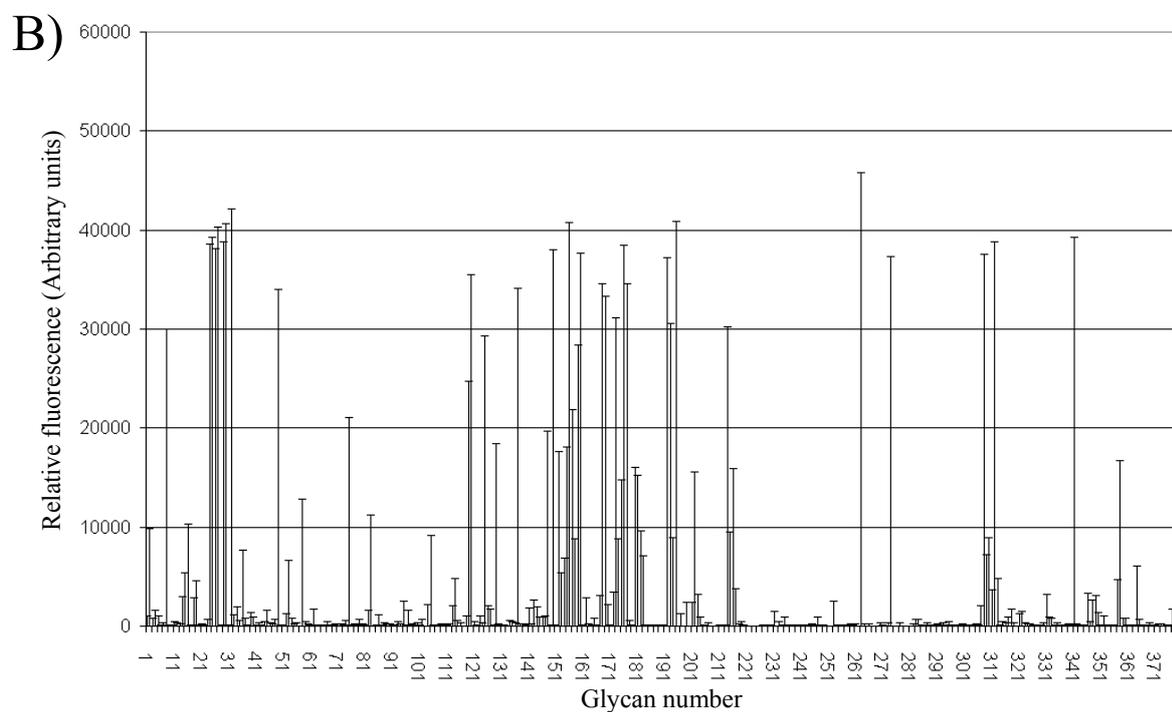
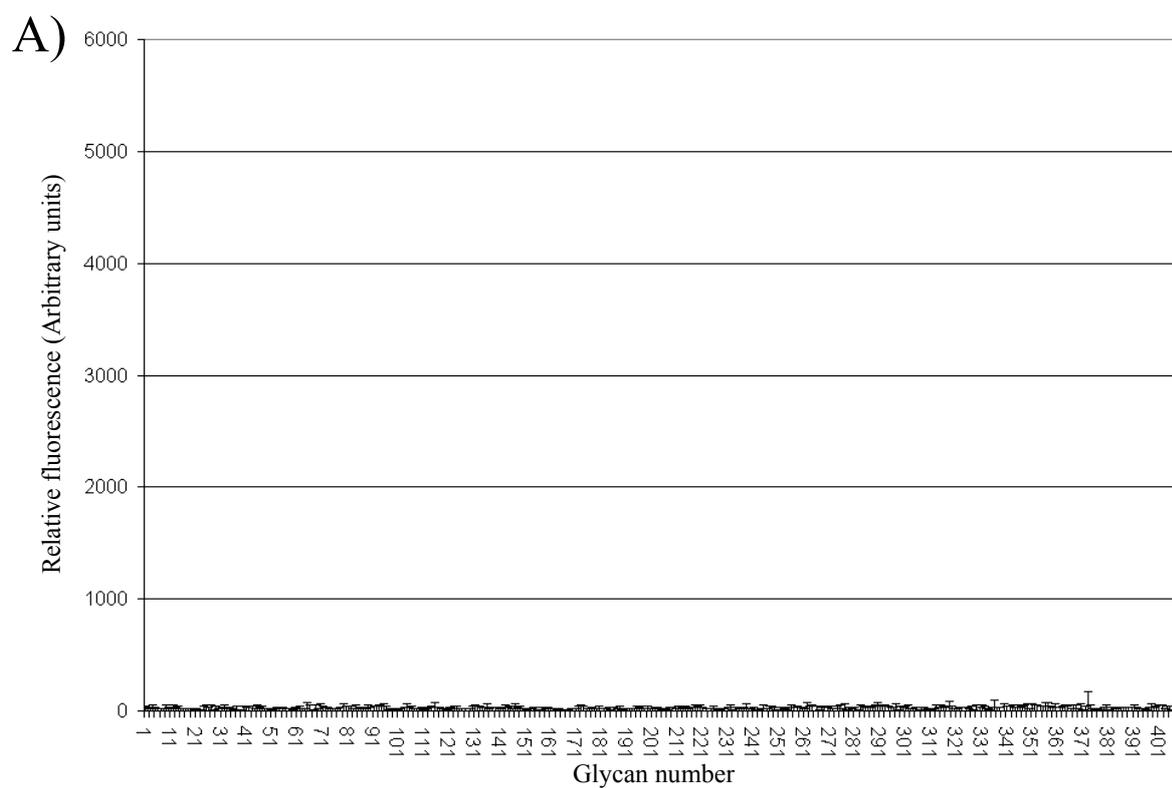
screening of glycan-binding proteins by antibody detection. SiaP-His<sub>6</sub> was examined on this array and the results are shown in Figure 4.8a (courtesy of the CFG).

These results indicate that SiaP-His<sub>6</sub> does not interact with any of the glycans displayed in the GlycoArray. A positive GlycoArray result provided by the CFG is shown in Figure 4.8b for comparison.

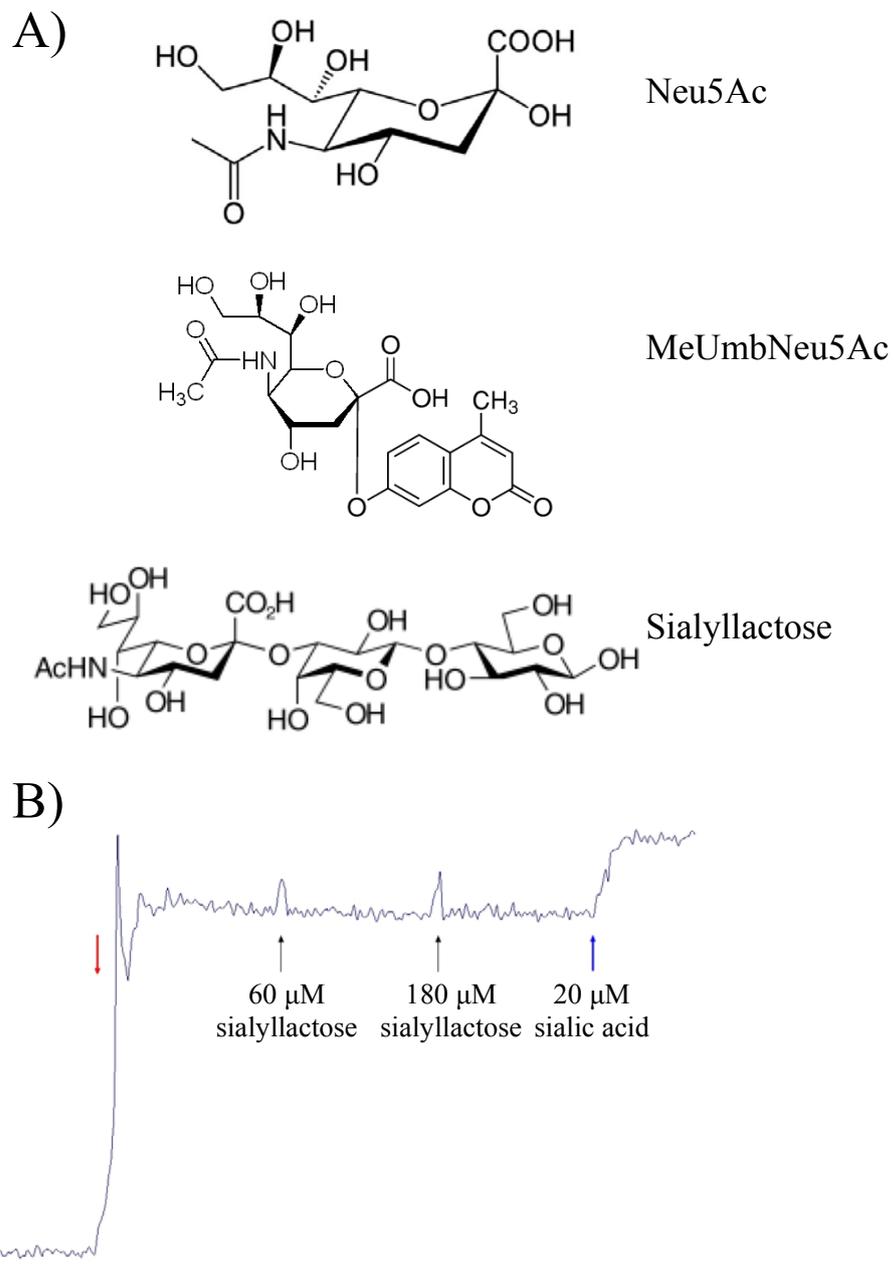
The negative results from the GlycoArray cast doubt on the previous finding that SiaP can bind sialyllactose (Severi *et al.*, 2005). To investigate this, binding of sialyllactose by SiaP-His<sub>6</sub> was determined by fluorescence signal change. A second modified sialic acid, methylumbelliferyl neuraminic acid (MeUmbNeu5Ac), was also investigated by fluorescence change. This is made up of sialic acid covalently modified by the addition of the small methylumbelliferyl fluorophore to the carboxylate group (Figure 4.9a). This fluorescently labelled sialic acid is commonly used as a substrate for sialidase assays, but the similar substrate 4-methylumbelliferyl- $\alpha$ -D-mannopyranoside has previously been used to investigate mannose-binding proteins, since it shows a fluorescence quench on its sequestration into a proteinacious environment (Thompson & Lakowicz, 1984, Kenoth *et al.*, 2003).

The interaction between SiaP-His<sub>6</sub> and sialyllactose was examined by the fluorescence change of SiaP-His<sub>6</sub> following the addition of up to 180  $\mu$ M sialyllactose. The interaction between SiaP-His<sub>6</sub> and MeUmbNeu5Ac was determined by the fluorescence change of the ligand following addition of 20  $\mu$ M SiaP-His<sub>6</sub>.

The fluorescence signal of SiaP-His<sub>6</sub> shows no change on the addition of sialyllactose (Figure 4.9b). The sequestration of the methylumbelliferyl fluorophore inside the protein upon ligand binding was expected to cause a decrease in ligand fluorescence (Thompson & Lakowicz, 1984, Kenoth *et al.*, 2003). Comparing the fluorescence emission spectra before and after the addition of SiaP-His<sub>6</sub> showed no change. This lack



**Figure 4.8:** Determination of glycan binding using a GlycoArray. **A)** Results from analysis of 200  $\mu\text{g/ml}$  SiaP-His<sub>6</sub> at a tenth of the scale of the positive result. **B)** Binding of a bacterial lectin to the GlycoArray. This is shown as an example of a glycan-binding protein.



**Figure 4.9:** The interaction of SiaP-His<sub>6</sub> with modified sialic acid. **A)** Structure of methylumbelliferyl neuraminic acid (MeUmbNeu5Ac). **B)** Fluorescence titration of SiaP-His<sub>6</sub> with sialyllactose. The red arrow indicated the addition of 0.2  $\mu$ M protein to the cuvette; sialyllactose was added to 60  $\mu$ M then 180  $\mu$ M (black arrows), followed by 20  $\mu$ M sialic acid (blue arrow).

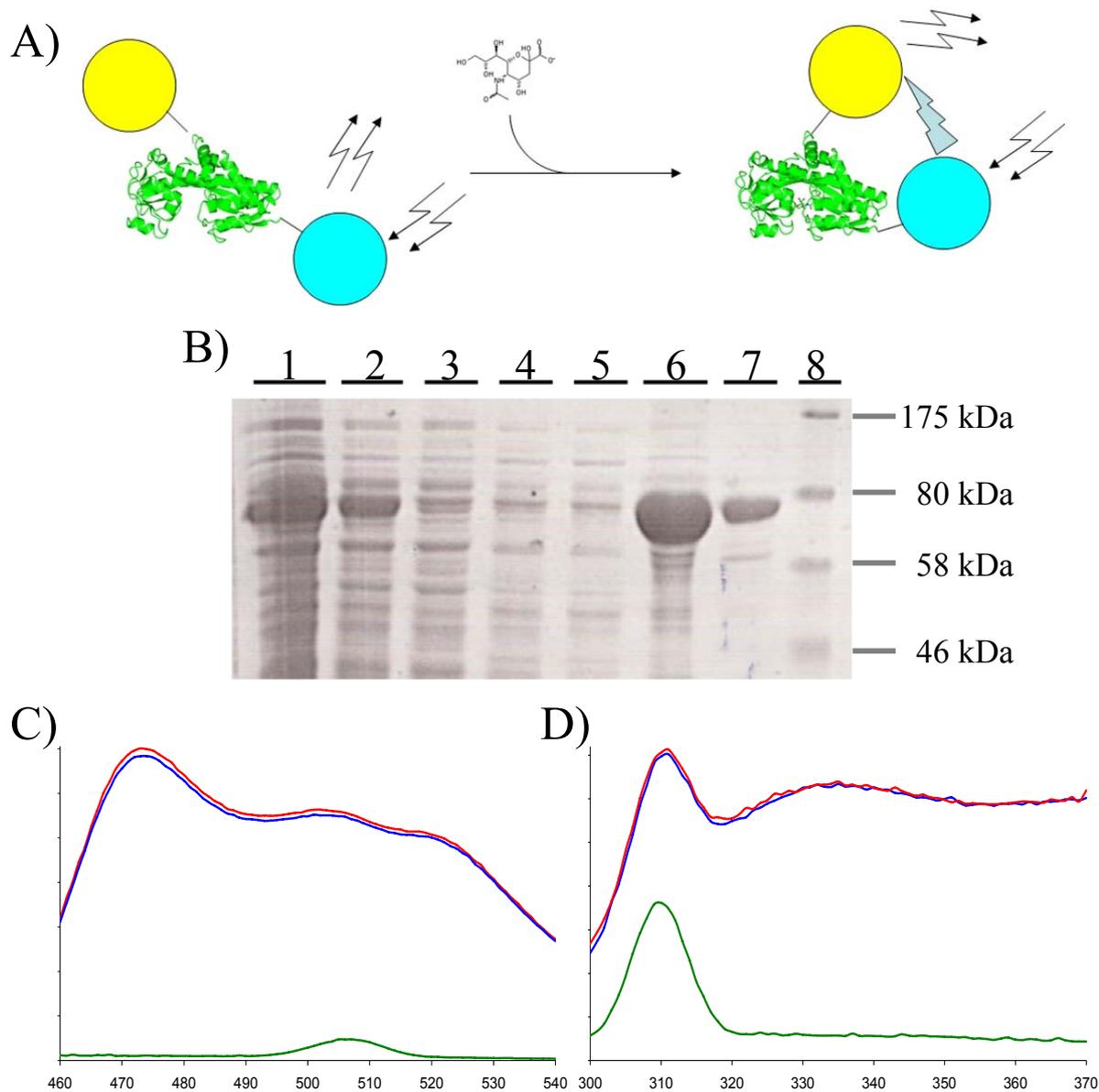
of interaction was also observed by surface plasmon resonance (SPR) of surface-attached SiaP-His<sub>6</sub> (Marcus Fischer, personal communication).

There is apparently no interaction between SiaP and sialyllactose or MeUmbNeu5Ac. In addition to the failure to find any conjugated sialic acid ligand for SiaP-His<sub>6</sub>, this indicates that SiaP specifically binds monomeric sialic acid.

#### 4.2.4 An in-solution approach for sialic acid detection

The results from the previous section show that SiaP is specific for monomeric sialic acid. This fundamentally changes the design for the SiaP-based sialic acid detector. Binding of a ligand by a substrate binding protein usually results in a large conformational change as the domains close around the ligand (Mao *et al.*, 1982). This can be taken advantage of in genetically-encoded nanosensors (Fehr *et al.*, 2002), where cyan- and yellow-fluorescent protein (CFP and YFP) are fused to the N- and C-termini of a truncated binding protein so that the conformational change on ligand binding causes a change in the FRET signal between the two fluorescent proteins (Figure 4.10a). Gu *et al.* (2006) created a nanosensor for inorganic phosphate, FLIPP<sub>i</sub>-260n, which used an enhanced CFP (eCFP) and a venus YFP (vYFP) fused to a phosphate binding protein (P<sub>i</sub>BP) from *Synechococcus* strain A. Creating a similar construct, FLIP-Neu5Ac, could produce a suitable sensor for sialic acid.

The *FLIPP<sub>i</sub>* gene construct was acquired through the plasmid database, Addgene Inc. The gene fusion for eCFP-P<sub>i</sub>BP-vYFP (FLIP-P<sub>i</sub>) was supplied on the IPTG-inducible plasmid, pRSET (Invitrogen), pRSET-FLIPP<sub>i</sub>-260n. The *p<sub>i</sub>bp* gene is inserted in the construct between two *KpnI* restriction enzyme (RE) sites (Gu *et al.*, 2006). After repeated failures in attempts to replace the *p<sub>i</sub>bp* gene with *siaP*, a new strategy was designed, where the upstream *KpnI* RE site was replaced by a *PinAI* RE site. After PCR amplification, *PinAI-siaP-KpnI* was blunt-end ligated into pBlunt0 (Invitrogen) to provide excess, correctly digested insert. This was successfully ligated into *PinAI* and *KpnI* digested pRSET-eCFP-vYFP vector to give pFLIP-Neu5Ac. *E. coli* BL21 (DE3) pLysS pFLIP-Neu5Ac



**Figure 4.10:** Construction of the FLIP-Neu5Ac nanosensor. **A)** The basis of the nanosensor. The *siaP* gene is inserted between eCFP and vYFP. On the addition of ligand, the FRET signal from the two fluorescent proteins changes due to the domain closure of the binding protein. **B)** Ni-affinity purification of FLIP-Neu5Ac. Lane 1: *E. coli* BL21 (DE3) pLysS pFLIP-Neu5Ac whole cell lysate; lane 2: clarified lysate; lane 3: Ni-column flow through; lane 4: Ni-column wash step; lanes 5–7: Ni-column elution fractions 1–3; lane 8: Molecular weight marker. **C)** Fluorescence emission spectra of FLIP-Neu5Ac (blue) and on the addition of 500  $\mu\text{M}$  sialic acid (red) when excited at 433 nm. The signal from the buffer is shown in green. **D)** Fluorescence emission spectra of FLIP-Neu5Ac (blue) and on the addition of 500  $\mu\text{M}$  sialic acid (red) when excited at 281 nm. The signal from the buffer is shown in green.

was grown in 500 ml M9 minimal medium supplemented with 0.4 % glucose at 25 °C to OD<sub>650</sub> 0.2 before induction with 1 mM IPTG in the dark for 48 hours. Harvested cells were resuspended in 50 mM potassium phosphate pH 8.0, sonicated for 10 minutes in total to lyse the cells and the supernatant was clarified by centrifugation. The FLIP-Neu5Ac-containing supernatant was incubated with 1 ml Ni-NTA resin (Qiagen) for 1 hour at 4 °C, washed with 50 mM Tris/HCl 300 mM NaCl (TBS) containing 20 mM imidazole and FLIP-Neu5Ac was eluted with 500 mM imidazole TBS. Sialic acid binding by FLIP-Neu5Ac was determined by fluorescence titration in 50 mM Tris/HCl pH 8.0 with excitation at 281 nm (tyrosine) and 433 nm (CFP).

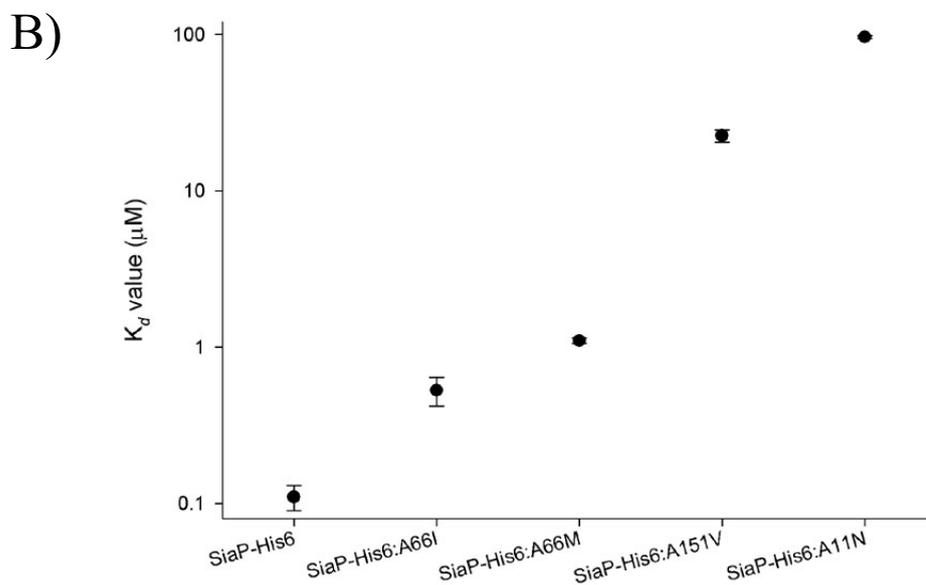
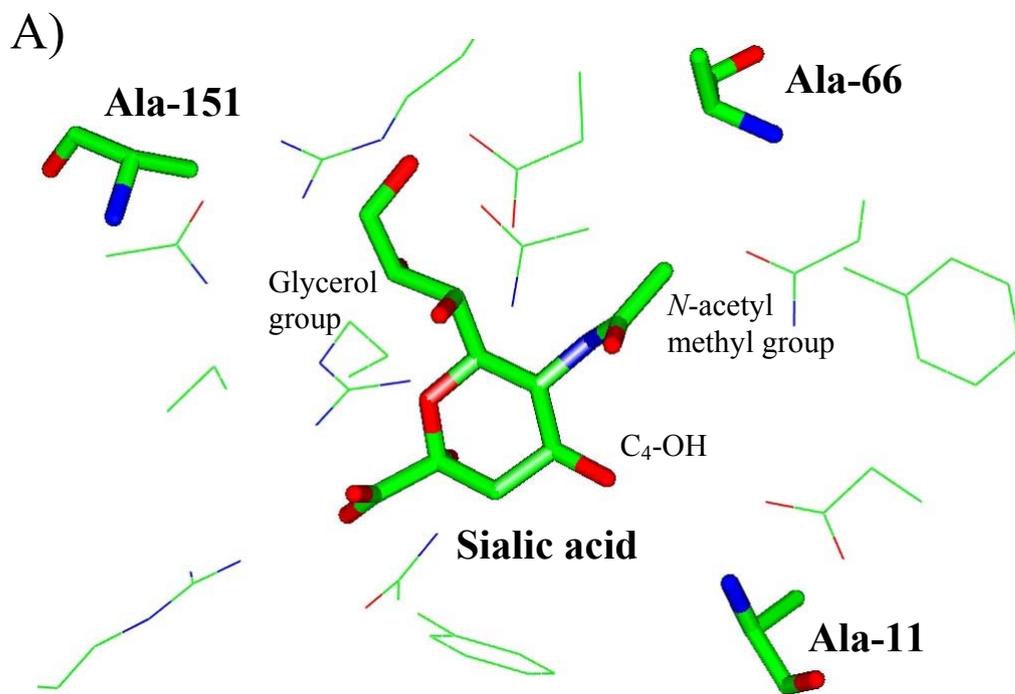
FLIP-Neu5Ac was expressed and purified in the same manner as FLIP-P<sub>i</sub>-260n (Gu *et al.*, 2006) (Figure 4.10b), eluting in the second fraction, which was bright yellow in appearance. The fluorescence emission spectra of 0.05 μM FLIP-Neu5Ac with and without 0.5 mM sialic acid were determined when exciting for tyrosine residues (281 nm) and CFP (433 nm). As can be seen, there is no reliable ligand-dependent change in the emission spectra when CFP or tyrosine residues are excited (Figure 4.14cd).

The appearance of three peaks around 475, 505 and 520 nm ( $\lambda_{\text{ex}}$  433 nm) suggest that eCFP and vYFP are correctly folded and close by in solution. Unfortunately, this shows no sensitivity to sialic acid. Nor could sialic acid binding be detected with excitation at 281 nm; this is likely to be due to the effect of introducing two proteins that are each similar in size to SiaP, whether by destabilising the fold or interfering with ligand binding.

### **4.3 Rational design of SiaP to modulate its sialic acid binding affinity**

#### **4.3.1 Binding site mutations designed to increase ligand affinity**

In collaboration with Professor Roderick Hubbard, the structure of sialic acid-bound SiaP was examined for changes in the protein structure that could increase contacts to the ligand, thereby potentially improving the affinity of binding. This analysis produced three possible changes to the protein that were targets for mutation (Figure 4.11a). The first candidate, Ala-11-Asn, was intended to form a hydrogen bond with the ring C<sub>4</sub>-hydroxyl



**Figure 4.11:** A family of binding site mutations designed to increase ligand affinity. **A)** Representation of the targets for increased interactions between Neu5Ac and the binding site of SiaP. **B)** Sialic acid-binding affinities of the ligand binding site mutant proteins. These were determined in triplicate in 50 mM Tris/HCl pH 8.0 at 37 °C.

group. The second and third targets, Ala-66 (to Ile and Met) and Ala-151 (to Val), aimed to provide a more hydrophobic environment for the CH<sub>3</sub> of the *N*-acetyl group and the glycerol backbone, respectively.

Using the same mutagenic technique as outlined previously, these four mutations were introduced into the *siaP-His<sub>6</sub>* allele present in the pAH16 construct. They were expressed and purified following the same protocol as for SiaP-His<sub>6</sub>. The affinities of the mutant proteins for sialic acid were determined by titration of the fluorescence signal of 0.05 μM protein. Where the *K<sub>d</sub>* value was above 1.1 μM, the protein concentration was increased to 0.25 μM to increase the fluorescence signal.

All of these mutant proteins gave clear fluorescence signal titration data, which are shown in Figure 4.11b. As can be seen, all of these mutations have caused a decrease in sialic acid binding affinity. The least damaging are the Ala-66 mutations, of which the shorter isoleucine mutant is least deleterious. The Ala-151-Val mutation causes a 200-fold decrease in affinity, probably by disturbing the water network around Arg-70 and Asn-154 and so affecting the entire environment around the glycerol group of the ligand. The most damaging of these mutations was Ala-11-Asn, originally intended to extend to the C<sub>4</sub>-hydroxyl and form a hydrogen bond. This causes approximately 1000-fold decrease in affinity and is likely not interacting with the C<sub>4</sub>-hydroxyl, but disturbing the ligand *N*-acetyl group.

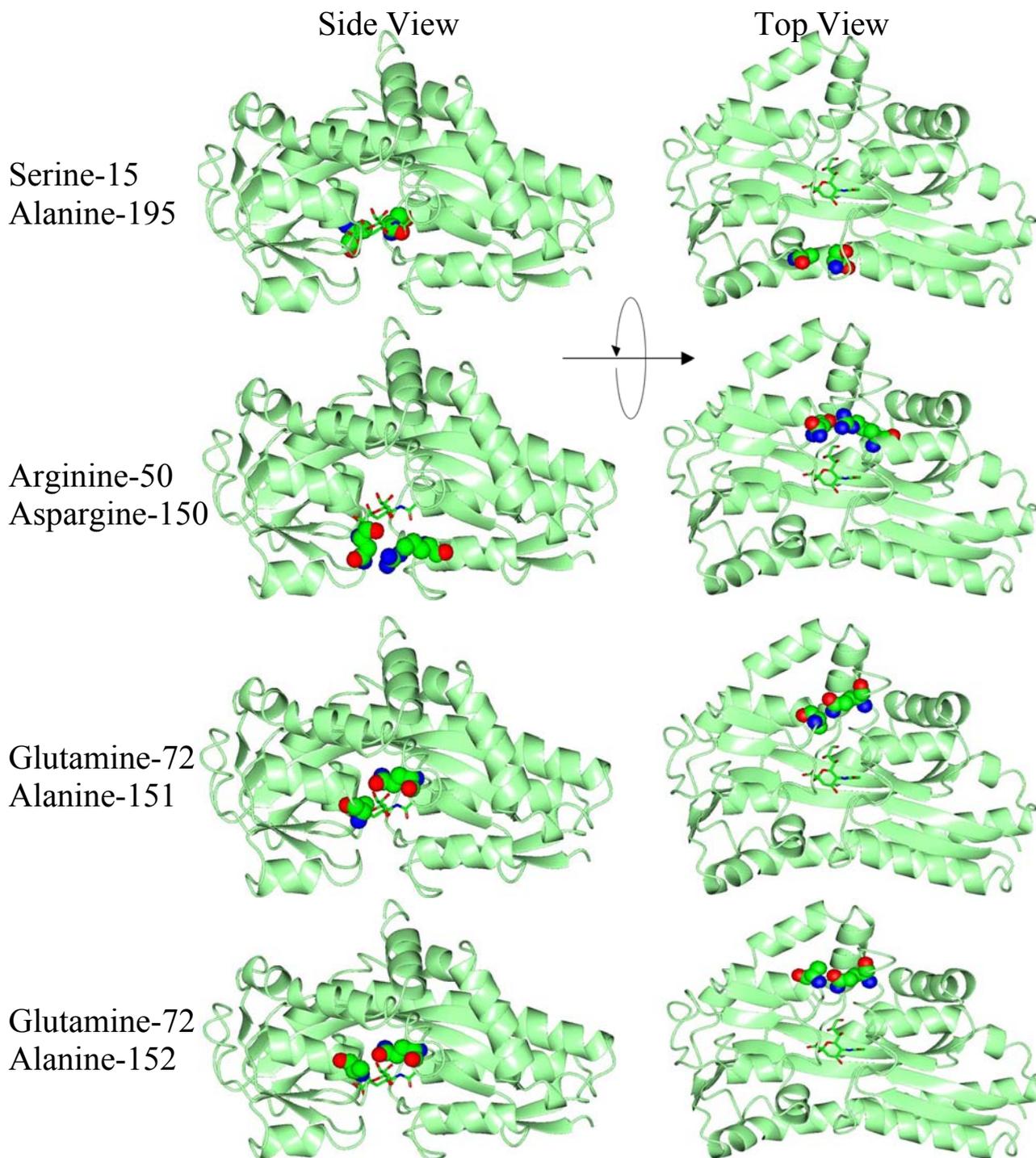
#### **4.3.2 Mutations designed to promote the closed conformation**

Binding proteins such as ABC PBPs and TRAP SBPs are believed to exist in equilibrium between the open and closed conformations, where the presence of ligand in the binding site stabilizes the closed conformation. There have been examples where destabilization of the open conformation has been used to increase the affinity of the binding protein for its ligand (Marvin & Hellinga, 2001, Telmer & Shilton, 2003). A different method of approach would be to stabilize the closed conformation, which has previously been accomplished by introducing disulphide bridges between the two domains of the protein

(Jacobson *et al.*, 1991, Stockner *et al.*, 2005, Zhang *et al.*, 1996). However, stabilizing the closed form has not been attempted by introducing opposite charges to form transient stabilizing interactions in the closed conformation. Four pairs of residues in SiaP were found to come close enough together when ligand bound to use as targets for this pair-wise approach (Figure 4.12 and Table 4.4). The first pair, Serine-15 and Alanine-195, is located near the hinge and are surface exposed in the open and closed conformations. It was decided to mutate both of these to both aspartate and lysine, which would also allow the control mutations of aspartate-aspartate and lysine-lysine. The second pair was Arginine-50 and Asparagine-150, which come very close together in ligand-bound SiaP and would only, theoretically, require the mutation of the asparagine to aspartate to form a salt bridge. The third and fourth pairs both share Glutamine-72, which comes close enough to Alanine-151 and Alanine-152, which could be made into a glutamate-lysine pair.

These combinations of double mutations were introduced into the *siaP-His<sub>6</sub>* allele in the pAH16 construct using a mutagenic megaprimer approach based on that of Kirsch & Joly (1998). This uses a forward primer containing one mutation and the reverse primer containing the second, both with a silent restriction site for screening, to make a PCR product of several hundred bases (Kirsch & Joly, 1998). This was purified and used as a megaprimer to introduce the double mutation into the *siaP* gene in pAH16. The mutant proteins were then expressed and purified following the same protocol as for SiaP-His<sub>6</sub>. The affinities of these mutant proteins for sialic acid were determined by titration of the fluorescence signal of 0.05  $\mu$ M protein, or 0.25  $\mu$ M protein when the  $K_d$  value was above 3.4  $\mu$ M sialic acid. All of these mutant proteins were expressed and purified in a similar manner to SiaP-His<sub>6</sub> and all gave clear titration of their fluorescence signal.

As can be seen in Table 4.4, all but one of these mutations caused a decrease in the sialic acid binding affinity. This single mutation, Asn-150-Asp, has no apparent effect on the ligand affinity, which is surprising, given the sensitivity of Ala-151 to mutation (above). The sensitivity of this location to mutation is highlighted in the comparison of Q72E;A151K and Q72E;A152K pairs, where switching from Ala-152-Lys to Ala-151-



**Figure 4.12:** Representation of the targets for the pair-wise mutants, designed to introduce Hydrogen bonds between the two domains of SiaP in the closed conformation. Sialic acid is shown as cylinders and the targets as spheres; carbon atoms are represented by green, oxygen by red and nitrogen by blue.

**Table 4.4:** Sialic acid-binding affinities of the pair-wise mutant proteins.

Protein	Intended interacting pair	$K_d$ ( $\mu\text{M}$ )
SiaP-His <sub>6</sub>	-	$0.11 \pm 0.02$
SiaP-His <sub>6</sub> :S15D;A195D	Asp-Asp	$15 \pm 7$
SiaP-His <sub>6</sub> :S15D;A195K	Asp-Lys	$76 \pm 4$
SiaP-His <sub>6</sub> :S15K;A195D	Lys-Asp	$9.3 \pm 1.9$
SiaP-His <sub>6</sub> :S15K;A195K	Lys-Lys	$38 \pm 9$
SiaP-His <sub>6</sub> :Q72E;A151K	Glu-Lys	$580 \pm 190$
SiaP-His <sub>6</sub> :Q72E;A152K	Glu-Lys	$3.4 \pm 1.3$
SiaP-His <sub>6</sub> :N150D	Arg-Asp	$0.10 \pm 0.02$

Lys causes a decrease in affinity of about 170-fold. The Ser-15;Ala-195 pairs have a strong detrimental effect, ranging from 85- to 690-fold decrease in affinity for S15K;A195D and S15D;A195K, respectively. Between these two extremes lie the control mutants, S15D;A195D and S15K;A195K, which cause 136- and 345-fold decreases. As can be seen, none of these mutations had the desired effect. This suggests that the interface between the two domains is very sensitive to the disruption caused by the introduction of such large, charged groups.

#### 4.4 Summary

The critical finding of this chapter, in relation to Authetix, is the finding that SiaP is specific for monomeric sialic acid (Section 4.2.3). This makes SiaP unsuitable for a LFD and fundamentally changes the design for a sialic acid detector. An attempt was made to construct a SiaP-based genetic nanosensor for sialic acid, but without success.

The buffer conditions for ligand binding were investigated, which resulted in an optimal set of conditions that gave SiaP a  $K_d$  value for Neu5Ac-binding of  $22 \pm 2$  nM (Figure 4.5). The thermodynamics of ligand binding were investigated using ITC and showed that ligand binding is enthalpically driven and associated with the release of many water molecules. This also suggested that the binding event is not a simple bi-molecular association as previously found, but could include multiple steps, as for DctP (Severi *et al.*, 2005, Walmsley *et al.*, 1992).

Rationally designed mutations were made in attempts to increase the ligand affinity of SiaP. Mutations in the binding pocket *all* had a deleterious effect, suggesting that the ligand binding site is a complex environment and sensitive to mutational change (Figure 4.11). A second group of mutations was developed, which aimed to introduce opposite charges on the binding cleft and promote the closed conformation of the protein (Figure 4.12). However, none of these caused an increase in affinity (Table 4.4).

## Chapter Five

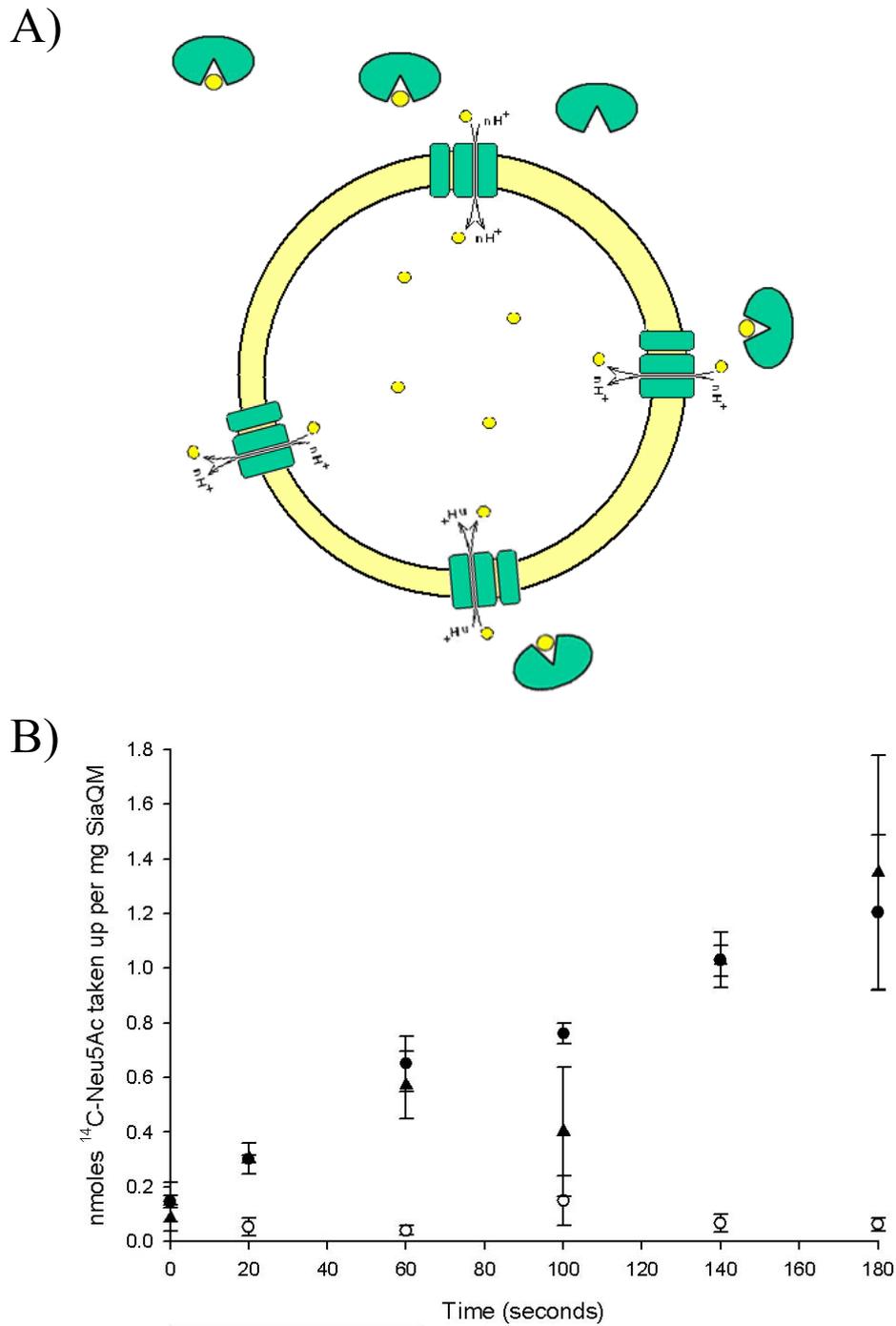
Investigation of the effect of mutations in  
SiaP upon transport by reconstituted  
SiaPQM

### 5.1 An *in vitro* assay for $^{14}\text{C}$ -Neu5Ac uptake by SiaP-SiaQM

Originally, the effects of mutations in *siaP* on Neu5Ac transport were to be investigated via their effect on growth in the *in vivo* assay mentioned previously. However, due to the growth seen for the extreme Arginine-147 mutations and the unforeseen complexity of cloning the *siaP* alleles into the final vector, an *in vitro* assay for SiaP activity was used (Mulligan *et al.*, 2009). This uses SiaP to supply  $^{14}\text{C}$ -labelled sialic acid ( $^{14}\text{C}$ -Neu5Ac) to SiaQM reconstituted into proteoliposomes (Figure 5.1a). The  $^{14}\text{C}$ -Neu5Ac transported into the lumen of the proteoliposomes can then be measured using scintillation counting. Originally, this assay used N-terminal decahistidine-tagged SiaP (His<sub>10</sub>SiaP) and so this had to be compared to SiaP-His<sub>6</sub> (Mulligan *et al.*, 2009).

Following the method of Mulligan *et al.* (2009), N-terminal decahistidine-tagged SiaQM was purified by Ni-affinity chromatography and reconstituted into proteoliposomes by rapid dilution. A homogenous suspension of 400 nm proteoliposomes containing the inside buffer was added to the reaction (outside) buffer, setting up an inwardly directed sodium gradient to energise the transporter. As the requirements for reconstituted SiaQM increased, this protein was kindly supplied by Judith Hawkhead. For the uptake assay, 5  $\mu\text{M}$  SiaP was equilibrated in the reaction buffer with 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac. SiaQM in proteoliposomes was added to a final concentration of 1.15  $\mu\text{M}$  and samples were taken 20 s later and every 40 s following this. Uptake of labelled sialic acid was halted by the addition of 1 mM unlabelled sialic acid and the transported  $^{14}\text{C}$ -Neu5Ac was measured by scintillation counting.

The assay was performed in triplicate for His<sub>10</sub>SiaP, SiaP-His<sub>6</sub> and the negative control, which used SiaP-His<sub>6</sub> and empty liposomes. These, not SiaQM-proteoliposomes, were chosen as the negative control due to the cost of producing SiaQM and that non-specifically associated SiaP had previously been shown to be the largest source of contaminating  $^{14}\text{C}$ -Neu5Ac (Christopher Mulligan, personal communication). As can be seen in Figure 5.1b, the uptake of  $^{14}\text{C}$ -Neu5Ac proceeds in a similar manner for both His<sub>10</sub>SiaP and SiaP-His<sub>6</sub> and both are clearly above the background level.



**Figure 5.1:** *In vitro*  $^{14}\text{C}$ -Neu5Ac uptake assay. **A)**  $^{14}\text{C}$ -Neu5Ac is bound by SiaP, delivered to SiaQM-containing proteoliposomes and transported into the lumen of the proteoliposomes. **B)**  $^{14}\text{C}$ -Neu5Ac uptake catalysed by SiaP-His<sub>6</sub> (black circles) and His<sub>10</sub>-SiaP (black triangles). The background intensity is shown as empty circles.

These results indicate that His<sub>10</sub>SiaP and SiaP-His<sub>6</sub> are equivalent in this assay and so the C-terminal hexahistidine-tag can be used in all future experiments. It is worth noting that the amounts of <sup>14</sup>C-Neu5Ac recovered are significantly lower than the highest intensities reported previously (Mulligan *et al.*, 2009), which is likely due to variations in the reconstitution efficiency of SiaQM.

## **5.2 The Arg-147 mutants catalyse no uptake *in vitro***

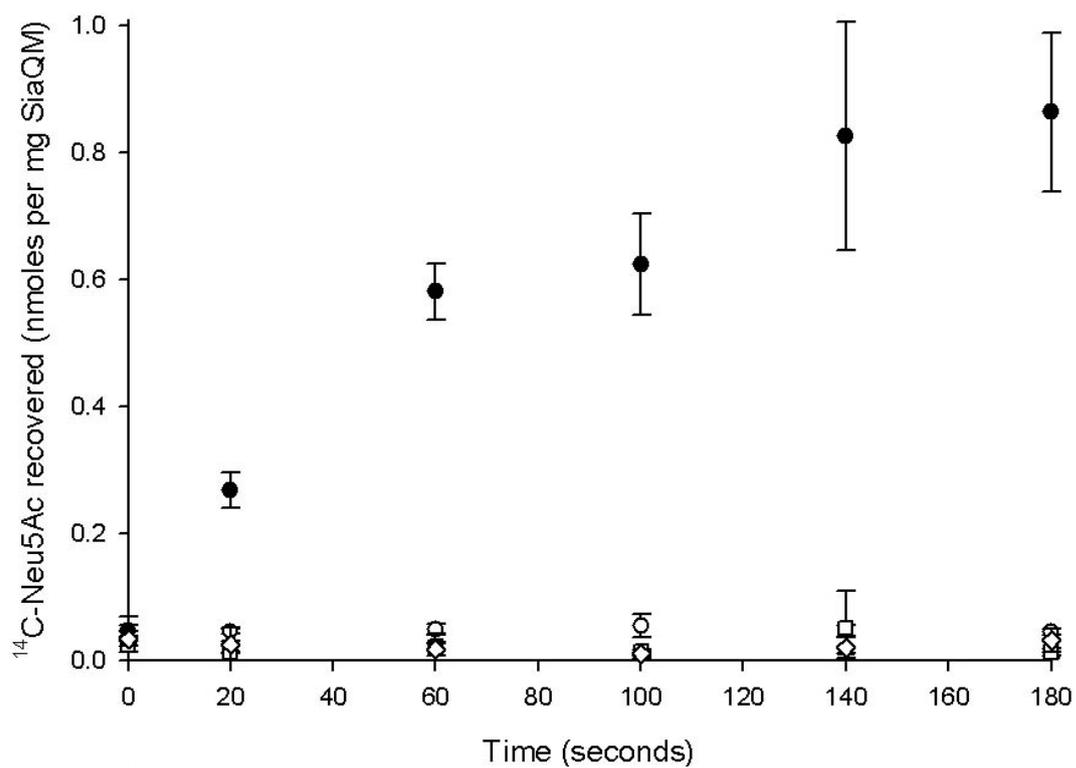
Previously, the Arginine-147 mutants of SiaP-His<sub>6</sub> had exhibited no detectable binding with up to 3 mM Neu5Ac (Figure 3.6). However, these same mutants are able to complement growth of the *E. coli*  $\Delta nanT$  strain when supplied *in trans* in the presence of *siaQM* (Figure 3.17). It was decided to investigate the transport activity of these mutant proteins in the reconstituted *in vitro* system.

The <sup>14</sup>C-Neu5Ac uptake assay was performed with 5  $\mu$ M of each of the SiaP-His<sub>6</sub>:R147 mutants and 5  $\mu$ M <sup>14</sup>C-Neu5Ac with samples taken every forty seconds with SiaP-His<sub>6</sub> as the positive control. As can be seen in Figure 5.2, the transport rates of the Arginine-147 mutants of SiaP-His<sub>6</sub> are indistinguishable from background.

This agrees with the *in vitro* sialic acid binding data that these mutants do not interact with sialic acid in the low  $\mu$ M range. It also suggests that the higher substrate concentrations and longer time periods of the *in vivo* growth assay are needed to detect transport by these mutants and that the Arg-147 is important for high affinity sialic acid transport but is not essential.

## **5.3 Position 170 is important for ligand binding and transport**

The *in vitro* investigation of the role of the Phenylalanine-170 ‘lid’ in ligand binding showed that this residue had to be an aromatic residue for high affinity sialic acid binding (Figure 3.7). The role in transport, rather than just ligand binding, could be investigated easily using the <sup>14</sup>C-Neu5Ac uptake assay.



**Figure 5.2:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by SiaP-His<sub>6</sub> (filled circles) and the SiaP-His<sub>6</sub>:R147 mutants (Ala, triangles; Glu, squares; Lys, diamonds). The background intensity is shown as empty circles.

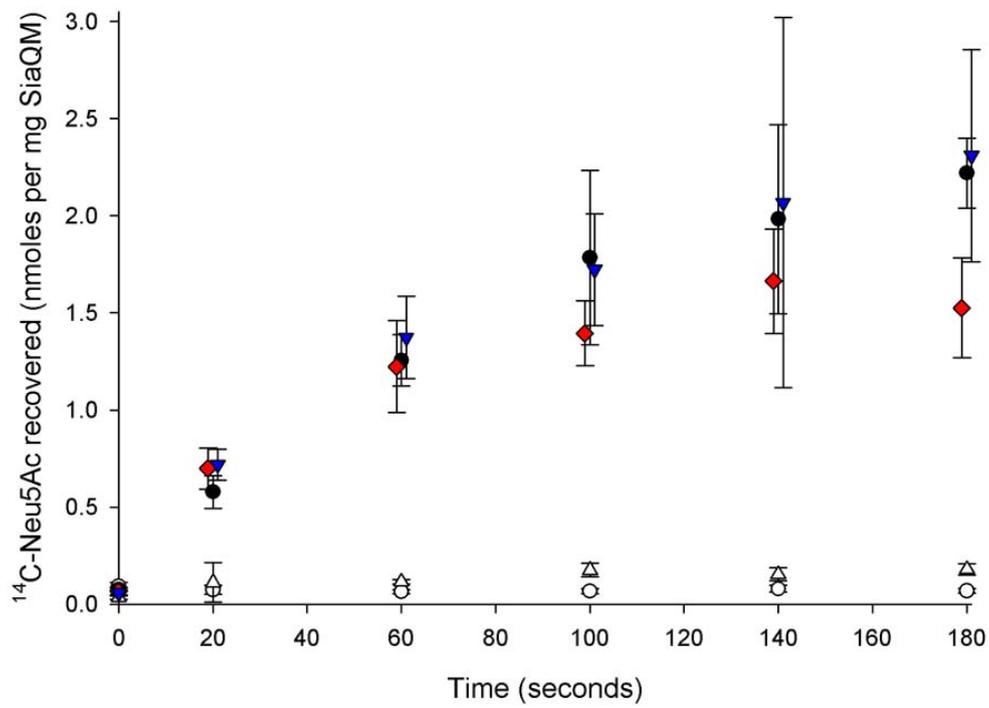
The transport of  $^{14}\text{C}$ -Neu5Ac catalysed by each of the SiaP-His<sub>6</sub>:F170 mutants was investigated using the same conditions described previously. Figure 5.3 shows the amounts of  $^{14}\text{C}$ -Neu5Ac recovered at each time point. As expected, SiaP-His<sub>6</sub>:F170A shows no uptake of substrate, since its  $K_d$  for sialic acid ( $340 \pm 70 \mu\text{M}$ ) is 68 times the available concentration, while the tyrosine and tryptophan mutants catalyse uptake at similar rates as SiaP-His<sub>6</sub>.

The working concentration of  $^{14}\text{C}$ -Neu5Ac used is five times the  $K_d$  values for the tyrosine and tryptophan mutants ( $1.13 \pm 0.05 \mu\text{M}$  and  $1.21 \pm 0.03 \mu\text{M}$ , respectively) and so these are expected to be ligand-bound. This is supported by the similarity between the uptake rates of SiaP-His<sub>6</sub> and the tyrosine mutant. The tryptophan mutant appears to catalyse transport at a similar initial rate, but this fails to accumulate to the same level.

The absence of transport by the low affinity-binding F170A mutant is not unexpected, since it is not saturated by the Neu5Ac used in the assay. Combined with the transport-positive phenotypes of the F170W and F170Y mutants indicate that the 170 position of SiaP has to have an aromatic group for binding and transport of sialic acid, but must be phenylalanine for highest affinity binding. The lower recovery of labelled Neu5Ac from the SiaP-His<sub>6</sub>:F170W-catalysed assay indicates that the larger aromatic side chain of this mutant has disturbed gating or regulation of the transport cycle in such a way that it is no longer able to accumulate such a high concentration gradient into the proteoliposome.

#### **5.4 Binding site mutations in SiaP reduce transport *in vitro* via their reduced occupancy**

The binding site mutants developed to introduce more side chain interactions with the ligand in Section 4.1.1 (Figure 4.1) all showed decreased affinity for sialic acid. Since all of these mutations are inside the core of the protein, they should not alter the interaction between SiaP and SiaQM. Due to this, these mutations could be used to investigate the effect of reduced sialic acid affinity on transport.



**Figure 5.3:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by SiaP-His<sub>6</sub> (black circles), SiaP-His<sub>6</sub>:F170A (empty triangles), SiaP-His<sub>6</sub>:F170W (red diamonds) and SiaP-His<sub>6</sub>:F170Y (blue down triangles). The background intensity is shown as empty circles.

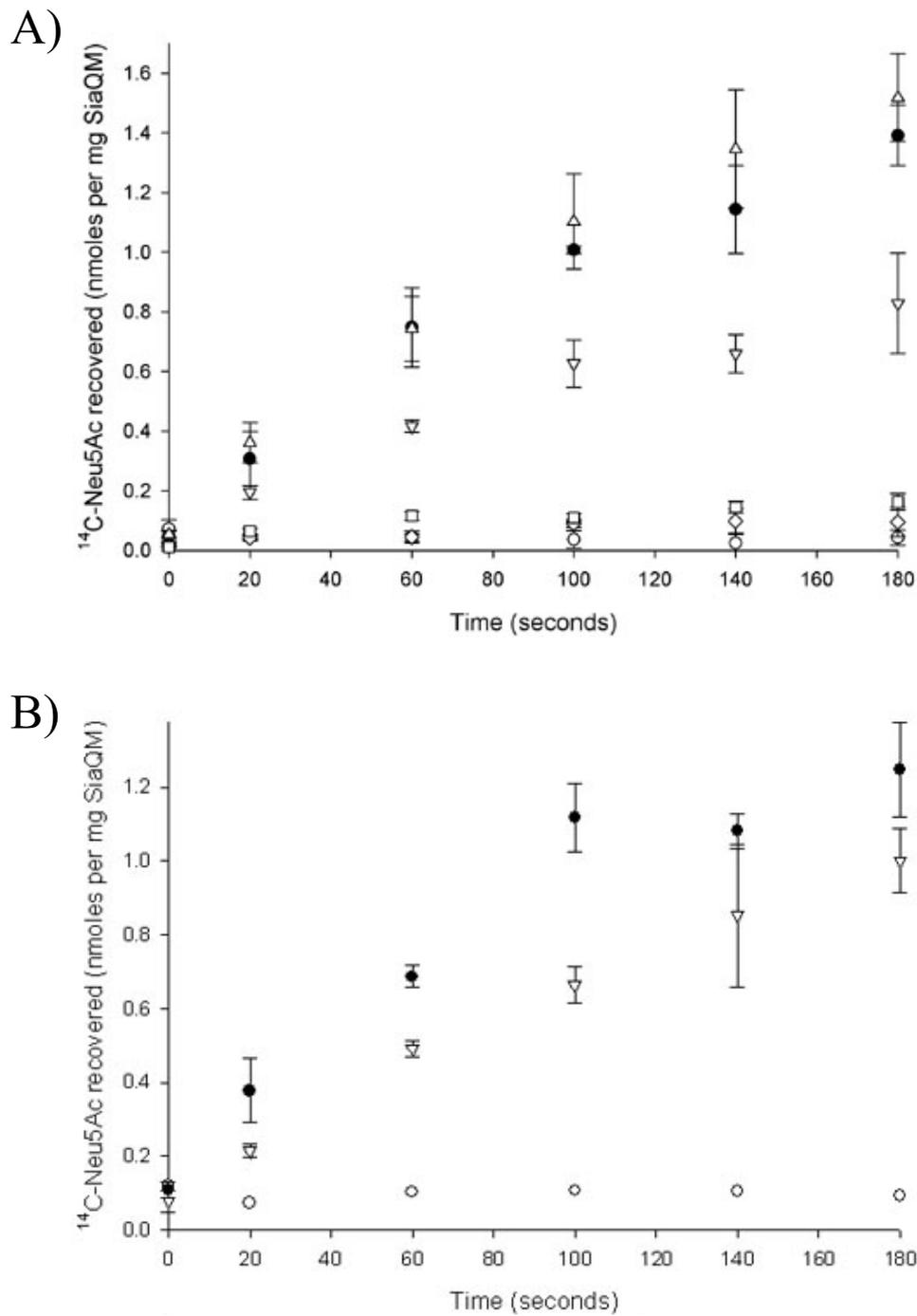
For these mutant proteins of varying affinity for sialic acid, the concentration of protein and  $^{14}\text{C}$ -Neu5Ac in each assay was kept at 5  $\mu\text{M}$  each. Uptake catalysed by SiaP-His<sub>6</sub>:A11N and SiaP-His<sub>6</sub>:A151V are at a similar level to the background, while that of SiaP-His<sub>6</sub>:A66I is similar to SiaP-His<sub>6</sub> (Figure 5.4a). Interestingly, transport with SiaP-His<sub>6</sub>:A66M is reduced compared to SiaP-His<sub>6</sub>.

The lack of detectable transport by the A11N and A151V mutants is unsurprising, given that their  $K_d$  values for sialic acid are 19 and 4.5 times higher than the concentration of  $^{14}\text{C}$ -Neu5Ac used in the assay. The transport by the A66I mutant is indistinguishable from SiaP-His<sub>6</sub>, indicating that the binding proteins are ligand-bound to a similar extent. The reduced uptake of the A66M mutant could be explained by the reduced occupancy of this mutant protein; the concentration of  $^{14}\text{C}$ -Neu5Ac used is only five times the  $K_d$  of the protein for sialic acid, compared to ten times the  $K_d$  of the Ala-66-Ile mutant.

To examine if the reduced uptake catalysed by the A66M mutant is due to reduced occupancy of the binding protein, the assay was repeated using a concentration of  $^{14}\text{C}$ -Neu5Ac ten times the  $K_d$  of the protein for sialic acid to reach a similar level of ligand-bound protein to the Ala-66-Ile mutant.

In this experiment, the transport of  $^{14}\text{C}$ -Neu5Ac by 5  $\mu\text{M}$  SiaP-His<sub>6</sub> or SiaP-His<sub>6</sub>:A66M was measured in the presence of 10  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac. The uptake of  $^{14}\text{C}$ -Neu5Ac by the Ala-66-Met under these conditions is much closer to that of SiaP-His<sub>6</sub>, but not identical (Figure 5.4b).

Doubling the concentration of  $^{14}\text{C}$ -labelled sialic acid brought the concentration to nine times the  $K_d$  of the mutant protein for sialic acid. This increase has brought the transport rate closer to that of SiaP-His<sub>6</sub> under the same conditions, indicating that the reduced transport rate seen previously was due to reduced occupancy of the binding protein under the same substrate-limiting conditions.



**Figure 5.4:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by SiaP-His<sub>6</sub> (black circles), SiaP-His<sub>6</sub>:A66I (up triangles), SiaP-His<sub>6</sub>:A66M (down triangles), SiaP-His<sub>6</sub>:A151V (diamonds) and SiaP-His<sub>6</sub>:A11N (squares). **A)** Uptake from 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac reaction concentration. The background intensity is shown as empty circles. **B)** Uptake from 10  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac reaction concentration. The background intensity is an average of two results and is shown as empty circles.

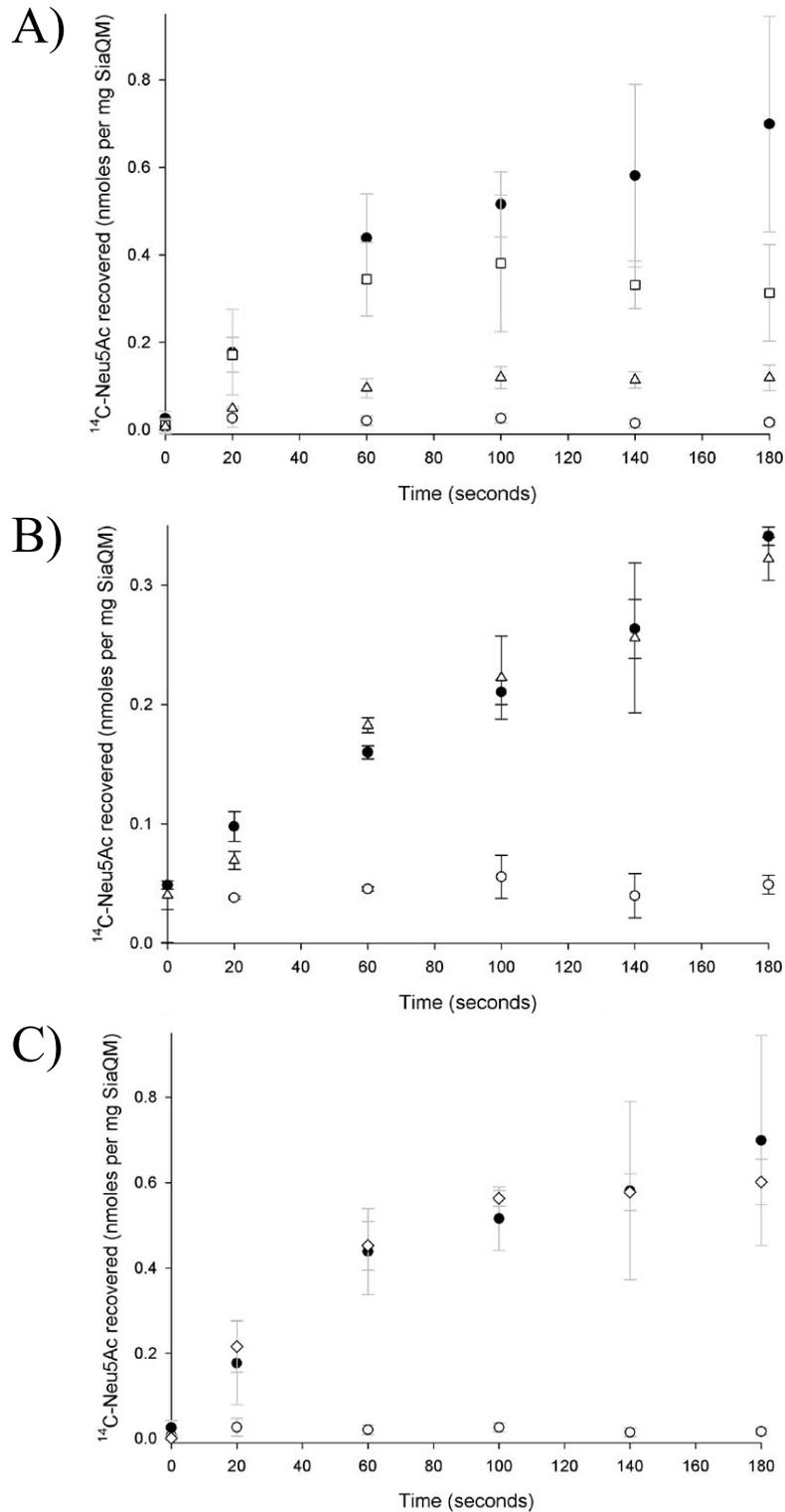
## 5.5 Surface mutations in SiaP affect transport by different mechanisms

In Section 4.1.2, double mutants of SiaP were developed where opposite charges were introduced that would come into proximity only in the closed form and so stabilise this conformation of SiaP (Figure 4.4). These double mutants introduced charged groups onto the surface of SiaP and so could be used to investigate the interaction between the binding protein and SiaQM. From the previous section, any effects of reduced ligand affinity could be compensated for by increased  $^{14}\text{C}$ -Neu5Ac concentration.

For the investigation of SiaP-His<sub>6</sub>:Q72E;A152K, SiaP-His<sub>6</sub>:S15D;A195K and SiaP-His<sub>6</sub>:S15K;A195D, the transport activity of 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac by 5  $\mu\text{M}$  of each protein was determined using the standard protocol. To investigate the effect of the lower affinity of SiaP-His<sub>6</sub>:Q72E;A152K, the uptake catalysed by 5  $\mu\text{M}$  of this protein was measured in the presence of 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac and 25  $\mu\text{M}$  Neu5Ac, giving a sialic acid concentration of 30  $\mu\text{M}$ .

The transport of  $^{14}\text{C}$ -Neu5Ac catalysed by the S15D;A195K double mutant is very low, but above background, while that of SiaP-His<sub>6</sub>:Q72E;A152K is much higher (Figure 5.5a). The reduced transport of the Q72E;A152K and S15D;A195K double mutants is not unexpected due to their reduced affinity for sialic acid (Figure 5.5a). To examine if this reduced rate is solely reduced to reduced occupancy of the binding proteins under these conditions, the concentration of sialic acid was raised to 30  $\mu\text{M}$ , at which the uptakes catalysed by the double mutant and SiaP-His<sub>6</sub> are identical (Figure 5.5b). This indicates that that the reduced uptake seen in the previous experiment was due to reduced occupancy of the binding protein.

Surprisingly, the uptake catalysed by the S15K;A195D double mutant was identical to that of SiaP-His<sub>6</sub> (Figure 5.5c). This is surprising since the  $K_d$  value of SiaP-His<sub>6</sub>:S15K;A195D for sialic acid is  $9.3 \pm 1.9 \mu\text{M}$ , which is twice the concentration of  $^{14}\text{C}$ -Neu5Ac used in this assay. It is not clear how this double mutant, which has a 78-fold decrease in ligand affinity, maintains wild type transport when the protein should not be saturated with ligand. It could be that there is a stronger interaction



**Figure 5.5:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by SiaP-His<sub>6</sub> with double mutations on the surface. The background intensity is shown as empty circles. **A)**  $^{14}\text{C}$ -Neu5Ac uptake catalysed by SiaP-His<sub>6</sub> (black circles), SiaP-His<sub>6</sub>:Q72E;A152K (squares) and SiaP-His<sub>6</sub>:S15D;A195K (triangles). **B)** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac from 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac mixed with 25  $\mu\text{M}$  Neu5Ac catalysed by SiaP-His<sub>6</sub> (black circles) and SiaP-His<sub>6</sub>:Q72E;A152K (triangles). **C)** The transport of  $^{14}\text{C}$ -Neu5Ac from 5  $\mu\text{M}$  reaction concentration catalysed by SiaP-His<sub>6</sub> (black circles) and SiaP-His<sub>6</sub>:S15K;A195D (diamonds).

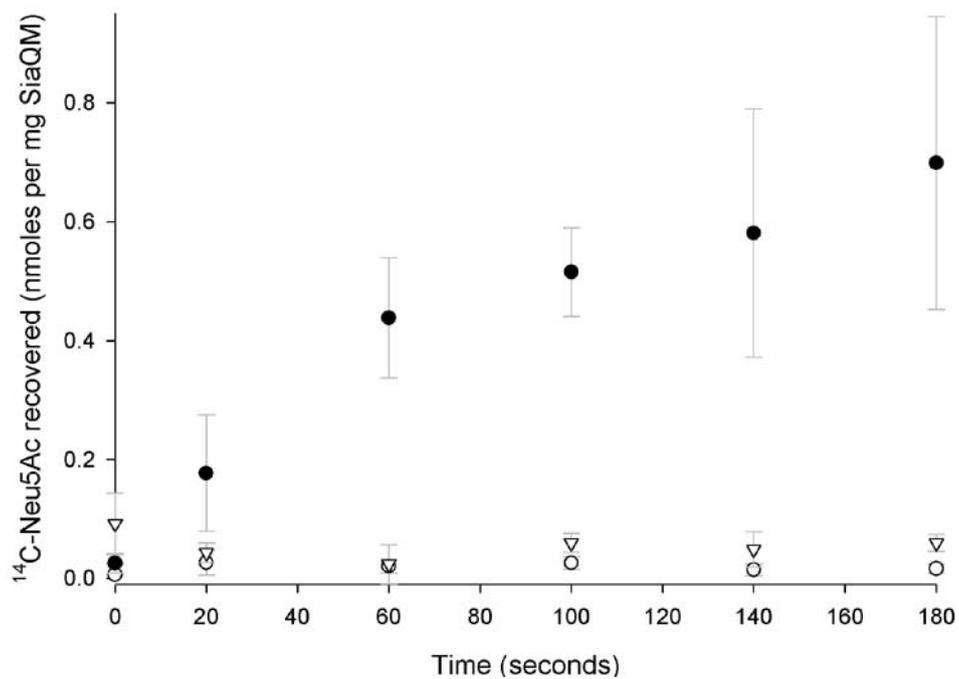
between SiaP-His<sub>6</sub>:S15K;A195D and SiaQM, thereby increasing the apparent concentration locally to the membrane protein. It could also be that the closed conformation has been disrupted in the mutant, increasing the probability of the ligand-bound protein returning to the open conformation. When in the transport complex, this conformationally less stable mutant could result in the substrate being delivered more quickly or successfully with each cycle of transport.

### **5.6 The Asn-150-Asp mutation abolishes transport without affecting ligand binding**

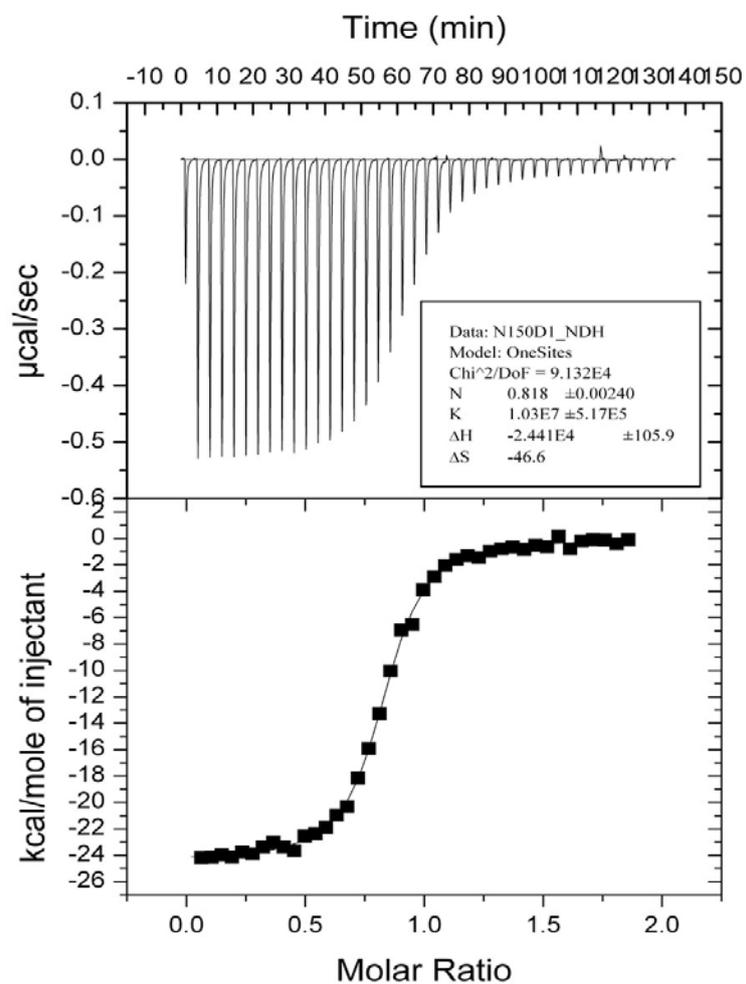
Along with the double mutants from Section 4.1.2, a point mutation was also made that was intended to have the same effect as the double mutants (Figure 4.4). The affinity of this mutant, SiaP-His<sub>6</sub>:N150D, for Neu5Ac was determined *in vitro* and found to have had no effect, with the  $K_d$  value remaining at  $0.10 \pm 0.02 \mu\text{M}$ . This was then investigated using the *in vitro* uptake assay with the standard protocol.

In Figure 5.7, the recovery of <sup>14</sup>C-Neu5Ac from the SiaP-His<sub>6</sub>:N150D assay is indistinguishable above background, indicating that the transport *in vitro* was abolished by this mutation. This lack of transport could be due to either an inactive protein preparation or an incorrect measurement of the affinity by fluorescence titration.

As an independent measure of the  $K_d$ , ITC was also performed in triplicate on 10  $\mu\text{M}$  protein in 50 mM Tris/HCl pH 8.0 at 37 °C with additions of 150  $\mu\text{M}$  sialic acid (Figure 5.7). These show that this mutation has had no effect on the affinity for sialic acid ( $K_d$   $0.10 \pm 0.02 \mu\text{M}$ ), which is surprising given the sensitivity of SiaP to all of the mutations made during the course of this investigation and the extreme effect of mutation to Ala-151 (Table 4.7).



**Figure 5.6:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by SiaP-His<sub>6</sub> (black circles) and SiaP-His<sub>6</sub>:N150D (down triangles). The background intensity is shown as empty circles.



**Figure 5.7:** ITC analysis of 10  $\mu\text{M}$  SiaP-His<sub>6</sub>:N150D in 50 mM Tris/HCl pH 8.0 at 37 °C. 150  $\mu\text{M}$  Neu5Ac was injected in 6  $\mu\text{l}$  aliquots following an initial injection of 2  $\mu\text{l}$ . In this titration,  $K = 1.03 \times 10^7 \text{ M}^{-1}$  and so the  $K_d$  value is 0.10  $\mu\text{M}$ .

It was decided to examine the effect of the Asn-150-Asp mutation on the transporter *in vivo* using the same system that was used to discover the functionality of SiaP-His<sub>6</sub> in Section 3.3 and the Arginine-147 mutants in Section 3.6. This is based on complementing the deletion of the *E. coli* native sialic acid transporter with *siaPQM* on a low copy number plasmid.

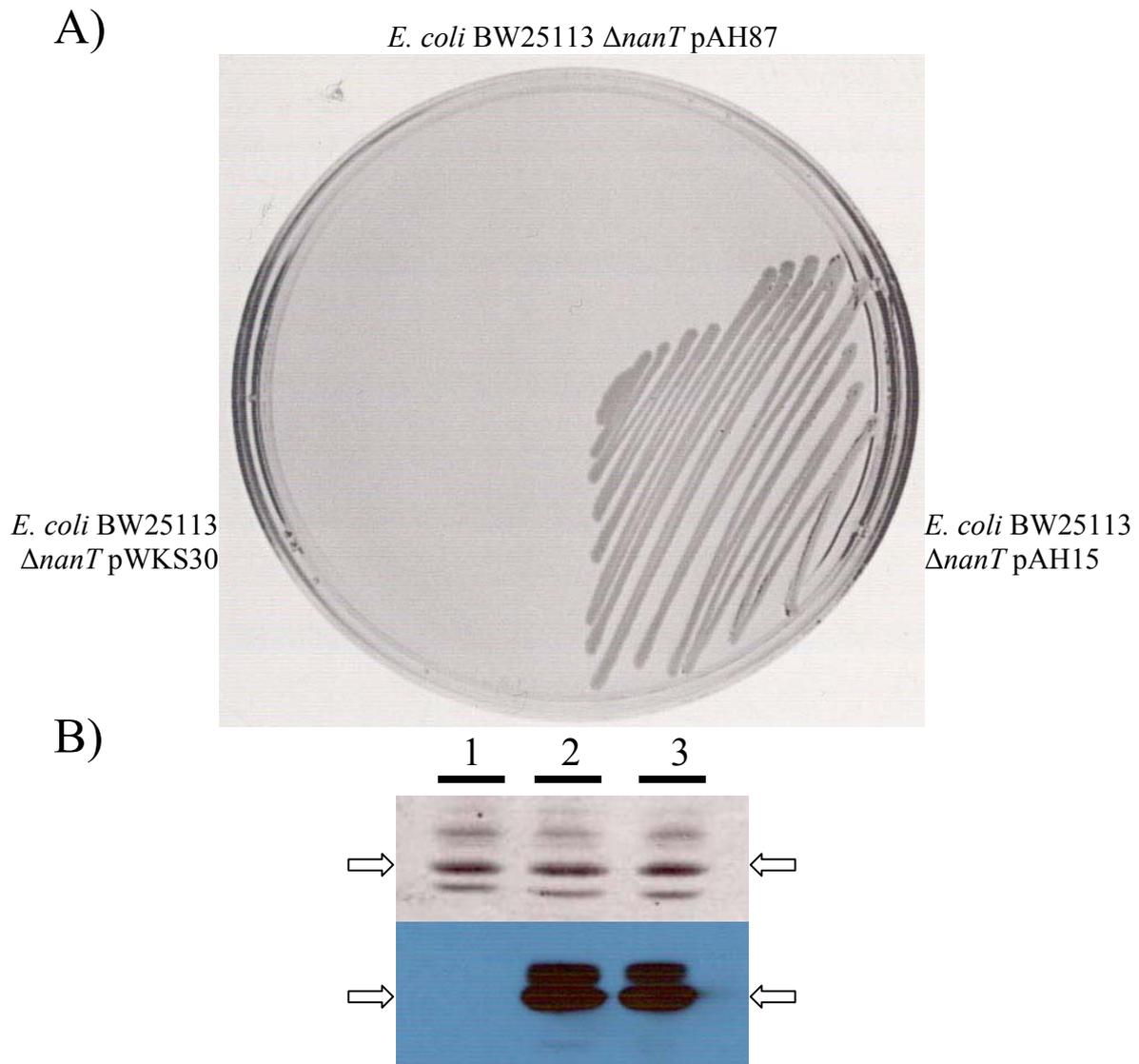
The Asn-150-Asp mutation was introduced directly into the *siaP-His<sub>6</sub>* allele in pAH15 (pWKS30-*siaP-His<sub>6</sub>-siaQM*), creating pAH87. This was then transformed into *E. coli* BW25113  $\Delta nanT$  and tested for growth on sialic acid as the sole carbon source with induction of the operon by IPTG. Expression of a hexahistidine-tagged SiaP was probed by Western blot analysis.

Introduction of this mutation to the whole transporter did not restore growth on sialic acid as the sole carbon source (Figure 5.8). This phenotype is far more extreme than that caused by mutation of the conserved Arg-147.

The failure of growth at this high sialic acid concentration and long incubation time indicate that this mutation has abolished either the binding protein–membrane protein interaction or ligand release from the SBP as ligand binding is unaltered. This *in vivo* null-growth phenotype was also seen in so-called dominant-negative MalE, where a mutated residue near the binding cleft sticks outwards in the closed conformation (Shilton *et al.*, 1996).

#### **5.6.1 SiaP-His<sub>6</sub>: N150D forms a transient, non-productive complex with SiaQM**

To investigate the possible interaction between SiaP-His<sub>6</sub>:N150D and SiaQM, a competition assay between SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D was set up. If an interaction occurred, then the presence of the transport-incompetent mutant protein would reduce transport of <sup>14</sup>C-Neu5Ac by the transport-competent protein.



**Figure 5.8:** A) Growth of *E. coli* BW25113  $\Delta nanT$  strains expressing *siaP-His<sub>6</sub>-siaQM* from pAH15, *siaP-His<sub>6</sub>:N150D-siaQM* from pAH87 or carrying the empty vector, pWKS30, on M9 minimal medium 1% agarose, 1 mM IPTG supplemented with 1 mg/ml sialic acid as the sole carbon source. B) **Top**, SDS PAGE gel of *E. coli* BW25113  $\Delta nanT$  strains expressing *siaPQM* from (1) pES7, (2) *siaP-His<sub>6</sub>-siaQM* from pAH15 and (3) *siaP-His<sub>6</sub>:N150D-siaQM* from pAH87 (3). **Bottom**, western blot with anti(tetrahistidine) antibodies showing expression of polyhistidine-tagged SiaP in lanes 2 and 3 with a population of uncleaved SiaP visible as a higher weight band. The MW of native SiaP is indicated by arrows.

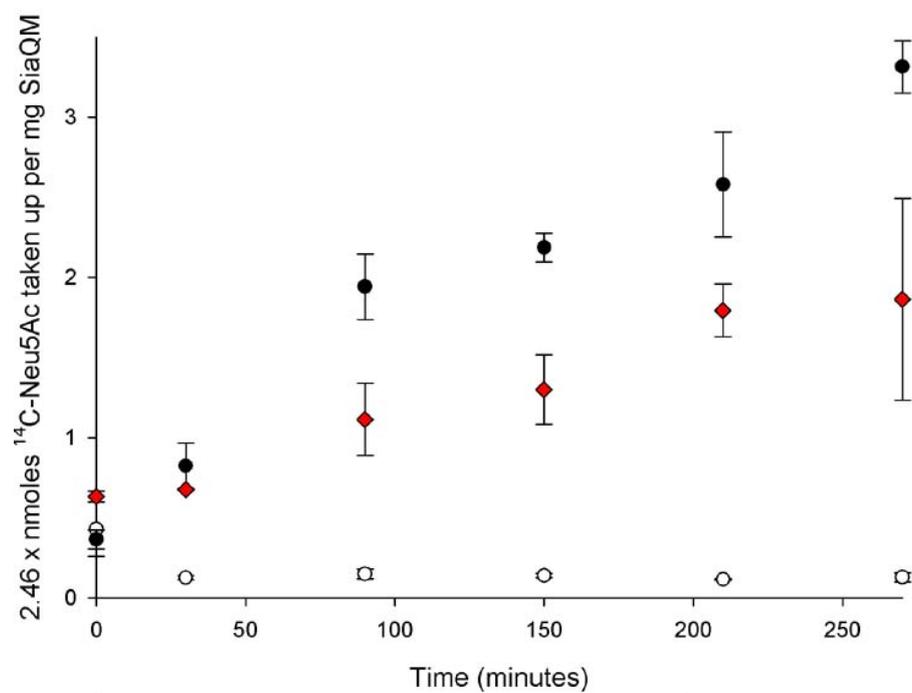
The assay was performed with 5  $\mu\text{M}$  SiaP-His<sub>6</sub> and again with a mix of 5  $\mu\text{M}$  SiaP-His<sub>6</sub> and 5  $\mu\text{M}$  SiaP-His<sub>6</sub>:N150D. To ensure that the binding protein would remain saturated throughout the experiment, the transport activity was measured in the presence of 20  $\mu\text{M}$  sialic acid, made up of 8.1  $\mu\text{M}$  <sup>14</sup>C-Neu5Ac and 11.9  $\mu\text{M}$  Neu5Ac.

As can be seen in Figure 5.9, the addition of the mutant protein to SiaP-His<sub>6</sub> at a ratio of 1:1 in the presence of an excess of sialic acid causes a reduction in the amount of <sup>14</sup>C-Neu5Ac recovered. This result shows that SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D are competing for the same available SiaQM. The association between the mutant protein and the membrane protein is non-productive and also transient, since transport still occurs via the activity of SiaP-His<sub>6</sub>.

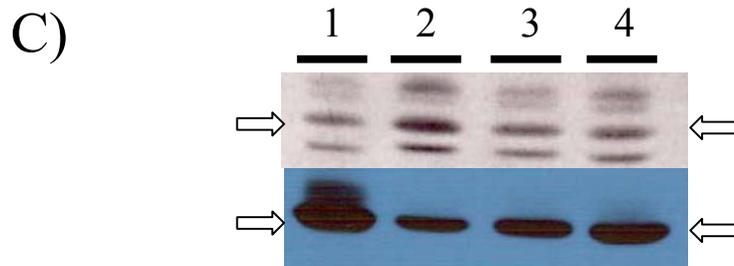
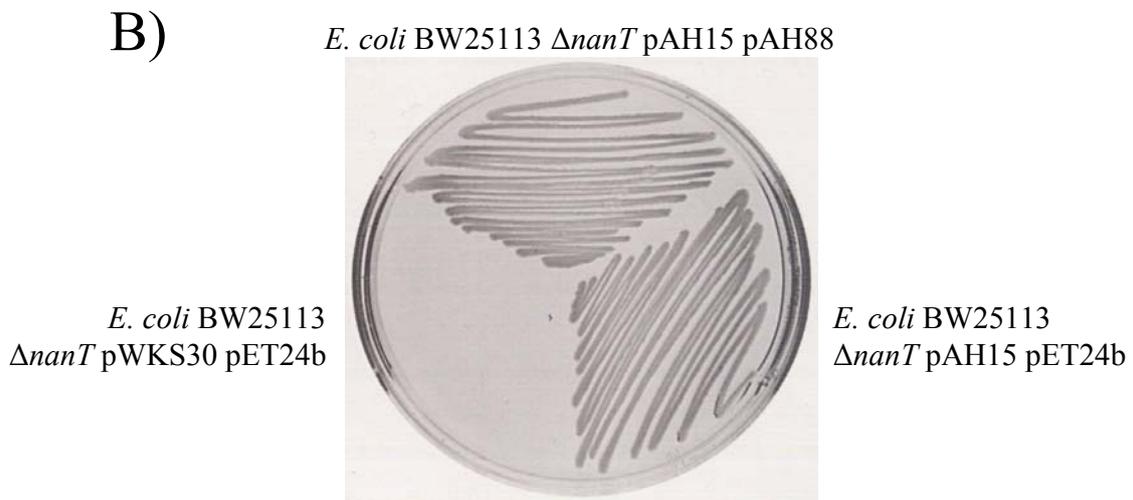
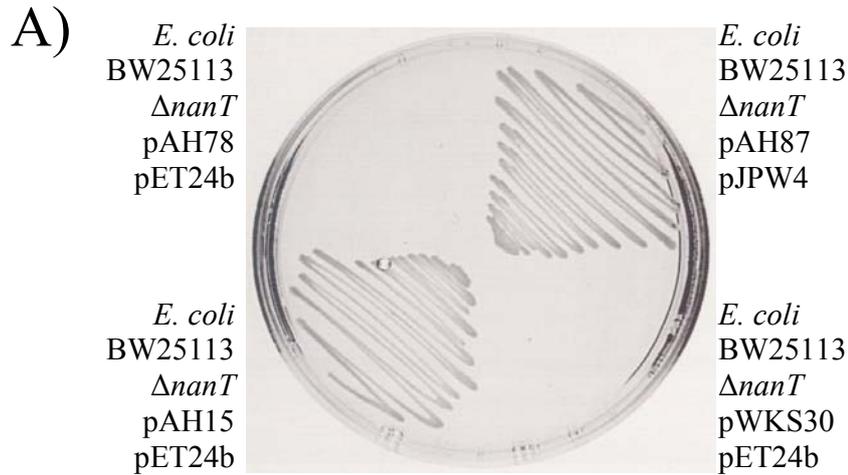
In the previous uptake assay, SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D were present at an equal concentration. To examine the interaction between SiaP-His<sub>6</sub>:N150D and SiaQM further, both SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D could individually be expressed to an excess over the other *in vivo*.

To monitor the effect on transport of supplying excess SiaP-His<sub>6</sub> to SiaP-His<sub>6</sub>:N150D–SiaQM, *E. coli* BW25113  $\Delta nanT$  was transformed with pAH78 (pWKS30-*siaP-His<sub>6</sub>:N150D-siaQM*) and pJPW4 (pET24b-*siaP-His<sub>6</sub>*). In contrast, the effect of excess SiaP-His<sub>6</sub>:N150D on the growth of SiaP-His<sub>6</sub>–SiaQM was investigated using *E. coli* BW25113  $\Delta nanT$  pAH15 (pWKS30-*siaP-His<sub>6</sub>-siaQM*) and pAH88 (pET24b-*siaP-His<sub>6</sub>:N150D*). These were tested for growth on sialic acid as the sole carbon source with induction of all proteins by IPTG. For all of these strains, expression of hexahistidine tagged-SiaP was probed by Western blotting.

The expression of *siaP-His<sub>6</sub>* in the strain carrying the  $\Delta nanT$  *siaP-His<sub>6</sub>:N150D*<sup>+</sup> *siaQM*<sup>+</sup> background restores growth on Neu5Ac (Figure 5.10ac). However, overexpression of *siaP-His<sub>6</sub>:N150D* in the  $\Delta nanT$  *siaP-His<sub>6</sub>*<sup>+</sup> *siaQM*<sup>+</sup> background does not abolish the growth of this strain on sialic acid (Figure 5.10bc). All of these double plasmid strains



**Figure 5.9:** The *in vitro* transport of  $^{14}\text{C}$ -Neu5Ac catalysed by  $5\ \mu\text{M}$  SiaP-His<sub>6</sub> (black circles) and a mix of  $5\ \mu\text{M}$  SiaP-His<sub>6</sub> and  $5\ \mu\text{M}$  SiaP-His<sub>6</sub>:N150D (red diamonds). The concentration of sialic acid used was  $8.1\ \mu\text{M}$   $^{14}\text{C}$ -Neu5Ac with  $11.9\ \mu\text{M}$  Neu5Ac to ensure an excess of ligand over protein. The background intensity is shown as empty circles.



**Figure 5.10:** Growth of *E. coli* BW25113  $\Delta nanT$  strains carrying low-copy number *siaPQM* mutant alleles and high-copy number *siaP* mutant alleles on M9 minimal medium 1% agarose, 1 mM IPTG supplemented with 1 mg/ml sialic acid as the sole carbon source. **A)** The effect of SiaP-His<sub>6</sub>, from pJPW4, on the growth phenotype of SiaP-His<sub>6</sub>:N150D-SiaQM from pAH87. **B)** The effect of overexpression of SiaP-His<sub>6</sub>:N150D, from pAH88, on the growth effect of SiaP-His<sub>6</sub>-SiaQM from pAH15. **C)** Expression of *E. coli* BW25113  $\Delta nanT$  strains carrying pET- and pWSK-based constructs: **1)** pAH15 (*siaP-His<sub>6</sub>-siaQM*) and pAH88 (*siaP-His<sub>6</sub>:N150D*); **2)** pAH15 (*siaP-His<sub>6</sub>-siaQM*) and pET24b (empty vector); **3)** pAH87 (*siaP-His<sub>6</sub>:N150D-siaQM*) and pET24b (empty vector); **4)** pAH87 (*siaP-His<sub>6</sub>:N150D-siaQM*) and pJPW4 (*siaP-His<sub>6</sub>*). Top, SDS PAGE gel. **Bottom**, western blot with anti(tetrahistidine) antibodies showing expression of polyhistidine-tagged SiaP in all lanes. The MW of native SiaP is indicated by arrows.

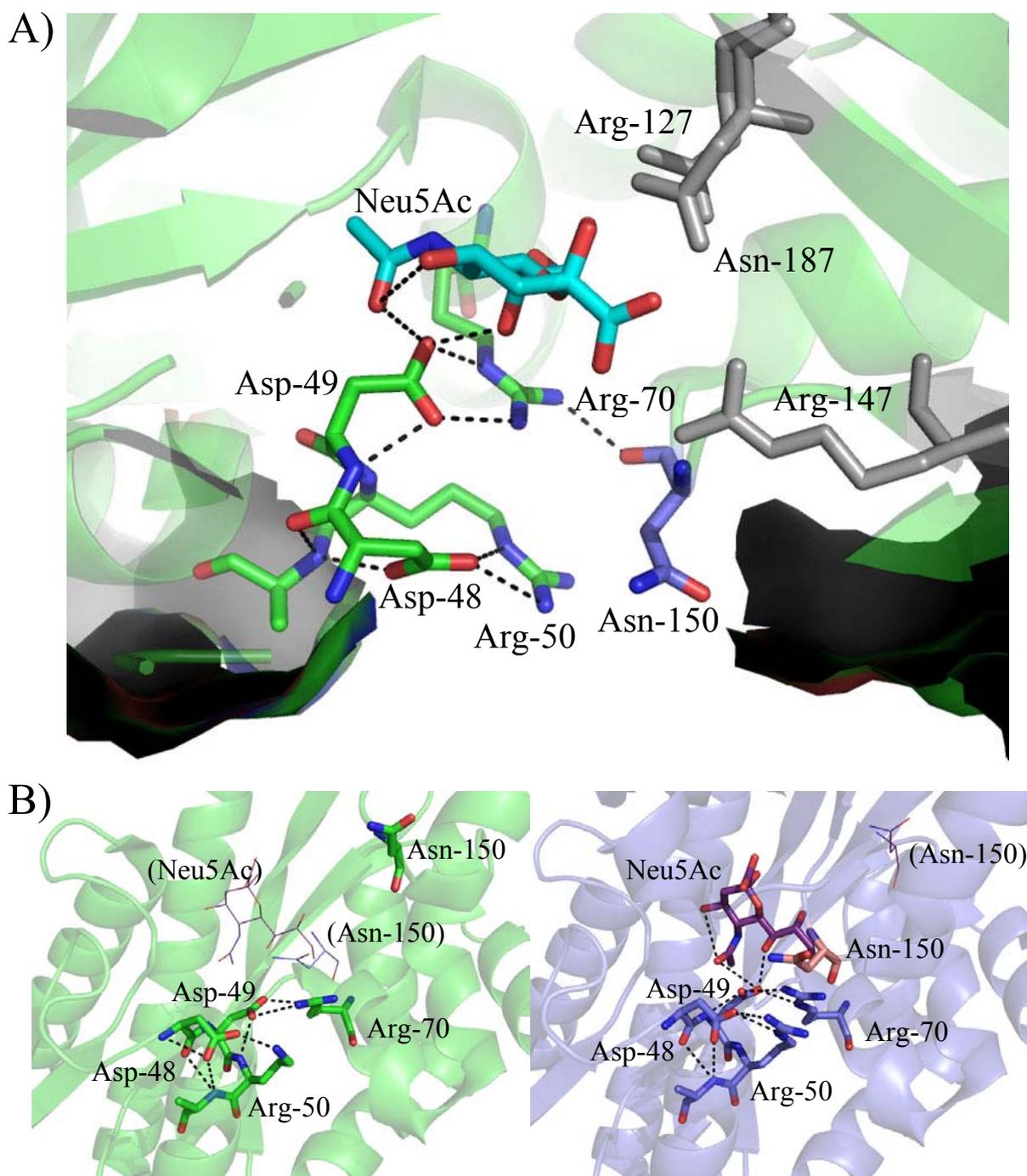
were tested for growth on M9 minimal medium agarose supplemented with 0.6% glucose with IPTG-induction and lack of growth without IPTG-induction on M9 minimal medium agarose supplemented with 1 mg/ml sialic acid.

The restoration of growth by expression of SiaP-His<sub>6</sub> in the strain expressing SiaP-His<sub>6</sub>:N150D–SiaQM indicates that SiaP-His<sub>6</sub>:N150D is not forming a permanent inactivating complex with SiaQM *in vivo* and that SiaP-His<sub>6</sub> can out-compete the non-productive mutant protein. However, with the supposed excess of SiaP-His<sub>6</sub>:N150D in the strain expressing *siaP-His<sub>6</sub>-siaQM*, growth is maintained. This could be due to the smaller population of transport-competent SiaP-His<sub>6</sub> maintaining a level of sialic acid uptake that allowed growth on the solid medium.

From these, it could be that the single atom change between asparagine and aspartate has disturbed the association between the binding protein and SiaQM or has interrupted the passage of the ligand in some way.

### **5.6.2 Asn-150 forms part of a network extending from the bound ligand to the surface of SiaP**

From the examination of the sialic acid-bound crystal structure of SiaP, it appears that  $\pi$ -stacking occurs between Asn-150 and Arg-50, which forms part of a hydrogen-bonded network with Asp-48 and Arg-70. This apparent hydrogen-bonded network extends from the surface of the protein to the bound sialic acid, via Asp-49 (Figure 5.11a). Arg-50 forms hydrogen bonds from an N $\eta$  and its N $\epsilon$  with one terminal oxygen from the carboxylate of Asp-48; N $\eta$  and its N $\epsilon$  of Arg-70 are coordinated by the terminal oxygens from the ligand-interacting Asp-49, while the remaining terminal amine of Arg-70 forms a hydrogen bond with the backbone carbonyl of Asn-150. Of these, only Asn-150 is part of Domain 2, which contains the ligand carboxylate-coordinating residues including Arg-147, and is only interacting with the remainder of the network via Asn–Arg stacking and one hydrogen bond.



**Figure 5.11:** Representation of the residues involved in the Arg-50 – Asn-150 network. **A)** The hydrogen bonded network extending from the bound Neu5Ac (cyan carbon atoms) to the surface of SiaP (partial surface representation). The residues involved in the network from domain 1 are shown with green carbon atom, while Asn-150 from domain 2 is shown with purple carbon atoms. Arg-127, Arg-147 and Asn-187 are shown in grey as a reference and a ribbon representation of SiaP is shown in green. **B)** The changes in the network upon ligand binding. In the open conformation, the relative positions of Neu5Ac and Asn-150 (marked in brackets) from the closed conformation are shown as lines. In the closed conformation, the relative position Asn-150 (marked in brackets) from the closed conformation is also shown as lines.

As can be seen in Fig 5.11b, this network undergoes a slight conformational change between the open and closed forms. In the open form of SiaP, the diaminomethyl group of Arg-50 is not represented in the crystal structure, due to disorder, and one terminal amine of Arg-70 interacts with the terminal oxygens of Asp-49. From this, it would seem that these residues could be acting as a network that takes part in the opening of SiaP and release of sialic acid on interaction with SiaQM.

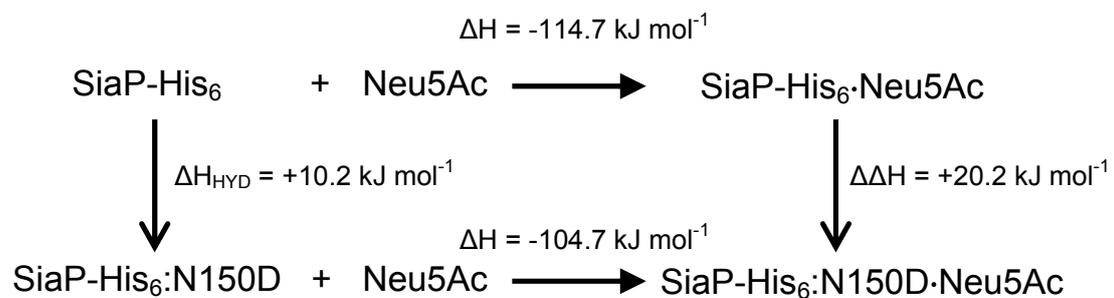
The ability of SiaP-His<sub>6</sub>:N150D to bind sialic acid with the same affinity and in apparently the same manner as SiaP-His<sub>6</sub> would suggest that this residue is not critical in domain closure. Since the mutant can interact with SiaQM, it would also seem that this mutant protein would be conformationally equivalent to SiaP.

The thermodynamic data from ITC of SiaP-His<sub>6</sub>:N150D and SiaP-His<sub>6</sub> are shown in Table 5.1. There is no significant effect of this mutant on  $K_d$  or  $\Delta G$ ; however,  $\Delta H$  and  $\Delta S$  both increase in the presence of this mutation by 9.6 kJ mol<sup>-1</sup> and 32.6 J mol<sup>-1</sup> K<sup>-1</sup>, respectively.

It is surprising that the Asn-150-Asp mutation appears to abolish transport of sialic acid without affecting the affinity of ligand binding. To confirm this surprising result, the uptake assay and ITC were repeated twice each with independent preparations of protein and found to give the same results. In addition to the  $\Delta H$  contributions that are identical between these two proteins, the change in  $\Delta H$  ( $\Delta\Delta H$ ) is made up of the hydration enthalpy difference ( $\Delta H_{\text{hyd}}$ ) between aspartate and asparagine plus the new, or lost, interaction(s). By considering the thermodynamic cycle, the resultant enthalpy change ( $\Delta\Delta H$ ) can be calculated, revealing the gain or loss of interactions caused by the mutation (Figure 5.12). Under these circumstances, it is possible to estimate the change in  $\Delta H_{\text{hyd}}$  ( $\Delta\Delta H_{\text{hyd}}$ ), assuming that the  $\Delta H_{\text{hyd}}$  of asparagine and aspartate in these proteins are equivalent to that of asparagine and aspartate chemical analogues in solution. This is not

**Table 5.1:** Thermodynamic data from the ITC analysis of SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D.

Buffer composition	$K_d$ ( $\mu\text{M}$ )	$\Delta H$ ( $\text{J mol}^{-1}$ )	$\Delta S$ ( $\text{J mol}^{-1} \text{K}^{-1}$ )	$-T\Delta S$ ( $\text{J mol}^{-1}$ )	$\Delta H - T\Delta S$ ( $\text{J mol}^{-1}$ )
SiaP-His <sub>6</sub>	$0.12 \pm 0.02$	$-114700 \pm 4400$	$-237.0 \pm 12.8$	$73500 \pm 4000$	$-41270 \pm 500$
SiaP-His <sub>6</sub> :N150D	$0.10 \pm 0.02$	$-104700 \pm 4300$	$-203.3 \pm 15.3$	$63000 \pm 4700$	$-41700 \pm 400$



**Figure 5.12:** The change in  $\Delta H$  ( $\Delta\Delta H$ ) between Neu5Ac binding by SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D. The enthalpies of binding and hydration are shown.  $\Delta\Delta H = -(-114.7) + 10.2 + (-104.7) = +20.2 \text{ kJ mol}^{-1}$ .

an unacceptable assumption, given that Asn-150 is on the surface of the protein and exposed to the bulk solvent. To do this, the terminal component of asparagine was taken to be ethanamide (acetamide) and for aspartate, this was taken to be ethanoic (acetic) acid. The  $\Delta H_{\text{hyd}}$  of acetamide is  $-40.59 \text{ kJ mol}^{-1}$  and the  $\Delta H_{\text{hyd}}$  of acetic acid is calculated as  $-30.43 \text{ kJ mol}^{-1}$ , giving the  $\Delta\Delta H_{\text{hyd}}$  of these two compounds as  $+10.2 \text{ kJ mol}^{-1}$  (Kang *et al.*, 1988). Adding these to the cycle shown in Figure 5.12, this gives the  $\Delta\Delta H$  of SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D binding to sialic acid to be equal to  $20.2 \pm 4.4 \text{ kJ mol}^{-1}$ . This value is equal to that of the new interactions formed in the closed mutant protein or lost by open conformation. Since the Asn-150 side chain only interacts with Arg-50 via  $\pi$ -stacking, it is more likely that this increase in  $\Delta H$  is due to Asp-150 having more interactions. This could be due to the formation of hydrogen bonds with other residues in the protein or in the bulk solvent. This could be demonstrated clearly in the crystal structure of this mutant protein.

It is also worth noting that the change in  $\Delta S$  of  $+33.7 \text{ J K}^{-1} \text{ mol}^{-1}$  (a change in  $T\Delta S$  of  $10.5 \text{ kJ mol}^{-1}$ ) between these two mutants indicates that the N150D mutant protein is either more disordered in the closed conformation or that the open form is more ordered.

The protein crystal structure of SiaP-His<sub>6</sub>:N150D was determined in the presence of sialic acid by Marcus Fischer (YSBL). For all residues except for Asp-150, the structures appear identical. Asp-150 can be observed in an identical position to Asn-150 in SiaP; however, it also appears that Asp-150 adopts a second conformation in approximately 60% of cases. This second conformation is oriented away from the network, projecting away from the cleft and into the bulk solvent.

With this data, it would seem that the N150D mutation does not prevent the interaction between the two proteins, but does prevent transport, probably by failing to open the domains and release the substrate. How this failure of transport might be caused by this

mutation could be explained by its incorrect interactions with the rest of the network and mis-positioning for the interaction with SiaQM.

From these results, it is clear that this position is of critical importance to the transport of sialic acid by SiaPQM, despite its lack of conservation in the DctP family (11%). The mutation of this residue appears to cause a failure of signalling between SiaP and SiaQM. Whether it is a signal of ligand-occupancy from SiaP to SiaQM or an opening/release instruction from SiaQM to SiaP is yet to be determined.

## 5.6 Summary

This chapter has focussed on the *in vitro* transport capabilities of the mutant library using the reconstituted system. The first to be investigated were the Arg-147 mutants, which catalysed no uptake of  $^{14}\text{C}$ -Neu5Ac (Figure 5.2). The Phe-170 mutants showed that this position needed to be an aromatic group for transport to be detectable and a phenylalanine was required for high affinity binding (Figure 5.3).

Mutations in the binding pocket of SiaP reduced the uptake of the labelled sialic acid, but this was demonstrated to be due to their reduced affinity for the ligand (Figure 5.4). In one of the surface-exposed double mutants, Q72E;A152K, this was also demonstrated to be the cause of its reduced transport (Figure 5.5ab). The S15K;A195D double mutant, despite its low affinity for sialic acid (9  $\mu\text{M}$ ), catalysed normal uptake from 5  $\mu\text{M}$   $^{14}\text{C}$ -Neu5Ac, which could be due to a stronger interaction with SiaQM or an altered equilibrium between the closed and open conformations (Figure 5.5c).

The most surprising finding from this chapter is the importance of the Asn-150 position to transport. Mutation of this asparagine to aspartate causes a complete loss of transport both *in vitro* (Figure 5.6) and *in vivo* (Figure 5.8), but appears to allow a transient, non-productive interaction with SiaQM (Figures 5.9 and 5.10). In the protein crystal structure of SiaP, Asn-150 seems to form a part of a hydrogen-bonded network that extends from the surface of the protein to the bound sialic acid (Figure 5.11) and so could be involved in signalling between SiaP and SiaQM.

## Chapter Six

### The utilisation of different sialic acid analogues

### 6.1 The ligand *N*-acetyl group is surrounded by a water network

As stated in Section 1.6, Neu5Ac is a member of the sialic acid family, which is a group of  $\alpha$ -keto acidic sugars (Figure 1.14). SiaP has previously been shown to bind Neu5Ac, Neu5Gc and KDN with  $K_d$  values in the  $\mu\text{M}$  range (Muller *et al.*, 2006). The most structurally similar analogue, glycolyl neuraminic acid (Neu5Gc), has an extra hydroxyl on the *N*-acetyl group and is bound by SiaP with a very similar  $K_d$  of 0.29  $\mu\text{M}$ . Also of interest is the de-aminated analogue, 3-deoxy-D-glycero-D-galacto-nonulosonic acid (KDN), which is bound with a  $K_d$  of 42  $\mu\text{M}$ . The differences between these three sialic acids are restricted to the *N*-acetyl group on C<sub>5</sub> and so could be investigated using the existing mutant library developed in Section 4.3.1. Of these, Ala-11-Asn, Ala-66-Ile and Ala-66-Met are changes in the region around the ligand *N*-acetyl group (Figure 4.11).

The affinities of these mutant proteins for different ligands were determined by titration of the fluorescence signal of 0.05  $\mu\text{M}$  protein, or 0.25  $\mu\text{M}$  protein when the  $K_d$  value was above 1.1  $\mu\text{M}$  (Table 6.1). The mutations cause a decrease in affinity, of which the Ala-11-Asn mutation is the most detrimental.

The effects of these mutations on the  $K_d$  values for the different sialic acids are shown in Table 6.2. As can be seen, both of the Ala-66 mutations cause a similar decrease in affinity for Neu5Ac and KDN. Of most interest is the effect of the Ala-11-Asn mutation. This causes an 873-fold decrease in affinity for Neu5Ac and a 198-fold decrease for Neu5Gc, but only a 12-fold decrease in the affinity for KDN. In the crystal structure of Neu5Ac-bound SiaP, Ala-11 is next to a hydrogen-bonded water network around the ligand *N*-acetyl group. And so, the introduction of a large, charged side chain such as asparagine would be very disturbing to ligands with the *N*-acetyl group present. When KDN is bound, this region would be far more accepting of the asparagine side chain, as shown by its decreased sensitivity to this mutation.

**Table 6.1:** The binding affinities of sialic acid analogues by SiaP-His<sub>6</sub>, SiaP-His<sub>6</sub>:A11N, SiaP-His<sub>6</sub>:A66I and SiaP-His<sub>6</sub>:A66M. Values in italics were determined for SiaP (Muller *et al.*, 2006).

Protein	<i>K<sub>d</sub></i> values		
	Neu5Ac	KDN	Neu5Gc
SiaP-His <sub>6</sub>	0.11 ± 0.02 μM	<i>42 μM</i>	<i>0.29 μM</i>
SiaP-His <sub>6</sub> :A11N	96 ± 2 μM	497 ± 6 μM	57.5 ± 3.2 μM
SiaP-His <sub>6</sub> :A66I	0.53 ± 0.11 μM	148 ± 3 μM	-
SiaP-His <sub>6</sub> :A66M	1.10 ± 0.04 μM	154 ± 34 μM	-

**Table 6.2:** The effect of the binding site mutations on the binding affinity of Neu5Ac, Neu5Gc and KDN compared to the wild-type SiaP-His<sub>6</sub>.

Protein	Decrease in affinity for ligand caused by the mutation		
	Neu5Ac	KDN	Neu5Gc
SiaP-His <sub>6</sub> :A11N	873 x	12 x	198 x
SiaP-His <sub>6</sub> :A66I	5 x	4 x	-
SiaP-His <sub>6</sub> :A66M	10 x	4 x	-

During the investigation of the physical properties of SiaP-His<sub>6</sub> in Section 4.1.4, the thermodynamics of ligand binding were determined experimentally at different temperatures (Figure 4.4) and used in Section 5.5.3 to investigate the effects of the N150D mutation on the ligand binding event (Figure 5.12). The components of this, such as enthalpy, entropy and heat capacity, could be determined to reveal the mechanics of ligand binding. The change in heat capacity ( $\Delta C_p$ ) of Neu5Ac binding by SiaP-His<sub>6</sub> was calculated from these results and revealed the  $\Delta C_p$  to be  $-2.74 \pm 0.19 \text{ kJ mol}^{-1} \text{ K}^{-1}$ , which would be negative due to large numbers of water molecules being displaced upon ligand binding. A similar analysis could be performed on KDN binding by SiaP-His<sub>6</sub> to reveal the extent of water displacement on ligand binding.

ITC was performed with various concentrations of SiaP-His<sub>6</sub> using ten times the concentration of KDN in triplicate at 15 °C and in duplicate at 30 °C. These constraints were due to the large amounts of protein needed for each run; 2.5 ml of protein sample had to be prepared at 110  $\mu\text{M}$  protein for runs at 15 °C and at 220  $\mu\text{M}$  protein for runs at 30 °C. Ligand-bound protein was recycled by guanidine denaturation.

For such low affinity binding, ITC required large amounts of protein. This was reduced by using guanidine hydrochloride to denature and refold the protein, releasing the bound ligand. The resultant, unliganded protein appeared to behave normally, that is, the thermodynamic data produced were similar. The ITC gave clear, repeatable results at both temperature points. The values for  $\Delta H$ ,  $\Delta S$  and  $\Delta G$  are shown in Table 6.3.  $\Delta C_p$  is given by the gradient of the graph of  $\Delta H$  against temperature.

Statistical thermodynamics shows that  $\Delta C_p$  is the fluctuation of enthalpy, and so is a direct measurement of the flexibility of the system (Cooper, 2005). An increase in this value ( $\Delta\Delta C_p$ ) represents an increase in flexibility, or fluctuation, which could be explained by the addition of water to the system. In the change from Neu5Ac to KDN binding this value is equal to  $+230 \text{ J mol}^{-1} \text{ K}^{-1}$ , which is made up of the additional water plus the difference in heat capacity between the two ligands. As in Section 5.5.3, this is also the case for both  $\Delta\Delta H$  and  $\Delta\Delta S$ .

**Table 6.3:** The thermodynamic data from the ITC analysis of Neu5Ac and KDN binding by SiaP-His<sub>6</sub>.

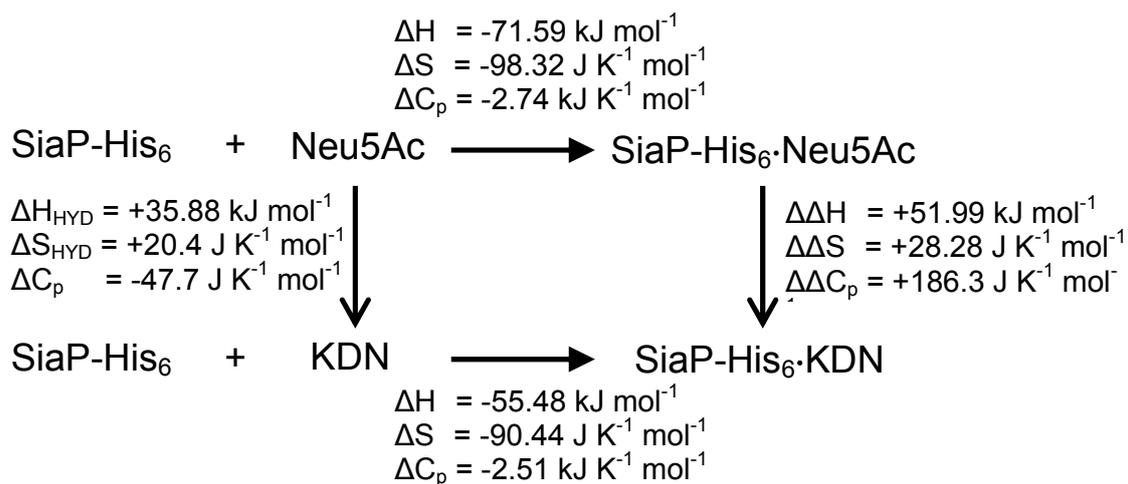
Ligand	Temp (°C)	K <sub>d</sub> (μM)	ΔH (kJ mol <sup>-1</sup> )	ΔS (J mol <sup>-1</sup> K <sup>-1</sup> )	-TΔS (kJ mol <sup>-1</sup> )	ΔH-TΔS (J mol <sup>-1</sup> )	ΔC <sub>p</sub> (kJ K <sup>-1</sup> mol <sup>-1</sup> )
Neu5Ac	30	0.058 ± 0.005	-84.4 ± 3.0	-139.9 ± 10.1	42.4 ± 3.1	-42.0 ± 3.0	-2.74 ± 0.19
	15	0.027 ± 0.003	-49.0 ± 1.6	-24.9 ± 5.7	7.2 ± 1.6	-41.8 ± 1.6	
KDN	30	6.36 ± 0.74	-30.4 ± 0.4	-6.0 ± 2.1	1.7 ± 0.6	-28.7 ± 0.6	-2.51 ± 0.03
	15	16.4	-68.0	-132.6	40.2	-27.8	

To determine the effective change in these thermodynamic values, the differences in the contributions of KDN and Neu5Ac were estimated using the values of methanol ( $\text{CH}_3\text{-OH}$ ) for KDN and *N*-methylacetamide ( $\text{CH}_3\text{-NHCOCH}_3$ ) for Neu5Ac (Della Gatta *et al.*, 1986, Graziano, 2001, Ooi *et al.*, 1987). The  $\Delta H_{\text{hyd}}$ ,  $\Delta S_{\text{hyd}}$  and  $\Delta C_p$  of methanol and *N*-methylacetamide and the resulting changes in  $\Delta H$ ,  $\Delta S$  and  $\Delta C_p$  are shown in Figure 6.1.

Adding these to the cycle shown in Figure 6.1, this gives the  $\Delta\Delta H$  from Neu5Ac binding to KDN binding to be equal to  $+51.98 \text{ kJ mol}^{-1}$ . This larger enthalpy indicates weaker binding. The change in  $\Delta C_p$  of  $+186.3 \text{ J mol}^{-1} \text{ K}^{-1}$  shows an increase in flexibility, which could be due to the addition of extra water, which is line with the increase in the entropy of the system ( $+60.5 \text{ J mol}^{-1} \text{ K}^{-1}$ ). All of these data suggest the presence of extra water involved in the binding of KDN compared to Neu5Ac.

It was previously shown that the A11N mutation had a less deleterious effect on KDN binding than Neu5Ac binding (Table 6.2). Unfortunately, it would be impractical to perform ITC on the binding on KDN by SiaP-His<sub>6</sub>:A11N due to the extreme protein needs for such a low affinity interaction (the expected  $K_d$  value at  $15 \text{ }^\circ\text{C}$  is  $\sim 115 \text{ } \mu\text{M}$ ). However, it was possible to perform ITC on Neu5Ac binding by the mutant, despite the excessive protein and ligand requirements.

ITC was performed on  $220$  or  $420 \text{ } \mu\text{M}$  SiaP-His<sub>6</sub>:A11N at  $15 \text{ }^\circ\text{C}$  and  $30 \text{ }^\circ\text{C}$ , respectively, with a ligand concentration equal to ten times the protein concentration. The protein concentration was chosen to maintain the lowest practical *c*-value ( $\sim 10$ ) for the titration curve. The extreme protein requirement restricted the number of repeats achievable to a duplicate dataset at both temperature points and so are reported merely as averages without standard deviations. The titrations gave clear, similar results and are shown in Table 6.4.



**Figure 6.1:** The thermodynamic cycles of Neu5Ac and KDN binding by SiaP-His<sub>6</sub>. The effect of the ligand modification is taken into account in the differences in the  $\Delta H$ ,  $\Delta S$  and  $\Delta C_p$  of both binding events.

For example,  $\Delta\Delta H(\text{SiaP-His}_6\cdot\text{Neu5Ac} \rightarrow \text{SiaP-His}_6\cdot\text{KDN})$

$$\begin{aligned}
 &= -\Delta H(\text{SiaP-His}_6\cdot\text{Neu5Ac}) + \Delta H_{\text{hyd}}(\text{Neu5Ac} \rightarrow \text{KDN}) + \Delta H(\text{SiaP-His}_6\cdot\text{KDN}) \\
 &= -(-71.59) + 35.88 + (-55.48) \\
 &= +51.99 \text{ kJ mol}^{-1}
 \end{aligned}$$

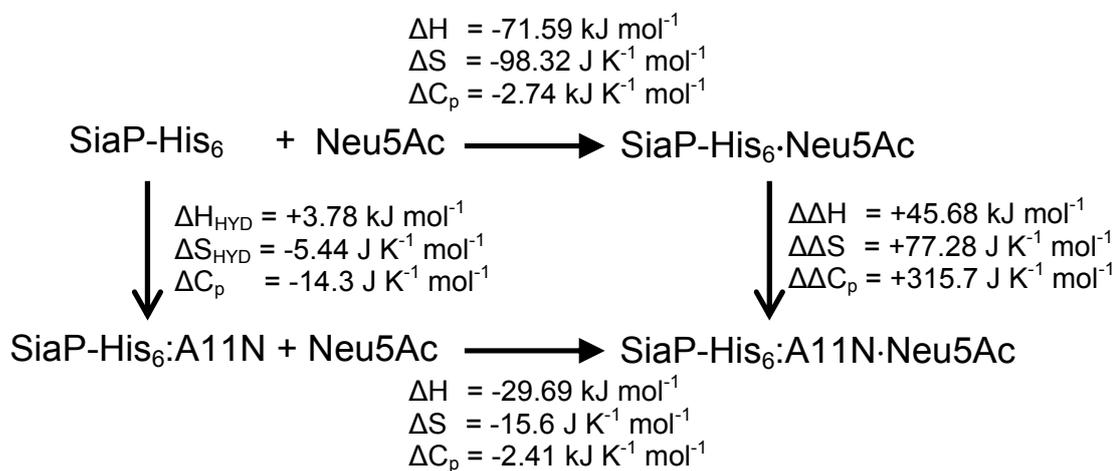
**Table 6.4:** The thermodynamic data from the ITC analysis of Neu5Ac binding by SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:A11N.

Protein	Temp (°C)	K <sub>d</sub> (μM)	ΔH (kJ mol <sup>-1</sup> )	ΔS (J mol <sup>-1</sup> K <sup>-1</sup> )	-TΔS (kJ mol <sup>-1</sup> )	ΔH-TΔS (J mol <sup>-1</sup> )	ΔC <sub>p</sub> (kJ K <sup>-1</sup> mol <sup>-1</sup> )
SiaP-His <sub>6</sub>	30	0.058 ± 0.005	-84.4 ± 3.0	-139.9 ± 10.1	42.4 ± 3.1	-42.0 ± 3.0	-2.74 ± 0.19
	15	0.027 ± 0.003	-49.0 ± 1.6	-24.9 ± 5.7	7.2 ± 1.6	-41.8 ± 1.6	
SiaP-His <sub>6</sub> :A11N	30	53.1	-5.8	65.7	-18.9	-24.7	-2.41 ± 0.02
	15	33.6	-41.9	-56.3	17.1	-24.8	

As previously, the differences in the values of  $\Delta H$ ,  $\Delta S$  and  $\Delta C_p$  are a combination of the differences in binding plus the direct differences introduced by the mutation of the amino acid side chains. The contributions of the alanine and asparagine residues were calculated using the method of Makhatadze and Privalov (Makhatadze & Privalov, 1993, Privalov & Makhatadze, 1993) (Dr Seishi Shimizu, personal communication) and are included in Figure 6.2.

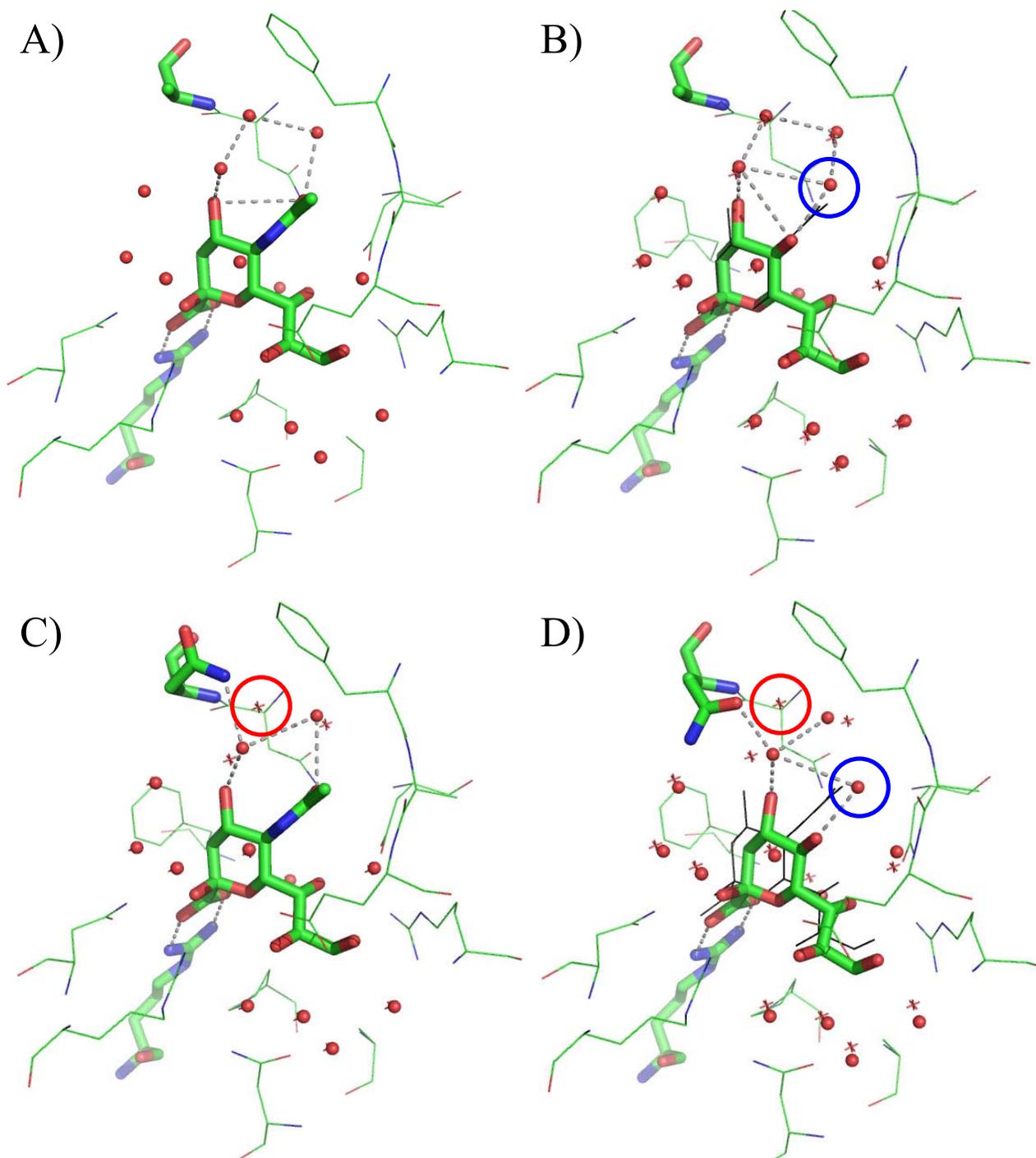
The cycle in Figure 6.2 gives the change in enthalpy to be an increase of  $45.7 \text{ kJ mol}^{-1}$ , and so binding is weaker. The large increases in  $\Delta S$  and  $\Delta C_p$  relate to large increases in disorder and flexibility. This could correspond to changes in the water network surrounding position 11, which would be revealed by the protein crystal structure.

The effects of these ligands and the A11N mutation on ligand binding were investigated further by comparing the protein crystal structures of SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:A11N with Neu5Ac and KDN bound (SiaP-His<sub>6</sub> with KDN and SiaP-His<sub>6</sub>:A11N with both ligands were determined by Marcus Fischer, YSBL). These structures reveal that the position of the methyl *N*-acetyl group from Neu5Ac is replaced by a single water molecule in the KDN-occupied binding site (Figure 6.3ab). This associated water molecule is the only addition between these two structures of SiaP-His<sub>6</sub> and fits in with the findings from the previous section. The increases in entropy and heat capacity both appear to correspond to the presence of this extra water molecule in the binding site of SiaP.



**Figure 6.2:** The thermodynamic cycles of Neu5Ac binding by SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:A11N. The effect of the amino acid modification is taken into account in the differences in the  $\Delta H$ ,  $\Delta S$  and  $\Delta C_p$  of both binding events. For example,  $\Delta\Delta H(\text{SiaP-His}_6\cdot\text{Neu5Ac} \rightarrow \text{SiaP-His}_6\text{:A11N}\cdot\text{Neu5Ac})$

$$\begin{aligned}
 &= -\Delta H(\text{SiaP-His}_6\cdot\text{Neu5Ac}) + \Delta H_{\text{hyd}}(\text{SiaP-His}_6 \rightarrow \text{SiaP-His}_6\text{:A11N}) + \Delta H(\text{SiaP-His}_6\text{:A11N}\cdot\text{Neu5Ac}) \\
 &= -(-71.59) + 3.78 + (-29.69) \\
 &= +45.68 \text{ kJ mol}^{-1}
 \end{aligned}$$



**Figure 6.3:** Representation of the ligand binding site from the crystal structures of SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:A11N bound to Neu5Ac and KDN. The ligands, Arg-147 and position 11 are shown as atom-colour cylinders, nearby residues as atom-colour lines and water molecules red spheres. The corresponding water molecules and ligand from the structure of SiaP-His<sub>6</sub> with Neu5Ac are shown as red crosses and black lines, respectively. Hydrogen bonds around the small network and Arg-147 are shown as dashed, grey lines. A) SiaP-His<sub>6</sub> with Neu5Ac. B) SiaP-His<sub>6</sub> with KDN. The extra water molecule is highlighted. C) SiaP-His<sub>6</sub>:A11N with Neu5Ac. The position of the displaced water molecule from the SiaP-His<sub>6</sub> structure is highlighted. D) SiaP-His<sub>6</sub>:A11N with KDN. The displaced and introduced water molecules are highlighted as in B and C.

The protein crystal structure of SiaP-His<sub>6</sub>:A11N bound with Neu5Ac shows the presence of the asparagine at position 11 (Figure 6.3c) that sterically intrudes upon the location of the three water molecules and displaces one of them, but without replacing it, as in the case of the ligand *N*-acetyl group (Figure 6.3b). The displaced water molecule did not interact directly with the ligand, but via the two other water molecules in this network of three. The explanation of the large decrease in affinity of this mutant for Neu5Ac seems to be limited to this location. The loss of the water molecule has caused change in the position of the two remaining water molecules, moving them away from their optimal positions.

The A11N mutation has a smaller effect on the  $K_d$  value for KDN than Neu5Ac (Table 6.2). The protein crystal structure of SiaP-His<sub>6</sub>:A11N bound with KDN shows a remarkable difference to the Neu5Ac-bound structure (Figure 6.3cd). Here, the Asn-11 residue is oriented towards the binding site and the ligand itself is positioned directly away from this residue, relative to the position of Neu5Ac. The charged hydroxyl group of the asparagine side chain is nearer to the location of the displaced water molecule and interacts with one of the remaining pair. A new three-water molecule hydrogen-bonded network exists with the remaining two from the original and the new water introduced by the absence of the ligand *N*-acetyl group. There are two other slight changes in the water molecules surrounding the ligand, but it would appear that this KDN-water network is more accepting of the A11N mutation than the Neu5Ac-network, including a translocation of the entire ligand, via the larger water network associated with KDN-binding.

## **6.2 *E. coli* can grow on different sialic acids as the sole carbon source using different sialic acid transporters**

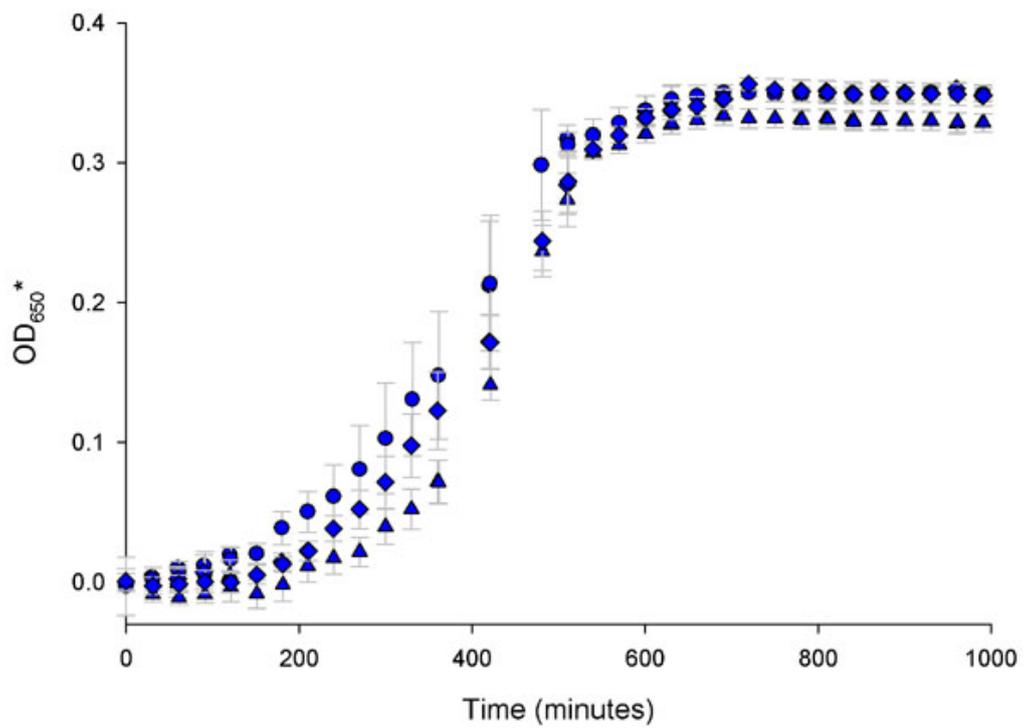
Growth of *E. coli* on Neu5Ac is well known and the use of Neu5Gc has previously been reported but not characterised in any way (Vimr & Troy, 1985). The use of KDN as the sole carbon source has not previously been examined in *E. coli*. The transport and catabolism of these by *E. coli* BW25113 could be investigated *in vivo* using the prototype incubated plate shaker mentioned in previous chapters. The small volume of the wells (<

1 ml) meant that growth on these expensive sialic acids could be investigated in a cost-effective manner. *E. coli* BW25113 was grown in M9 minimal medium supplemented with 0.5 mg/ml of each sialic acid as the sole carbon source. Growth was measured by the incubated plate shaker every 30 minutes at 31 °C with shaking at 250 rpm.

As can be seen in Figure 6.4, growth occurs on all three of these sialic acids to the same final density and in the order of Neu5Ac, KDN and Neu5Gc. Growth on Neu5Gc is not unexpected, given its similarity in structure to Neu5Ac. This structural similarity would explain the great similarity in growth rates between these two carbon sources. Growth on KDN as the sole carbon sources has not been reported previously and it should be noted that growth on this is at a similar rate to growth on Neu5Ac and Neu5Gc.

These differences in growth rate could have been due to differences in transport and catabolism or in the induction of the *nan* and *nag* operons by the substrate and its downstream products. To determine which of these occurs, the effect of pre-induction of these operons was examined by incubation with Neu5Ac before inoculation into the plate.

*E. coli* BW25113 was grown in M9 minimal medium containing 0.4% glycerol before being incubated in M9 1 mg/ml Neu5Ac for four hours. The cells were then harvested, washed and inoculated into M9 minimal medium with 0.5 mg/ml of each sialic acid as the sole carbon source. The growth of these was measured by the incubated plate shaker every 30 minutes at 35 °C with shaking at 250 rpm.



**Figure 6.4:** Growth curves of *E. coli* BW25113 on 0.5 mg/ml Neu5Ac (blue circles), Neu5Gc (blue up triangles) and KDN (blue diamonds).

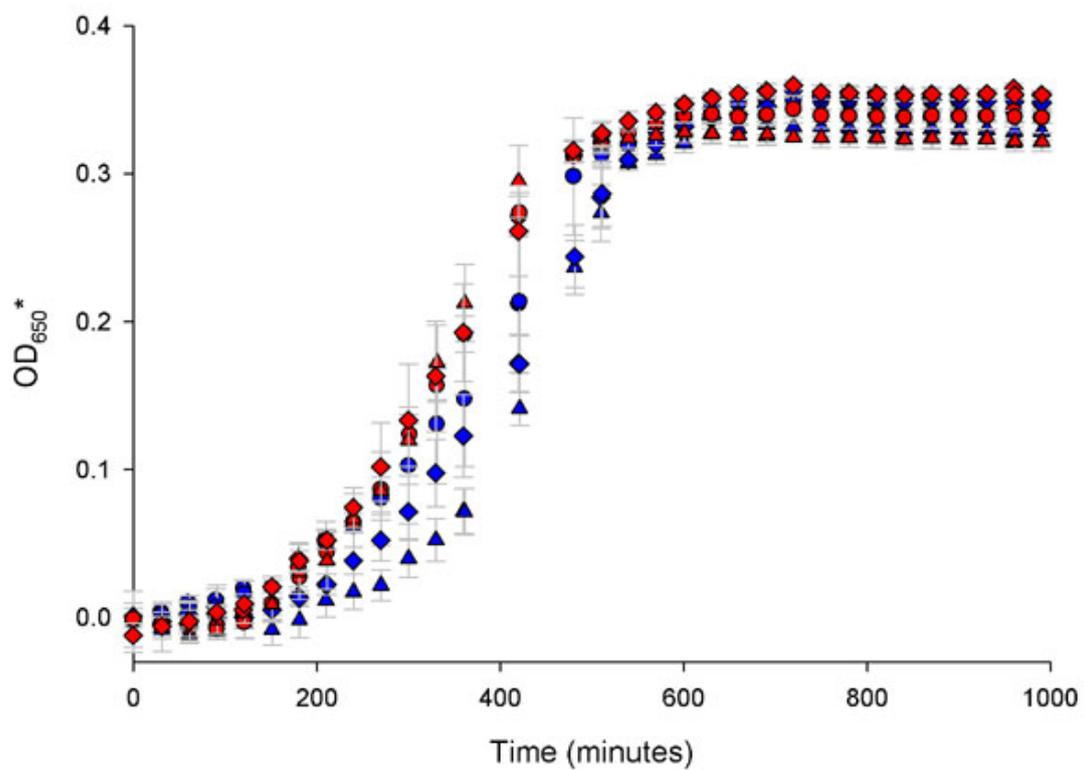
As can be seen in Figure 6.5, the growth rates following induction are above that of the uninduced on Neu5Ac and appear to be identical.

This increase in growth rate and apparent similarity between the three substrates following induction indicates that the differences in the growth rates in Figure 6.4 are due to the differences in induction by the different sialic acids and their downstream products.

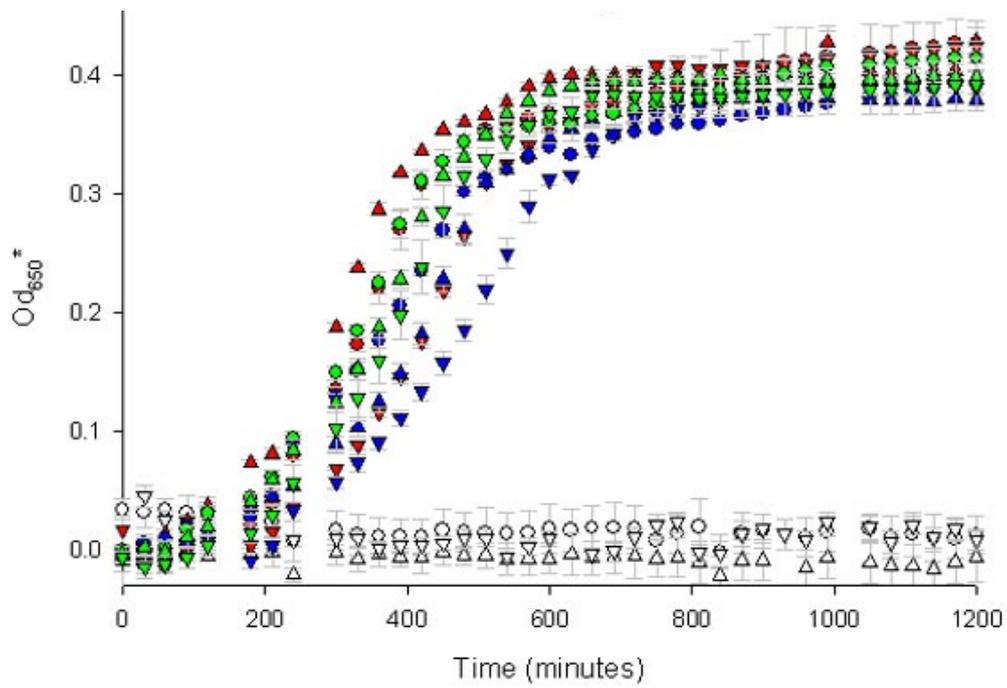
In the previous experiment, growth of *E. coli* on the sialic acids was mediated by the endogenous sialic acid transporter, NanT, which is a member of the Major Facilitator Superfamily (MFS). Using the previous *E. coli*  $\Delta nanT$  deletion strain and low copy number plasmid system, *siaPQM* was shown to allow growth on Neu5Ac (Figure 3.5). These, and another binding protein-independent transporter (STM1128) from the Sodium Solute Symporter (SSS) family, were examined for growth on the three sialic acids by expression from this low copy number plasmid.

As in the standard protocol, *E. coli* BW25113  $\Delta nanT$  expressing *nanT*, *siaPQM* or *stm1128* from a low copy number plasmid were grown in M9 minimal medium without FeSO<sub>4</sub>, supplemented with 0.4% glycerol to remove catabolic repression. These were harvested, washed and inoculated into M9 minimal medium (without FeSO<sub>4</sub>) with 0.5 mg/ml of each sialic acid as the sole carbon source. Their growth at 31 °C was measured by the incubated plate shaker every 30 minutes.

Growth on these three sialic acids with expression of the three sialic acid transporters is shown in Figure 6.6. Under all of these conditions, growth occurred at varying rates (Table 6.5 for doubling times,  $T_{GEN}$ ).



**Figure 6.5:** Growth curves of un-induced (blue) and pre-induced (red) *E. coli* BW25113 on 0.5 mg/ml Neu5Ac (circles), Neu5Gc (up triangles) and KDN (diamonds).



**Figure 6.6:** Growth curves of *E. coli* BW25113  $\Delta nanT$  strains expressing, from a low copy number plasmid, NanT (red), SiaPQM (blue) and Stm1128 (green). The sole carbon sources were: 0.5 mg/ml Neu5Ac (circles), Neu5Gc (up triangles) and KDN (down triangles). Growth of *E. coli* BW25113  $\Delta nanT$  pWKS30 on these carbon sources is shown by empty symbols.

**Table 6.5:** Growth rates of *E. coli* BW25113  $\Delta nanT$  strains expressing sialic acid transporters as indicated.

Transporter	0.5 mg/ml Neu5Ac		0.5 mg/ml Neu5Gc		0.5 mg/ml KDN	
	Growth rate ( $\text{min}^{-1}$ )	T <sub>GEN</sub> (min)	Growth rate ( $\text{min}^{-1}$ )	T <sub>GEN</sub> (min)	Growth rate ( $\text{min}^{-1}$ )	T <sub>GEN</sub> (min)
NanT	$8.18 \pm 0.10 \times 10^{-3}$	$122 \pm 2$	$9.39 \pm 0.10 \times 10^{-3}$	$107 \pm 3$	$6.10 \pm 0.15 \times 10^{-3}$	$164 \pm 4$
Stm1128	$8.41 \pm 0.10 \times 10^{-3}$	$119 \pm 2$	$8.17 \pm 0.17 \times 10^{-3}$	$122 \pm 3$	$7.40 \pm 0.10 \times 10^{-3}$	$135 \pm 2$
SiaPQM	$5.61 \pm 0.05 \times 10^{-3}$	$178 \pm 2$	$5.20 \pm 0.13 \times 10^{-3}$	$192 \pm 5$	$4.62 \pm 0.06 \times 10^{-3}$	$216 \pm 3$

With constitutive expression, all of these different families of transporters transport different sialic acids and allow growth at similar rates with doubling times varying between about 100 and 200 minutes. This similar transport of different ligands suggests that multiple sialic acids could be relevant substrates for these organisms.

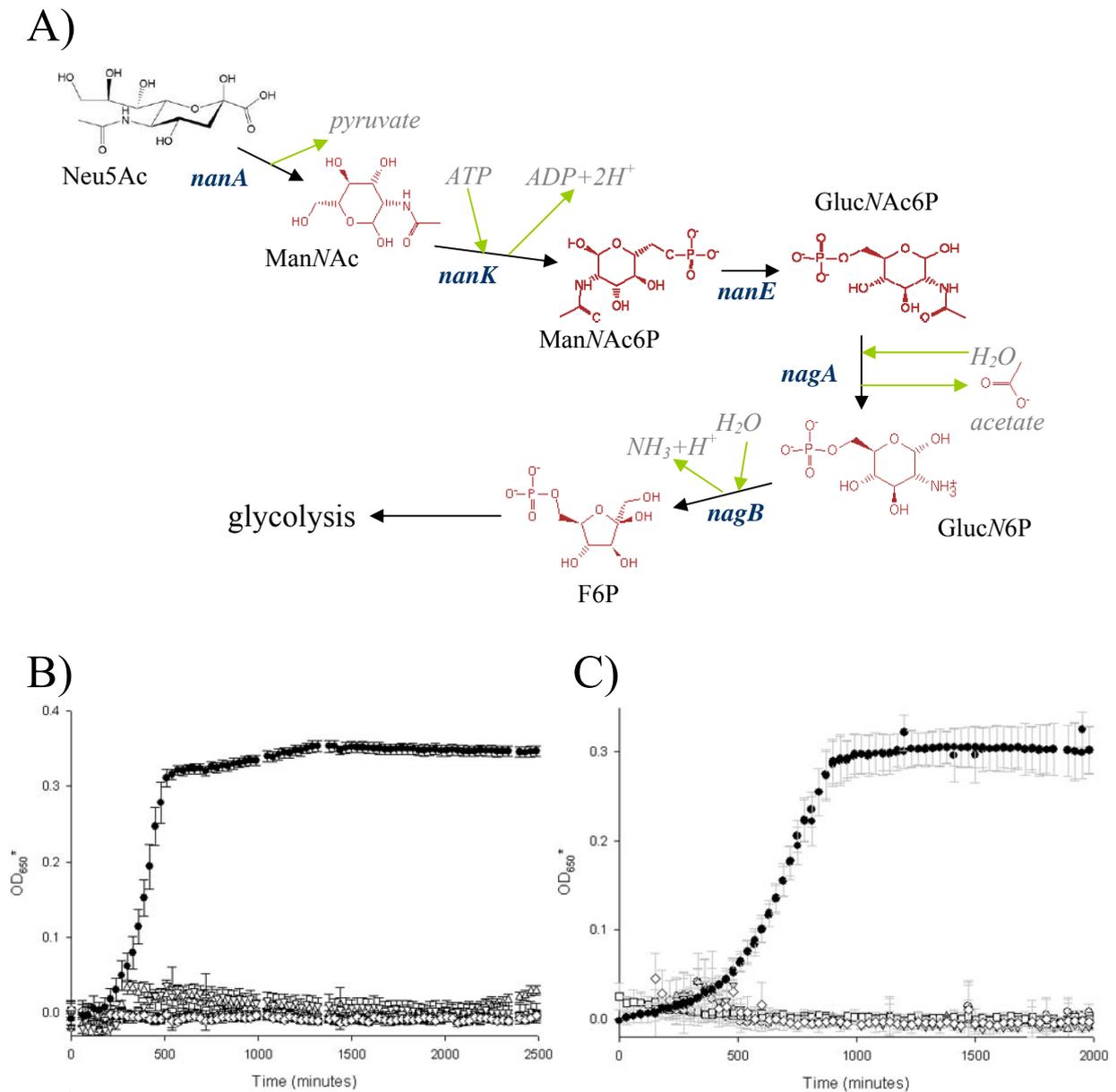
The doubling times calculated for each transporter with Neu5Ac and Neu5Gc are all very similar, the major differences are seen with KDN as the substrate. Of these transporters, SiaPQM stands out as consistently allowing the slowest growth on all three substrates, which could be due to the involvement of the binding protein requiring a second association step in the transport cycle. STM1128 allows a very similar growth rate on all three sialic acids, while NanT catalyses the fastest growth on Neu5Gc. This could be evidence of differences in the ligand-binding mechanism of these two proteins, where there are, relatively, less critical interactions around the *N*-acetyl group or at least better acceptance of its replacement such that its absence does not reduce the rate of uptake.

### **6.3 The catabolism of sialic acids requires relevant components of the sialometabolic pathway in *E. coli***

#### **6.3.1 The catabolism of Neu5Gc requires the *nan* and *nag* genes and releases glycolate**

Growth of *E. coli* on Neu5Ac is dependent on the presence of the sialometabolic genes *nanA*, *nanK*, *nanE*, *nagA* and *nagB* (Figure 6.7a) (Plumbridge & Vimr, 1999, Vimr & Troy, 1985). Intracellular Neu5Ac is processed by these three enzymes from the *nan* operon, the inducer of which might not be Neu5Ac (Kalivoda *et al.*, 2003). The product of these is *N*-acetyl-D-glucosamine-6-phosphate (GlucNAc6P), which induces the *nag* pathway and is processed to D-fructose-6-phosphate (F6P). The metabolism of Neu5Gc would also be expected to follow this pathway.

Strains of *E. coli* BW25113 containing kanamycin cassette replacement deletions of the sialometabolic genes or *ytfQ* (positive control) were grown in M9 minimal medium



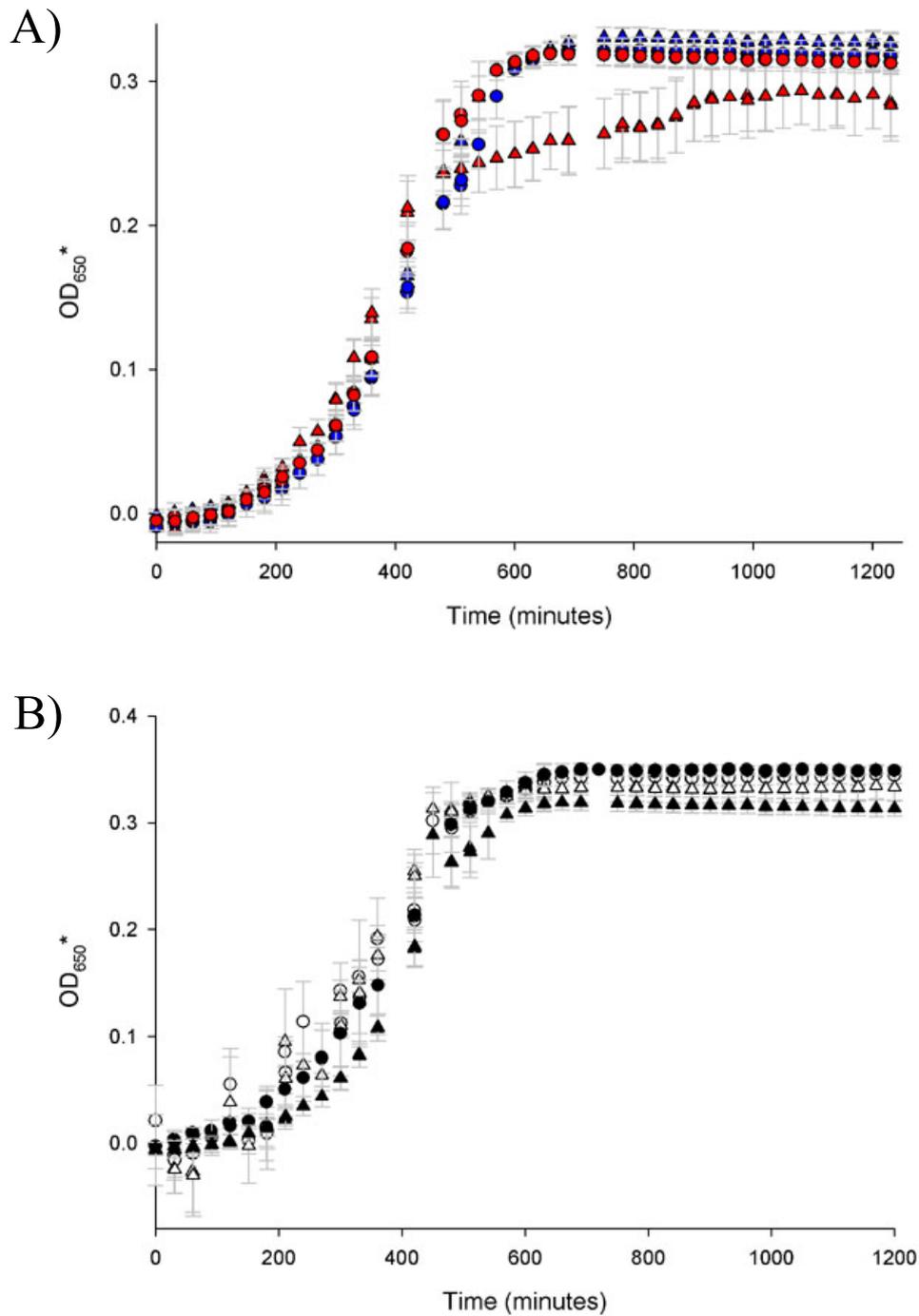
**Figure 6.7:** Growth on Neu5Ac and Neu5Gc of *E. coli* strains with deletions in the sialometabolic pathway. **A)** The sialometabolic pathway of *E. coli*. **ManNAc**, *N*-acetyl-D-mannosamine; **ManNAc6P**, *N*-acetyl-D-mannosamine-6-phosphate; **GlucNAc6P**, *N*-acetyl-D-glucosamine-6-phosphate; **GlucN6P**, D-glucosamine-6-phosphate; **F6P**, D-fructose-6-phosphate. **B)** Growth on 0.5 mg/ml Neu5Ac of *E. coli* BW25113 with kanamycin cassette insertions in place of *nanA*, *nanK*, *nanE*, *nagA*, *nagB* (empty symbols) and *yjhS* (black circles). **C)** Growth on 0.5 mg/ml Neu5Gc of *E. coli* BW25113 with kanamycin cassette insertions in place of *nanA*, *nanK*, *nanE*, *nagA*, *nagB* and *yjhS*.

supplemented with 0.4% glycerol to remove catabolic repression. These cells were harvested, washed and inoculated into M9 minimal medium supplemented with 0.5 mg/ml Neu5Ac or Neu5Gc as the sole carbon source. Growth was then measured by the incubated plate shaker. Growth of the control and deletion strains on Neu5Ac and Neu5Gc is shown in Figure 6.7bc. From these, there is no growth on either of the sialic acids as the sole carbon source without the intact sialometabolic pathway.

The absence of growth by all of the sialometabolic deletion strains shows that Neu5Gc follows this pathway just as Neu5Ac. The only expected difference is the production of the carbon source glycolate instead of acetate by NagA. This was examined using a deletion of the *glcE* gene, which encodes the FAD-binding subunit of glycolate oxidase and is essential for this activity (Pellicer *et al.*, 1996).

As in the standard protocol, *E. coli* BW25113 containing kanamycin cassette replacement deletions of *ytfQ* and *glcE* were inoculated into M9 minimal medium supplemented with 0.5 mg/ml Neu5Ac or Neu5Gc as the sole carbon source. Growth was then measured by the incubated plate shaker. Growth of the control and *glcE* deletion strains on Neu5Ac and Neu5Gc is shown in (Figure 6.8a). This shows a reduction in the final optical density of the *glcE* deletion strain when grown on Neu5Gc, suggesting the production of glycolate during Neu5Gc catabolism contributes to growth.

It has recently been suggested that the final gene in the *nan* operon, *yhcH*, is essential for growth on Neu5Gc by converting it to Neu5Ac (Roy *et al.*). Using the incubated plate shaker, this could quickly and easily be investigated. As in the standard protocol, the strains containing kanamycin cassette replacement deletions of *ytfQ* and *yhcH* were inoculated into M9 minimal medium supplemented with 0.5 mg/ml Neu5Ac or Neu5Gc as the sole carbon source. Growth was then measured by the incubated plate shaker.



**Figure 6.8:** The effect of different deletions in *E. coli* BW25113 on growth on Neu5Gc. **A)** Growth of *E. coli* BW25113 with kanamycin cassette insertions in place of *ytfQ* (circles) and *glcE* (triangles) on 0.5 mg/ml Neu5Ac (blue) and Neu5Gc (red). **B)** Growth of *E. coli* BW25113 with kanamycin cassette insertions in place of *ytfQ* (black) and *yhcH* (white) on 0.5 mg/ml Neu5Ac (circles) and Neu5Gc (triangles).

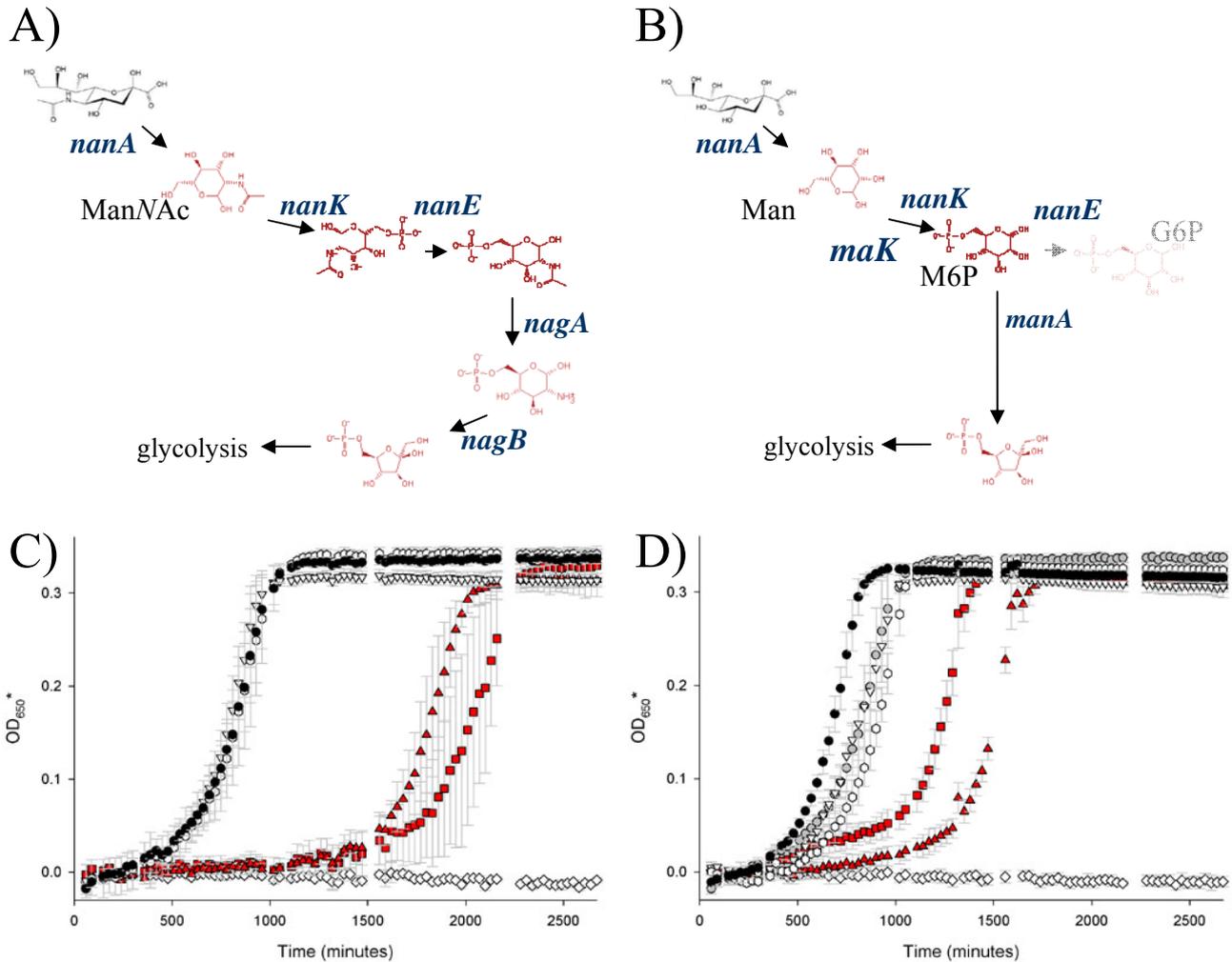
Growth of the control and *yhcH* deletion strains on Neu5Ac and Neu5Gc is shown in (Figure 6.8b). As can be seen, the *yhcH* deletion strain can grow on both Neu5Gc and Neu5Ac, demonstrating that this gene is non-essential for catabolism of these sialic acids. This is not unexpected given the structural similarity between these two compounds and that NanA has been shown to process both of these (Comb & Roseman, 1960).

### 6.3.2 The catabolism of KDN requires *nanA*

The examination of the sialometabolic pathway implies that KDN would require only the first enzyme of the *nan* genes, NanA (Figure 6.9ab). The lyase would remove pyruvate from KDN leaving D-mannose (Man), which would then enter mannose degradation via Mannofructokinase, MaK (Kornberg *et al.*, 2000).

As previously, strains of *E. coli* BW25113 containing sialometabolic enzyme deletions were grown in M9 minimal medium 0.4% glycerol before being harvested, washed and inoculated into M9 minimal medium with 0.5 mg/ml KDN as the sole carbon source. To investigate the effect of induction, this was repeated but before inoculation into wells of the plate, the cells were incubated for four hours in M9 minimal medium 1 mg/ml Neu5Ac. The growth of the deletion strains under both of these conditions was measured by the incubated plate shaker every 30 minutes at 35 °C with shaking at 250 rpm.

Growth of the deletion strains on KDN reveals that *nanA* is essential, while *nagA* and *nagB* are not (Figure 6.9c). *In vitro*, NanA releases pyruvate from KDN (Judith Hawkhead, personal communication). As expected, *nanK* and *nanE* are non-essential but, unexpectedly, their deletion causes a very long delay in growth. This delay could have been due to poor induction of *nan* operon by KDN and its downstream products. To examine this, the deletion strains were grown under the same conditions except for a 4-hour incubation with Neu5Ac before inoculation into the plate. Figure 6.9d shows the growth of these pre-induced cells. Again, *nanA* is essential; *nagA* and *nagB* are nonessential for growth, but these strains do not experience the reduction in lag time seen in the control deletion strain (*ytfQ*). Pre-induction of the *nanK* and *nanE* deletions strains



**Figure 6.9:** The effects of deletions in the sialometabolic pathway on growth of *E. coli* on KDN. **A)** The sialometabolic pathway of *E. coli*. **ManNAc**, N-acetyl-D-mannosamine. **B)** The expected catabolic pathway taken by KDN in *E. coli*. **Man**, D-mannose; *maK*, mannofructokinase; M6P, D-mannose-6-phosphate; **G6P**, D-glucose-6-phosphate; *manA*, M6P isomerase. G6P is transparent since this reaction has not been confirmed, just that M6P can interact with NanE to inhibit ManNAc6P-GluNAc6P epimerisation (Larrion *et al.*, 2007). **C)** Growth on 0.5 mg/ml KDN of *E. coli* BW25113 with kanamycin cassette insertions in place of *nanA* (empty diamonds), *nanK* (red squares), *nanE* (red up triangles), *nagA* (empty down triangles), *nagB* (empty circles) and *yjhS* (black circles). **D)** Growth of the same strains on 0.5 mg/ml KDN after four hours incubation with 1 mg/ml Neu5Ac. Symbols are the same except for *yjhS* deletion (black circles) are also shown without pre-induction (grey circles).

causes a decrease in the delay, which is much more pronounced for the *nanK* deletion.

It would appear that *nanK* and *nanE* are non-essential for growth on KDN, but are able to process the KDN products down stream of NanA to accelerate, or avoid an inhibition of, growth. The interaction of NanK and NanE with D-mannose and D-mannose-6-phosphate (M6P), respectively have been shown *in vitro* and so this could potentially be occurring *in vivo* (Ferrero *et al.*, 2007, Larion *et al.*, 2007).

With pre-induction by incubation with Neu5Ac, this delay is diminished but still present, suggesting that these enzymes are useful in processing KDN. Assuming that the incubation with Neu5Ac has resulted in full induction of the *nan* operon, it could be that NanK is providing an initial pathway to D-mannose-6-phosphate from D-mannose and that NanE may process D-mannose-6-phosphate to Glucose-6-phosphate (G6P) before the intracellular concentration becomes inhibitory. However, induction for the operon may not be full under these conditions, depending on which of Neu5Ac, *N*-acetyl-D-mannosamine and *N*-acetyl-D-mannosamine-6-phosphate are the true *nan* inducers. Both deletion strains experience a Neu5Ac-dependent reduction in the lag time, which, following this argument, supports that either Neu5Ac or the first *nan* product, *N*-acetyl-D-mannosamine, is the *nan* inducer. However, it has previously been suggested that these are not the true inducers and so this long lag phase could be caused by the slow production of an unknown inducer further downstream.

Alternatively, the poor growth of these deletion strains could be due to a detrimental effect of the kanamycin cassette on the down stream genes in the operon. This would implicate the final *nan* gene, *yhcH*, in the utilisation of KDN.

### **6.3 Summary**

This section has focussed on the binding, transport and ultimately the catabolism of the three different sialic acids Neu5Ac, Neu5Gc and KDN. Firstly, the binding of these was investigated using the existent mutant library (Table 6.1). From this followed the examination of the water environment around the ligand *N*-acetyl group using

thermodynamic and crystallographic methods (Figure 6.3). Secondly, sialic acid transporters from three different transporter families (MFS, SSS and TRAP-T) were shown to transport the different sialic acids and restore growth to *E. coli* BW25113  $\Delta nanT$  (Table 6.5). Finally, the catabolism of these sialic acids was investigated using deletions in the sialometabolic pathway of *E. coli* BW25113. These showed that the *N*-acetyl and *N*-glycolyl containing Neu5Ac and Neu5Gc were catabolised by the full sialometabolic pathway, whereas the de-aminated analogue KDN required only the first *nan* gene with the kanamycin cassette insertions at *nanK* and *nanE* causing a very long, induction-sensitive lag phase (Figures 6.7 and 6.9).

# Chapter Seven

## Discussion

This discussion chapter is split into three main themes. The first covers sialic acid binding by SiaP, while the second discusses the transport of sialic acid by the whole transporter in the context of both the binding site and surface mutations. The third section includes the growth of *E. coli* on the various sialic acids that were available and their transport by different transporters.

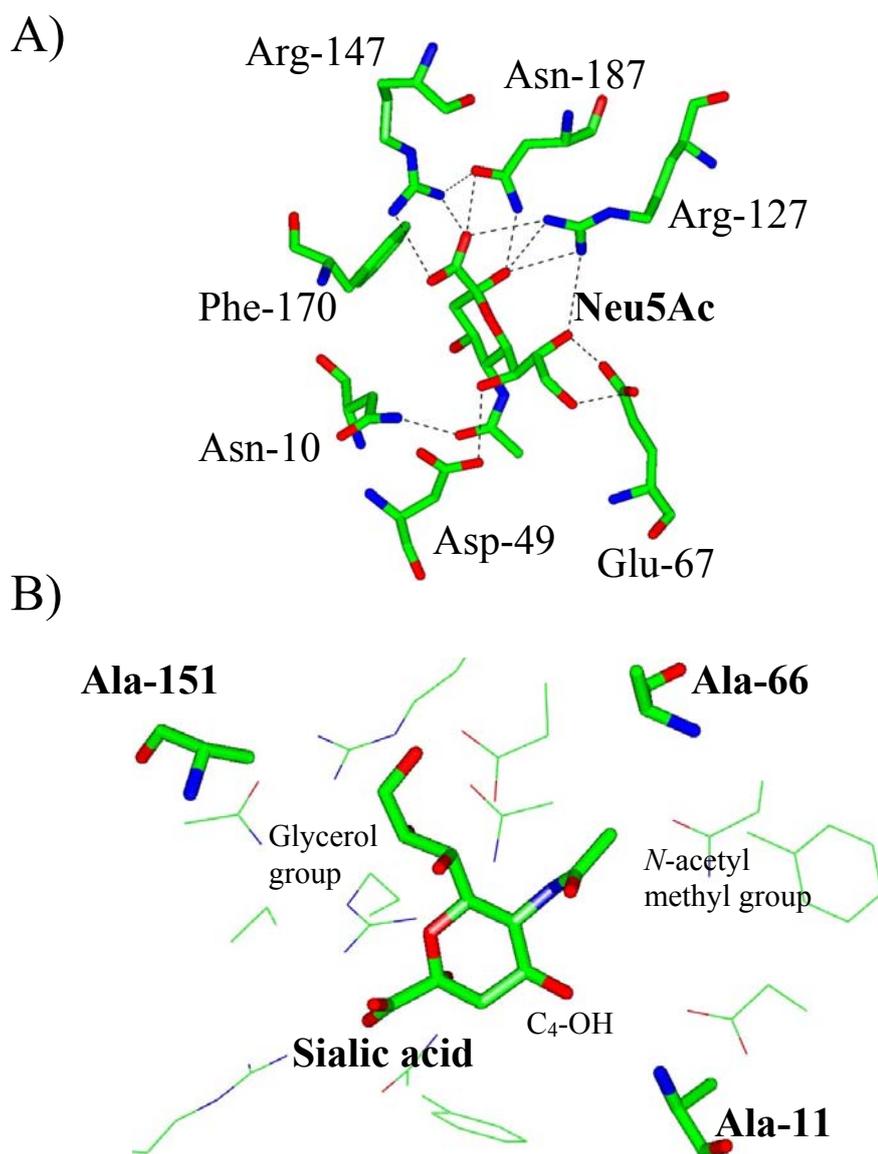
## **7.1 The binding of sialic acid by SiaP is sensitive to conditions and mutation**

During the course of this project, the binding site of SiaP was investigated using a wide variety of techniques. These ranged from the determination of the thermodynamics of binding different ligands to the contribution of individual amino acid residues in the binding of the ligand and transport in the intact transport complex. Using a titration of the protein's intrinsic fluorescence, it was confirmed that SiaP binds Neu5Ac with a  $K_d$  value of  $0.12 \pm 0.02 \mu\text{M}$  in 50 mM Tris/HCl pH 8.0 at 37 °C.

### **7.1.1 Can SiaP function without domain closure?**

Co-ordination of the carboxylate group by an arginine residue is a common method of sialic acid binding and the DctP family SBPs contain a totally conserved arginine residue that is involved in organic acid carboxylate binding. The protein crystal structure of Neu5Ac-bound SiaP showed three residues involved in ligand carboxylate stabilisation: Arg-127, Arg-147 and Asn-187. Arg-147 corresponds to the totally conserved arginine residue and forms a bipartite salt bridge with the ligand carboxylate group (Figure 7.1a). These positions were mutated to the amino acids shown in Table 3.1. It was found that growth was allowed at a sialic acid concentration as low as 650  $\mu\text{M}$  (Section 3.7), while no binding could be detected *in vitro* at Neu5Ac concentrations up to 3 mM in the *siaP-His<sub>6</sub>* background (Table 3.2) and 5 mM in the reporter background (Table 3.3). This apparent contradiction was investigated at length during this work.

No interaction with Neu5Ac could be detected *in vitro* for any of these mutants except for the most conservative change to Arg-127 (to lysine), which gave a  $K_d$  value of about 1 mM (Table 3.2). Using the ligand binding reporter variant (*siaP-His<sub>6</sub>:F170W*) as



**Figure 7.1:** Representation of all of the targets for mutation within the binding site of SiaP, based on 3B50.pdb (Johnston *et al.*, 2008). The ligand and the amino acids of interest are shown as atom coloured cylinders; hydrogen bonds are shown as dashed black lines. **A)** Amino acid residues that form direct electrostatic interactions with the Neu5Ac ligand. **B)** Amino acid residues within the binding site that do not form electrostatic interactions with the ligand. Other binding site residues are shown as atom coloured lines for clarity.

background, this R127K mutant gave a  $K_d$  value of 1.5 mM, while no interaction with Neu5Ac could be detected for any other mutant except for N187A, which gave a  $K_d$  value of just  $244 \pm 8 \mu\text{M}$ . This surprising result could be due to a beneficial interaction between these two close positions ( $< 4 \text{ \AA}$ ) or the reporter variant revealing the previously undetectable fluorescence signal change, as intended. This insertion of tryptophan residues to generate reporter variants has been used previously in studies of ligand binding systems such as calmodulin and the nucleotide binding site of F1-ATPase (Kilhoffer *et al.*, 1992, Weber *et al.*, 1998).

The Arg-147 mutants were investigated further for sialic acid binding using a variety of different methods including the recovery of bound  $^{14}\text{C}$ -Neu5Ac, ITC, gel shift assays and thermal denaturation. From these, no interaction between sialic acid and the mutant proteins could be detected. All of the methods used are based on the closure of the domains of the SBP and so it was concluded that this was not occurring, due to either a disruption of the closing mechanism or that the  $K_d$  for sialic acid was significantly above the attainable sialic acid concentrations. However, it has been shown that maltose binding protein binds reduced, oxidised and cyclic maltodextrins without the need for domain closure (Hall, Ganesan *et al.* 1997; Hall, Thorgeirsson *et al.* 1997). The extent of closure could be investigated by paramagnetic relaxation enhancement (PRE) NMR with labelled cysteine residues at positions away from the binding cleft. The closing mechanism could have been disrupted by a disturbance of a ‘trigger’ residue in the binding site that is used to detect the presence of the ligand and promote the closure of the domains (Sharff *et al.*, 1992). In SiaP, this was previously proposed to be the Arg-127 residue as this is located in the hinge region and comes into contact with the ligand (Muller *et al.*, 2006); this could be investigated further using by introducing the R127A mutation in the *in vivo* system to examine its phenotype compared to those of the Arg-147 mutants.

At this point, it is worth reiterating the failure of hydrophobic interaction chromatography to separate SiaP:R147A and SiaP:R147K (Figure 3.2) and the differences in mobility on native PAGE (Figure 3.15). The retention of the proteins by the HIC column is due to an

improved interaction with the matrix, which could be due to a greater hydrophobic surface being exposed in these mutant proteins, that is, slight unfolding.

The mobility on native PAGE also varies between the mutants. The hexahistidine tag has no effect on the migration of the native protein, while SiaP:R147E, SiaP-His<sub>6</sub>:R147A and SiaP-His<sub>6</sub>:R147K show a slight retardation, probably due to a slight increase in size or overall charge. This difference is slightly less in the case of SiaP:R147E, which purified normally by HIC. However, the combination of the hexahistidine tag and the R147E mutation causes a large increase in electrophoretic mobility. The explanation for these results is not clear at this point, but the nature of the open conformations of these mutant proteins may hold clues to this.

To investigate the apparent absence of domain closure, transport by the Arg-147 mutants was investigated using the *in vivo* assay. This transport-positive phenotype was confirmed independently via an *in vivo* uptake assay of <sup>14</sup>C-Neu5Ac by *E. coli* BW25113 *AnanT* over a similar time course as the *in vitro* uptake assay in proteoliposomes. These showed the accumulation of <sup>14</sup>C-Neu5Ac in the same pattern as the *in vivo* growth phenotype (Dr. Emmanuel Severi, personal communication).

The interaction of open conformation SiaP with SiaQM has been shown previously, where an excess of unliganded SiaP can catalyse the reverse transport of <sup>14</sup>C-Neu5Ac from loaded SiaQM-proteoliposomes (Mulligan *et al.*, 2009). This would suggest that these mutants might bind the ligand with one domain, as in the initial binding of maltose by MalE (Spurlino *et al.*, 1991), or are fulfilling a mechanistic role as a scaffold, generating a transport-competent complex in which ligand binding by the SBP is not critical.

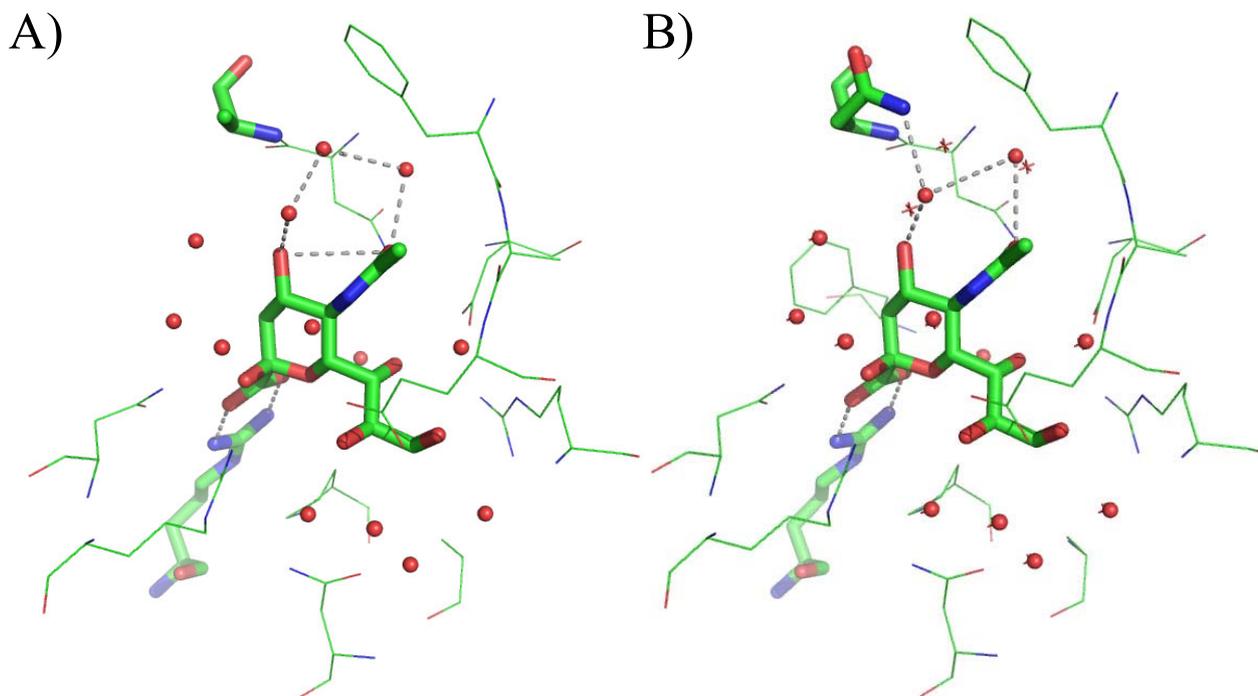
To determine if these mutants were capable of binding sialic acid and undergoing domain closure, the protein crystal structures were determined in the presence of a high concentration of sialic acid (Marcus Fischer, YSBL). All three were found in the closed, ligand-bound conformation. From this, it is likely that these apparently transport-competent mutants could be supplying Neu5Ac via the small, transient population of closed, ligand-bound complexes that occur in an unsaturated equilibrium. This could be

investigated by PRE using the same methods as Bermejo *et al.* and Tang *et al.* (2007). This also shows that the totally conserved arginine residue is critical only for high affinity ligand binding, and is not mechanistically essential for domain closure.

### **7.1.2 The sialic acid binding affinity of SiaP is sensitive to conditions and mutation of the binding site**

Throughout the course of this work, fifteen mutations have been made at positions around the ligand binding site (Figure 7.1). Mutations of Arg-127, Arg-147 and Asn-187 have been investigated for ligand binding using both the His<sub>6</sub>-tagged native protein and the reporter variant as backgrounds to the mutation (Sections 3.5 and 3.7). All of these have shown undetectable or extremely low affinity (approximately 9000-fold decrease or greater) for Neu5Ac (except for SiaP-His<sub>6</sub>:F170W;N187A at  $244 \pm 8 \mu\text{M}$ ). In Section 4.3.1, four mutations were targeted to positions around the ligand binding site that did not form direct interactions with the bound Neu5Ac (Figure 7.1b). These were all found to cause a decrease in the affinity for the ligand by between 5 – 900-fold.

In Section 4.3.1, it was hoped that the SiaP-ligand interactions were a compromise between binding and eventual release into the transporter, so that, for *in vitro* applications, the affinity could be increased at the expense of ligand release. However, the mutations that were introduced, even if they formed the intended interactions at all, have actually disturbed the binding of the ligand, causing a decrease in affinity of 5–900-fold. The most rigorously investigated of these was the Ala-11-Asn mutation in Section 6.1. The protein crystal structure of this (Figure 7.2) shows that, in fact, the mutated residue forms no direct interactions with the ligand, but instead displaces a secondary water molecule and alters the co-ordination of two nearby water molecules from the shell around the ligand. Such detrimental effects as this would suggest that the binding site and its associated water molecules form very specific, but also very sensitive, interactions with sialic acid.

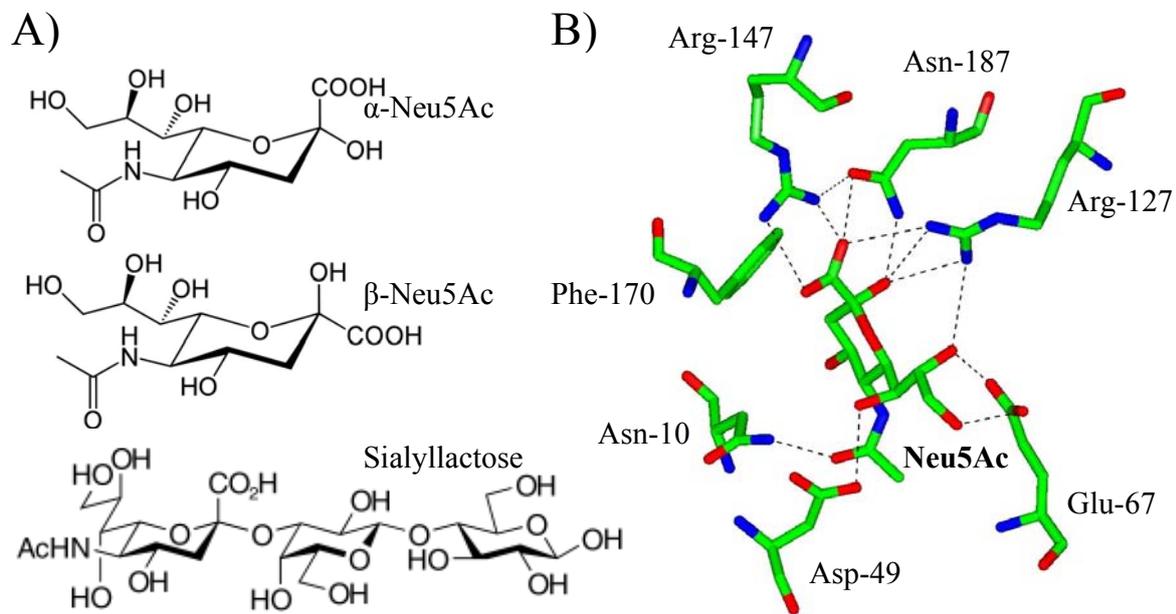


**Figure 7.2:** Representation of the ligand binding site from the crystal structures of (A) SiaP-His<sub>6</sub> and (B) SiaP-His<sub>6</sub>:A11N bound to Neu5Ac. Neu5Ac, Arg-147 and position 11 are shown as atom-colour cylinders, nearby residues as atom-colour lines and water molecules red spheres. In the SiaP-His<sub>6</sub>:A11N structure (B) the corresponding water molecules and ligand from the structure of SiaP-His<sub>6</sub> (A) are shown as red crosses and black lines, respectively. Hydrogen bonds around the small network and Arg-147 are shown as dashed, grey lines.

The sensitivity of sialic acid binding by SiaP was also found during the investigation of the biophysical properties of the protein. The sensitivity to ethanol is not unexpected, since this organic solvent disturbs the secondary structure of proteins by promoting the formation of  $\alpha$ -helical content (Section 4.1.1) (Buhrman *et al.*, 2003, Deshpande *et al.*, 2005, Knubovets *et al.*, 1999). During the examination of the pH-dependence of ligand binding, a small change in affinity was noted between pH 7 and 7.5. This could not be attributed to a change in the protonation state of any component of the system and so it is likely that the change in affinity is due to the sensitivity of ligand binding to the ionic strength of the buffer. The temperature sensitivity (Figure 4.4a) is also not unexpected for the enthalpically-driven ligand binding event in a binding protein that undergoes a large conformational change to a closed form that is stabilised by weak bonds between the domain-domain and protein-ligand interfaces and the release of large numbers of water molecules from the protein and the ligand.

Great temperature sensitivity would be expected of all DctP-type SBPs, since all of the known substrates are small, hydrophilic molecules, of which sialic acid is the largest. Indeed, it was odd that SiaP had previously been reported to bind to the even larger, conjugated sialic acid, sialyllactose, with a high affinity (Muller *et al.*, 2006). In Section 4.2.3, this was refuted and it is now clear that SiaP cannot bind to sialic acids that are modified or conjugated at the reducing sugar terminus (C<sub>2</sub>-OH). An examination of the protein crystal structure does support this, since these modifications at the anomeric C<sub>2</sub> position would prevent the co-ordination of the ligand carboxylate by the relocation of the carboxylate group to the  $\alpha$ -position and the presence of the modification at the  $\beta$  position (Figure 7.3).

This monomeric ligand specificity might be expected of the DctP family, in contrast to other SBPs such as MalE, which can bind to ligands of various lengths from maltose (glucose-1,4-glucose) up to amylose (glucose oligomer) and even cyclodextrins (cyclic glucose oligomers) (Ferenci, 1980). The ability of MalE to bind longer chain ligands means that the interactions made on binding can accept the modifications to the reducing hydroxyl of the sugar ring. In the case of sialic acid, modification of the reducing



**Figure 7.3:** Conjugated sialic acids could not be bound by SiaP. **A)** The structures of the  $\alpha$ - and  $\beta$ -anomers of Neu5Ac and sialyllactose, where the carboxylate group is in the  $\beta$ -position. **B)** The  $\beta$ -Neu5Ac occupied binding site of SiaP. The ligand and the co-ordinating residues are shown as atom coloured cylinders with hydrogen bonds as dashed lines.

terminus disturbs the environment around the carboxylate group and can reposition it to the  $\alpha$ -anomer (Figure 7.3).

## **7.2 The transport of sialic acid by SiaPQM is dependent upon ligand binding by SiaP and the correct interaction between the SiaP and SiaQM**

In Section 4.3.1, a group of ligand binding site mutants was developed that were intended to increase the affinity for Neu5Ac by introducing more interactions with the bound ligand (Figure 7.1b). Unfortunately, all of these caused a decrease in affinity, but did result in a library of binding site mutants with decreasing affinity for Neu5Ac, down to the 100  $\mu$ M range. Since these mutations were all within the binding site of the protein, they could be used to investigate the effect of decreasing affinity on transport without directly affecting the interaction between SiaP and SiaQM.

Unsurprisingly, it was found that decreasing affinity caused a decrease in the rate of transport at 5  $\mu$ M ligand (Figure 5.4). Using the example of SiaP-His<sub>6</sub>:A66M mutant, near maximal uptake could be restored by saturation of the binding protein using a sialic acid concentration  $\sim$ 10-times the  $K_d$  value. This demonstrates that the rate of uptake is dependent upon the ligand occupancy of the binding protein.

In Section 4.3.2, a group of mutations on the surface of SiaP was developed that were intended to decrease the  $K_d$  value of the protein for the ligand by stabilising the closed conformation of the protein. All but one of these (N150D) caused a reduction in the affinity of the protein for sialic acid. The *in vitro* uptake assay was used to investigate the effect of several of these mutations on the interaction with SiaQM (Section 5.4.1). In this group, the reduced transport of the Q72E;A152K double mutant was shown to be caused by the decreased affinity for Neu5Ac, while N150D and S15K;A195D abolished and permitted uptake, respectively (Figures 5.5 and ND uptake).

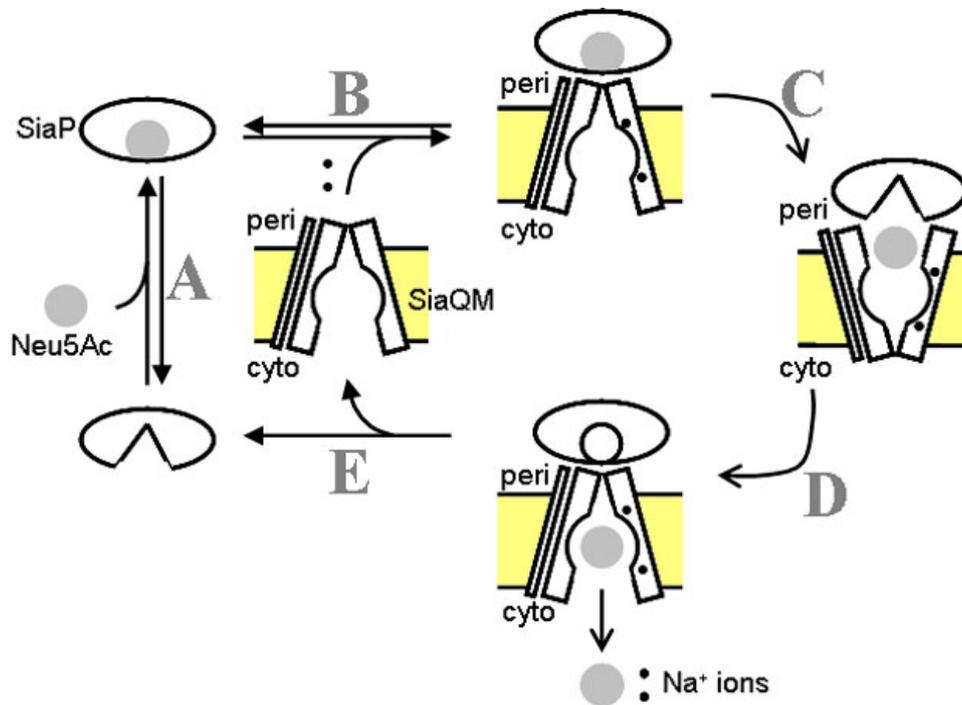
The S15K;A195D double mutant catalysed-uptake was indistinguishable from the control, despite the saturation state of this mutant ( $K_d$   $9 \pm 2$   $\mu$ M). Previous findings

suggested that a reaction concentration of 90  $\mu\text{M}$  Neu5Ac should have been needed to reach maximal uptake, given its  $K_d$  value. This surprising result of an unsaturated binding protein (approximately 35% occupancy) catalysing a normal uptake rate could be due to either of two different methods; a general scheme for transport is proposed in Figure 7.4 to clarify the effects of these, and further, mutations on transport.

The first method could be that these mutations have improved the duration or stability of the interaction between the binding protein and SiaQM, which would alter the association or dissociation of Step B in the scheme in Figure 7.4. This could increase the apparent concentration in proximity to the membrane component by maintaining the interaction between the two components throughout the transport cycle. This could be a sustained association, giving a similar effect as transporters with SBP fusions (van der Heide & Poolman, 2002), but would require the substrate to be able enter the binding protein. Alternatively, the interaction of a single domain could have been improved, allowing the SBP to open while associated to the permease, as in the *S. typhimurium* histidine ABC transporter (Ames *et al.*, 1996). To examine this, the strength of the interaction between this mutant SBP and SiaQM could be determined by surface plasmon resonance (SPR) or crosslinking between the two components.

The second could be that this double mutant has destabilised the closed conformation of SiaP, promoting the open conformation. The  $K_d$  value is made up of the on- ( $k_{\text{on}}$ ) and off-rates ( $k_{\text{off}}$ ) of the equilibrium shown as Step A in Figure 7.4. A destabilised closed conformation would increase the off-rate, that is, the reverse of Step A, resulting in a lower  $K_d$  value without affecting the on-rate. This could be determined by stopped-flow spectroscopy to investigate the on- and off-rates.

In the protein crystal structure of Neu5Ac-bound SiaP, the loop containing Phe-170 has moved relative to the rest of the domain compared to the open conformation of SiaP and this Phe-170 forms a hydrophobic interaction with the ligand. Investigation of this residue has suggested that this position acts as a lid for the binding site and also has a role in transport. When truncated to alanine, the ligand affinity decreases  $\sim 3000$ -fold and transport *in vitro* is undetectable (Table 3.2 and Figure 5.3). When replaced by polar



**Figure 7.4:** The proposed sialic acid transport scheme of SiaPQM. **A)** SiaP exists in an equilibrium between the open and closed, ligand-bound forms. **B)** Closed, ligand-bound SiaP and at least two Na<sup>+</sup> ions associate with SiaQM. **C)** The complex adopts the periplasm-open conformation and the substrate moves part way through the membrane. **D)** The complex resets to the cytoplasm-open conformation. **E)** SiaP is ligand-free and could still be associated with SiaQM in an open, closed or partially-open conformation. At this point, ligand-free SiaP, the substrate and the Na<sup>+</sup> ions dissociate from the complex in an unknown order.

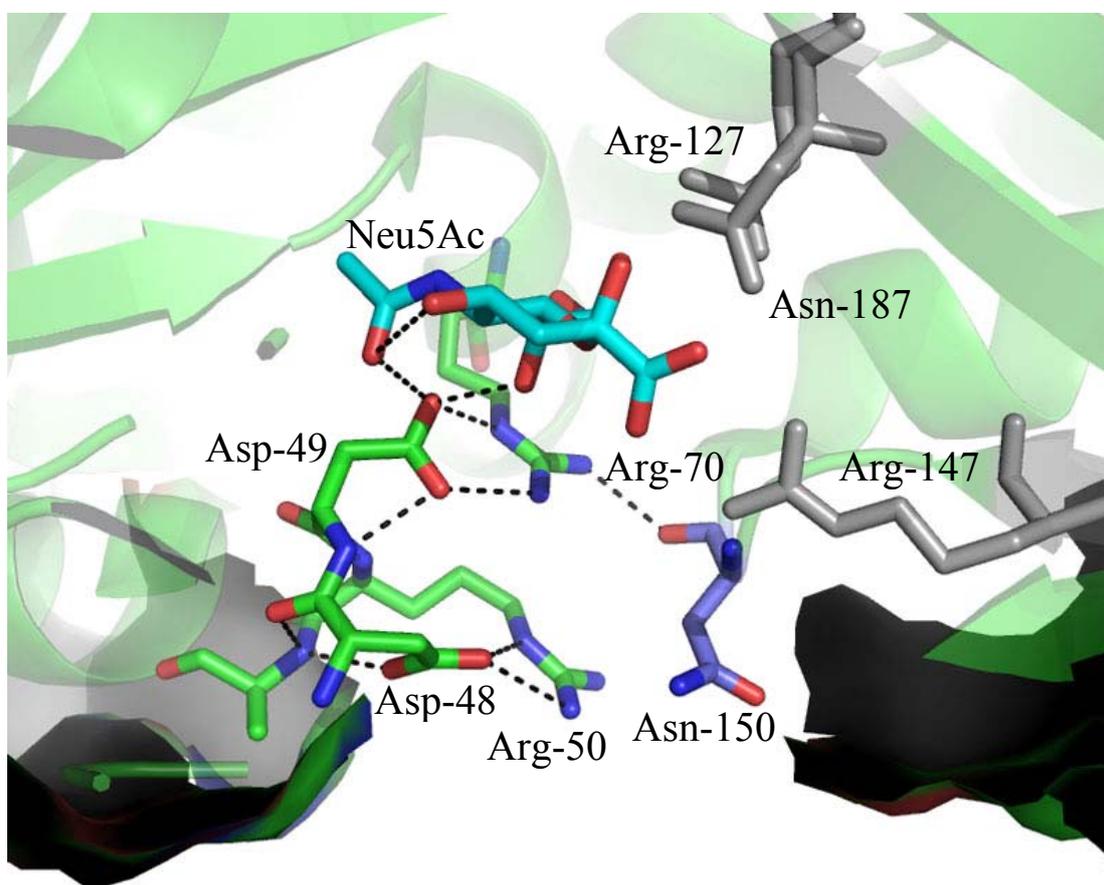
aromatic groups, the affinity is reduced ~10-fold and transport is maintained. Only the phenylalanine ring at this position combines high affinity binding with transport activity. The reduced ability of the tryptophan mutant to accumulate  $^{14}\text{C}$ -Neu5Ac to high levels while maintaining a similar initial rate could be due to the increased size of this residue disturbing its role as a gate in the translocation channel or in the regulation of the transport cycle in some way (Step D in Figure 7.4). It seems less likely that this effect is due to a disturbance of the SBP domain closure or the reduced affinity for the ligand, since the mutant SBP should be saturated and the initial rate of uptake was identical to SiaP-His<sub>6</sub>. Aromatic residues have been found to play key roles in secondary active transport, where they help form a transport-intermediate occluded state by forming part of periplasmic and cytoplasmic gates, for example LeuT and vSGLT, which use phenylalanine and tyrosine residues for this purpose (Faham *et al.*, 2008, Yamashita *et al.*, 2005).

In the group of surface mutations developed in Section 4.3.2, an Asn-150-Asp point mutation was made that was intended to form an aspartate-arginine pair (Table 4.4). In contrast to all other mutations made in the course of this project, this mutation did not have a deleterious effect on the affinity for Neu5Ac, despite its proximity to the apparently sensitive Ala-151 residue (Figure 4.11b). It was found that this N150D mutation abolished transport of Neu5Ac both *in vitro* (Figure 5.6) and *in vivo* (Figure 5.8). However, this could be restored *in vivo* by the expression *in trans* of *siaP-His<sub>6</sub>* (Figure 5.10) and was shown to compete with an equal concentration of SiaP-His<sub>6</sub> *in vitro* (Figure 5.9). This would indicate that this mutant is able to interact with SiaQM in a transient, non-productive manner and so is blocking Step C in the proposed transport scheme (Figure 7.4). ITC analysis of the ligand binding event revealed an increase in enthalpy and entropy between Neu5Ac-binding by SiaP-His<sub>6</sub> and SiaP-His<sub>6</sub>:N150D (Table 5.1). This could be due to the formation of hydrogen bonds with other residues or more likely a component of the bulk solvent, such as transient interactions with Na<sup>+</sup> ions.

From an examination of the protein crystal structure, Asn-150 appears to form a part of a hydrogen-bonded network with Asp-48, Arg-50 and Arg-70, shown in Figure 7.5. The extreme effect of this mutation would suggest that this position is critical for substrate transport and, although the precise mechanism is unclear at this point, it would appear to be interrupting Step C in the transport scheme (Figure 7.4). It could be that this is near to or forms part of the SiaPQM substrate translocation channel and so interferes with the passage of the substrate. This could also be a part of a ligand-sensing mechanism or a part of the opening-signal regulating the transport cycle. It is also worth mentioning that in many protein crystal structures of SiaP, it has been noted that there is a small, unidentifiable electron density connected to Asn-150 that is not present in the Asp-150 structure (Marcus Fischer, personal communication). If this unknown density is not an artefact, it could be that this associated structure is critical to the transport of sialic acid by SiaPQM.

The structure of the N150D mutant showed that there could be two distinct positions of the Asp-150. The first population of these is an identical orientation to Asn-150 in the Neu5Ac-bound SiaP structure. The second population comprises about 60% of the total and is oriented away from the network and extends outwards, away from the protein. This could fit with the increase in enthalpy and entropy since this exposed side chain could interact with more water molecules.

The extended side chain orientation of the Asp-150 could be similar to the structure of a dominant-negative MalE mutant that was determined by Shilton *et al.* (1996), where a W230R mutation results in this residue projecting from the surface into a region of MalE known to be critical for the interaction with MalFGK<sub>2</sub>. However, this is not truly analogous, since Trp-230 is a binding site residue and the W230R mutant causes a large decrease in ligand affinity. The protein crystal structure shows that this mutated residue interferes with the previously-determined protein-protein interaction region that is away from the Trp-230 position.



**Figure 7.5:** Representation of the residues involved in the Arg-50 – Asn-150 The hydrogen bonded network extending from the bound Neu5Ac (cyan carbon atoms) to the surface of SiaP (partial surface representation). The residues involved in the network from domain 1 are shown with green carbon atoms, while Asn-150 from domain 2 is shown with purple carbon atoms. Arg-127, Arg-147 and Asn-187 are shown in grey as a reference and a ribbon representation of SiaP is shown in green.

This N150D mutation has not explained the exact role of the Asn-150 residue in the transport cycle. This could be investigated by replacing Asn-150 with alanine to monitor the effect of the side chain on the transport of sialic acid. It would also be beneficial to investigate the remaining residues in the proposed network with such mutations as Arg-50 and Arg-70 to alanine, aspartate or lysine and Asp-48 and Asp-49 to alanine or asparagine. It would also be possible to demonstrate a physical interaction between the components of the transporter by attempting to cross-link SiaP:N150C and SiaQM with a hetero-bifunctional cross-linking reagent.

In fact, this has recently been achieved and has demonstrated a direct interaction between this SBP and the membrane domains (C. Mulligan, personal communication). This assay could now be used to determine the protein-protein interactions with the previous mutations from this work. It would be important to investigate the association of the Arg-147 and Phe-170 mutants, as well as the SiaP-His<sub>6</sub>:S15K;A195D double mutant, which was suspected of altering this association.

### **7.3 Utilisation of sialic acids by bacteria**

As discussed in Section 1.6.1, bacteria can use sialic acids as a source of carbon, nitrogen and energy. Their transport can be mediated by ABC, secondary active and SBP-dependent secondary transporters. Once inside the cell, Neu5Ac is catabolised in *E. coli* by the *nan* and *nag* operon products to produce fructose-6-phosphate, which enters the glycolysis pathway (Figure 1.14).

As stated in Section 6.2, the use of Neu5Ac by *E. coli* as the sole carbon source has been characterised previously, while the use of Neu5Gc had been reported but not characterised in any way (Vimr & Troy, 1985) and the use of KDN had not been examined in *E. coli*. It was found that all of these three sialic acids were capable of supporting growth (Figure 6.4a). Neu5Ac allowed growth to occur soonest, closely followed by KDN and then Neu5Gc. These differences were shown to be due to the induction effects of these different sialic acids, since pre-induction with Neu5Ac resulted in identical growth for all substrates (Figure 6.4b). Growth on Neu5Gc was found to follow the sialometabolic pathway as Neu5Ac, but with release of glycolate instead of

acetate by NagA, as expected (Figures 6.7 and 6.8). For growth on KDN, the first *nan* gene, *nanA*, was the only sialometabolic gene found to be essential for growth (releasing pyruvate and D-mannose), while deletion of *nanK* and *nanE* were found to cause a long lag phase before growth (Figure 6.9).

Unfortunately, the inducer for the *nan* operon is uncertain due to reports that cast suspicion and doubt upon Neu5Ac, ManNAc and ManNAc-6P<sub>i</sub> (Johnston *et al.*, 2007, Kalivoda *et al.*, 2003), but from the growth of *E. coli* BW25113 it is clear that Neu5Ac or its products are better than KDN or its products, which are better than Neu5Gc or its products. This result for KDN is surprising if it is a downstream product that is the inducer since, for KDN, this is mannose. Mannose is not normally present in the cytoplasm as it is transported via a PTS transporter to give mannose-6-phosphate (Man6P) and if present would be acted upon by a promiscuous kinase activity to give Man6P. With the much longer lag phase in Neu5Gc-growth, it would seem that the presence of the extra hydroxyl group is detrimental to intracellular Neu5Gc-induction and so binding of this involves accurate recognition of the C<sub>5</sub> position, but this is not critical for the enzymatic action of the *nan* and *nag* gene products.

The *nanK* and *nanE*-deletion effects on KDN-dependent growth would be due to their capacity to interact with mannose and Man6P, respectively, producing Man6P and possibly glucose-6-phosphate (Figure 6.9) (Ferrero *et al.*, 2007, Larion *et al.*, 2007). The effects of these enzymes would be to reduce the growth-inhibition effects of high intracellular concentrations of mannose and Man6P. The reduction but not avoidance of this lag phase following pre-induction could be due to poorer expression or induction of the remaining *nan* genes from the deletion strains or that NanK and NanE are playing an important role under these conditions. This would suggest that NanK can supplement promiscuous hexose kinase activity and NanE reduces the inhibitory effect of Man6P build up, possibly by producing glucose-6-phosphate that would enter the pentose phosphate pathway.

Pre-induction with Neu5Ac causes a reduction in the lag time experienced by both of these deletion strains, presumably since Neu5Ac or ManNAc are the inducers. Combined with the growth of *E. coli* BW25113 on KDN, above, it would seem that sialic acids are the *nan* inducers.

To investigate these, the effect of complete induction could be investigated using a deletion of the *nan* regulator, *nanR*. This would require the construction of a series of double deletions by bacteriophage P1 transduction of *nanR::kan* into a clean deletion (excision of the kanamycin cassette) of each of the sialometabolic genes.

An alternative explanation for the growth defect was that the presence of the kanamycin cassette had a detrimental effect on the expression of *yhcH*, the final gene in the *nan* operon. This could be investigated by examining the growth of *E. coli* BW25113 *yhcH::kan* or by using a clean deletion of *nanK* and *nanE*, so removing the potential polar effects of this cassette.

The sialic acid transporters NanT (MFS), STM1128 (SSS) and SiaPQM (TRAP-T) are from three different families of secondary transporters; the major difference between these is that only SiaPQM is SBP-dependent. Like SiaPQM, STM1128 is sodium ion gradient-driven, whereas NanT is proton-driven. All of these transporters restored growth on each of the three different sialic acids at different rates (Table 6.5). A clear difference can be seen with SiaPQM, where growth is consistently the slowest. This is likely due to the required association between the membrane protein and the binding protein. The rate of growth allowed by STM1128 is very similar for each carbon source, while NanT allowed the fastest growth on Neu5Gc rather than Neu5Ac. This would seem to indicate that there are differences in the ligand binding mechanism of these two proteins, where there are presumably differences in the co-ordination of the ligand *N*-acetyl group where all of these transporters are sensitive to its absence and NanT is more accepting of its increase in size. How this would relate structurally to the *N*-acetyl group co-ordination mechanism by SiaP cannot be determined at this point.

That these very different transporters from different organisms can transport these different sialic acid analogues is an important finding with widespread implications. The capability to transport these three different sialic acids suggests that they can all be used by the source bacteria, *E. coli*, *H. influenzae* and *S. typhimuriumi*. These pathogens all contain the *nan* genes, *nanA*, *nanK* and *nanE*, which can catabolise these sialic acids. It has recently been shown that another *nan* gene, *nanS* (*yjhS*), can metabolise 9-O-acetyl-*N*-acetylneuraminic acid to Neu5Ac (Steenbergen *et al.*, 2009). In their examination of nearly 2000 bacterial genomes, Almagro-Moreno & Boyd (2009) found that the *nan* genes were mostly restricted to pathogenic and commensal  $\gamma$ -proteobacteria, *Fusobacterium* and Firmicutes, and they demonstrated the ability to utilise Neu5Ac by several *Salmonella*, *Vibrio* and *Yersinia* species. These results start to suggest that *nan*-containing species, which are mostly pathogens, could be exposed to and using a wide variety of sialic acids.

## References

- Abramson, J., I. Smirnova, V. Kasho, G. Verner, S. Iwata & H. R. Kaback, (2003a) The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport. *FEBS Lett* **555**: 96-101.
- Abramson, J., I. Smirnova, V. Kasho, G. Verner, H. R. Kaback & S. Iwata, (2003b) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**: 610-615.
- Adler, J., G. L. Hazelbauer & M. M. Dahl, (1973) Chemotaxis toward sugars in *Escherichia coli*. *J Bacteriol* **115**: 824-847.
- Agusti, R., M. E. Giorgi, V. M. Mendoza, C. Gallo-Rodriguez & R. M. de Lederkremer, (2007) Comparative rates of sialylation by recombinant trans-sialidase and inhibitor properties of synthetic oligosaccharides from *Trypanosoma cruzi* mucins-containing galactofuranose and galactopyranose. *Bioorg Med Chem*.
- Akiyama, N., K. Takeda & K. Miki, (2009) Crystal structure of a periplasmic substrate-binding protein in complex with calcium lactate. *J Mol Biol* **392**: 559-565.
- Almagro-Moreno, S. & E. F. Boyd, (2009) Insights into the evolution of sialic acid catabolism among bacteria. *BMC Evol Biol* **9**: 118.
- Ames, G. F. & J. Lever, (1970) Components of histidine transport: histidine-binding proteins and hisP protein. *Proc Natl Acad Sci U S A* **66**: 1096-1103.
- 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.
- Anandakrishnan, R. & A. Onufriev, (2008) Analysis of basic clustering algorithms for numerical estimation of statistical averages in biomolecules. *J Comput Biol* **15**: 165-184.
- Andreescu, S. & L. A. Luck, (2008) Studies of the binding and signaling of surface-immobilized periplasmic glucose receptors on gold nanoparticles: a glucose biosensor application. *Anal Biochem* **375**: 282-290.
- Angata, T., E. H. Margulies, E. D. Green & A. Varki, (2004) Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms. *Proc Natl Acad Sci U S A* **101**: 13251-13256.
- Angata, T. & A. Varki, (2002) Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev* **102**: 439-469.
- Aubrey, R. & C. Tang, (2003) The pathogenesis of disease due to type b *Haemophilus influenzae*. *Methods Mol Med* **71**: 29-50.
- 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.
- Benson, D. E., D. W. Conrad, R. M. de Lorimier, S. A. Trammell & H. W. Hellinga, (2001) Design of bioelectronic interfaces by exploiting hinge-bending motions in proteins. *Science* **293**: 1641-1644.
- Bermejo, G. A., M. P. Strub, C. Ho & N. Tjandra, Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein. *Biochemistry* **49**: 1893-1902.
- Berntsson, R. P., M. K. Doeven, F. Fusetti, R. H. Duurkens, D. Sengupta, S. J. Marrink, A. M. Thunnissen, B. Poolman & D. J. Slotboom, (2009) The structural basis for peptide selection by the transport receptor OppA. *EMBO J* **28**: 1332-1340.
- 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*.
- Bohm, G., R. Muhr & R. Jaenicke, (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng* **5**: 191-195.
- 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.
- 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.
- Buhrman, G., V. de Serrano & C. Mattos, (2003) Organic solvents order the dynamic switch II in Ras crystals. *Structure* **11**: 747-751.
- Carlin, A. F., S. Uchiyama, Y. C. Chang, A. L. Lewis, V. Nizet & A. Varki, (2009) Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. *Blood* **113**: 3333-3336.
- 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.
- CFG, (2010) Functional Glycomics Gateway [www.functionalglycomics.org/static/consortium/resources/resourcecoreh.shtml](http://www.functionalglycomics.org/static/consortium/resources/resourcecoreh.shtml).
- Chen, C., A. A. Malek, M. J. Wargo, D. A. Hogan & G. A. Beattie, The ATP-binding cassette transporter Cbc (choline/betaine/carnitine) recruits multiple substrate-binding proteins with strong specificity for distinct quaternary ammonium compounds. *Mol Microbiol* **75**: 29-45.
- Chen, J., G. Lu, J. Lin, A. L. Davidson & F. A. Quioco, (2003) A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol Cell* **12**: 651-661.

- Chervenak, M. C. & E. J. Toone, (1995) Calorimetric analysis of the binding of lectins with overlapping carbohydrate-binding ligand specificities. *Biochemistry* **34**: 5685-5695.
- Comb, D. G. & S. Roseman, (1960) The sialic acids. I. The structure and enzymatic synthesis of N-acetylneuraminic acid. *J Biol Chem* **235**: 2529-2537.
- Condemine, G., C. Berrier, J. Plumbridge & A. Ghazi, (2005) Function and expression of an N-acetylneuraminic acid-inducible outer membrane channel in *Escherichia coli*. *J Bacteriol* **187**: 1959-1965.
- Corfield, A. P., S. A. Wagner, J. R. Clamp, M. S. Kriaris & L. C. Hoskins, (1992) Mucin degradation in the human colon: production of sialidase, sialate O-acetyltransferase, N-acetylneuraminidase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. *Infect Immun* **60**: 3971-3978.
- Corfield, T., (1992) Bacterial sialidases--roles in pathogenicity and nutrition. *Glycobiology* **2**: 509-521.
- Cuneo, M. J., A. Changela, A. E. Miklos, L. S. Beese, J. K. Krueger & H. W. Hellinga, (2008) Structural analysis of a periplasmic binding protein in the tripartite ATP-independent transporter family reveals a tetrameric assembly that may have a role in ligand transport. *J Biol Chem* **283**: 32812-32820.
- Datsenko, K. A. & B. L. Wanner, (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.
- Daus, M. L., S. Berendt, S. Wuttge & E. Schneider, (2007a) 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.
- Daus, M. L., M. Grote, P. Muller, M. Doebber, A. Herrmann, H. J. Steinhoff, E. Dassa & E. Schneider, (2007b) ATP-driven MalK dimer closure and reopening and conformational changes of the "EAA" motifs are crucial for function of the maltose ATP-binding cassette transporter (MalFGK2). *J Biol Chem* **282**: 22387-22396.
- Davidson, A. L. & J. Chen, (2004) ATP-binding cassette transporters in bacteria. *Annu Rev Biochem* **73**: 241-268.
- 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., 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.
- 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. M. C., (1986) *Data for biochemical research*, p. xii,580p. Clarendon, Oxford.

- de Lorimier, R. M., Y. Tian & H. W. Hellinga, (2006) Binding and signaling of surface-immobilized reagentless fluorescent biosensors derived from periplasmic binding proteins. *Protein Sci* **15**: 1936-1944.
- Dean, D. A., A. L. Davidson & H. Nikaido, (1989a) Maltose transport in membrane vesicles of *Escherichia coli* is linked to ATP hydrolysis. *Proc Natl Acad Sci U S A* **86**: 9134-9138.
- Dean, D. A., J. D. Fikes, K. Gehring, P. J. Bassford, Jr. & H. Nikaido, (1989b) Active transport of maltose in membrane vesicles obtained from *Escherichia coli* cells producing tethered maltose-binding protein. *J Bacteriol* **171**: 503-510.
- Della Gatta, G., G. Barone & V. Elia, (1986) Enthalpies of solvation for N-alkylamides in water and in carbon tetrachloride at 25°C. *Journal of Solution Chemistry* **15**: 157-167.
- Deshpande, A., S. Nimsadkar & S. C. Mande, (2005) Effect of alcohols on protein hydration: crystallographic analysis of hen egg-white lysozyme in the presence of alcohols. *Acta Crystallogr D Biol Crystallogr* **61**: 1005-1008.
- Deuschle, K., S. Okumoto, M. Fehr, L. L. Looger, L. Kozhukh & W. B. Frommer, (2005) Construction and optimization of a family of genetically encoded metabolite sensors by semirational protein engineering. *Protein Sci* **14**: 2304-2314.
- 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.
- Duan, X., J. A. Hall, H. Nikaido & F. A. Quioco, (2001) Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. *J Mol Biol* **306**: 1115-1126.
- Duan, X. & F. A. Quioco, (2002) Structural evidence for a dominant role of nonpolar interactions in the binding of a transport/chemosensory receptor to its highly polar ligands. *Biochemistry* **41**: 706-712.
- Faham, S., A. Watanabe, G. M. Besserer, D. Cascio, A. Specht, B. A. Hirayama, E. M. Wright & J. Abramson, (2008) The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na<sup>+</sup>/sugar symport. *Science* **321**: 810-814.
- Fehr, M., W. B. Frommer & S. Lalonde, (2002) Visualization of maltose uptake in living yeast cells by fluorescent nanosensors. *Proc Natl Acad Sci U S A* **99**: 9846-9851.
- Ferenci, T., (1980) The recognition of maltodextrins by *Escherichia coli*. *Eur J Biochem* **108**: 631-636.
- Ferrero, M. A., H. Martinez-Blanco, F. F. Lopez-Velasco, C. Ezquerro-Saenz, N. Navasa, S. Lozano & L. B. Rodriguez-Aparicio, (2007) Purification and characterization of GlcNAc-6-P 2-epimerase from *Escherichia coli* K92. *Acta Biochim Pol* **54**: 387-399.
- 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.
- 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.
- 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.
- Gerber, S., M. Comellas-Bigler, B. A. Goetz & K. P. Locher, (2008) Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science* **321**: 246-250.
- 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.
- Gordon, J. C., J. B. Myers, T. Folta, V. Shoja, L. S. Heath & A. Onufriev, (2005) H<sup>++</sup>: a server for estimating pK<sub>a</sub>s and adding missing hydrogens to macromolecules. *Nucleic Acids Res* **33**: W368-371.
- Gould, A. D. & B. H. Shilton, Studies of the maltose transport system reveal a mechanism for coupling ATP hydrolysis to substrate translocation without direct recognition of substrate. *J Biol Chem*.
- Gould, A. D., P. G. Telmer & B. H. Shilton, (2009) Stimulation of the maltose transporter ATPase by unliganded maltose binding protein. *Biochemistry* **48**: 8051-8061.
- Graziano, G., (2001) Hydration Thermodynamics of N-Methylacetamide. *Journal of the Physical Society of Japan* **70**: 2234.
- Gu, H., S. Lalonde, S. Okumoto, L. L. Looger, A. M. Scharff-Poulsen, A. R. Grossman, J. Kossmann, I. Jakobsen & W. B. Frommer, (2006) A novel analytical method for in vivo phosphate tracking. *FEBS Lett* **580**: 5885-5893.
- 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.
- Guntas, G., S. F. Mitchell & M. Ostermeier, (2004) A molecular switch created by in vitro recombination of nonhomologous genes. *Chem Biol* **11**: 1483-1487.
- 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. & Y. Kakinuma, (1985) Primary and secondary transport of cations in bacteria. *Ann N Y Acad Sci* **456**: 375-383.
- Hazelbauer, G. L. & J. Adler, (1971) Role of the galactose binding protein in chemotaxis of *Escherichia coli* toward galactose. *Nat New Biol* **230**: 101-104.
- Hofnung, M., (1974) Divergent operons and the genetic structure of the maltose B region in *Escherichia coli* K12. *Genetics* **76**: 169-184.
- Hofnung, M., D. Hatfield & M. Schwartz, (1974) malB region in *Escherichia coli* K-12: characterization of new mutations. *J Bacteriol* **117**: 40-47.
- Hollenstein, K., D. C. Frei & K. P. Locher, (2007) Structure of an ABC transporter in complex with its binding protein. *Nature* **446**: 213-216.
- 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.
- Hor, L. I. & H. A. Shuman, (1993) Genetic analysis of periplasmic binding protein dependent transport in *Escherichia coli*. Each lobe of maltose-binding protein interacts with a different subunit of the MalFGK2 membrane transport complex. *J Mol Biol* **233**: 659-670.
- 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.
- Huvent, I., H. Belrhali, R. Antoine, C. Bompard, C. Locht, F. Jacob-Dubuisson & V. Villeret, (2006) 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.
- 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.
- 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.
- Jacobson, B. L., J. J. He, P. S. Vermersch, D. D. Lemon & F. A. Quioco, (1991) Engineered interdomain disulfide in the periplasmic receptor for sulfate transport reduces flexibility. Site-directed mutagenesis and ligand-binding studies. *J Biol Chem* **266**: 5220-5225.
- Jacso, T., M. Grote, M. L. Daus, P. Schmieder, S. Keller, E. Schneider & B. Reif, (2009) Periplasmic loop P2 of the MalF subunit of the maltose ATP binding cassette transporter is sufficient to bind the maltose binding protein MalE. *Biochemistry* **48**: 2216-2225.
- 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.
- Johnston, J. W., A. Zaleski, S. Allen, J. M. Mootz, D. Armbruster, B. W. Gibson, M. A. Apicella & R. S. Munson, Jr., (2007) Regulation of sialic acid transport and catabolism in *Haemophilus influenzae*. *Mol Microbiol* **66**: 26-39.
- Jones, P. M., M. L. O'Mara & A. M. George, (2009) ABC transporters: a riddle wrapped in a mystery inside an enigma. *Trends Biochem Sci* **34**: 520-531.
- Jung, H., (2001) Towards the molecular mechanism of Na<sup>(+)</sup>/solute symport in prokaryotes. *Biochim Biophys Acta* **1505**: 131-143.
- Kaback, H. R., M. Sahin-Toth & A. B. Weinglass, (2001) The kamikaze approach to membrane transport. *Nat Rev Mol Cell Biol* **2**: 610-620.

- Kadaba, N. S., J. T. Kaiser, E. Johnson, A. Lee & D. C. Rees, (2008) The high-affinity E. coli methionine ABC transporter: structure and allosteric regulation. *Science* **321**: 250-253.
- Kalivoda, K. A., S. M. Steenbergen, E. R. Vimr & J. Plumbridge, (2003) Regulation of sialic acid catabolism by the DNA binding protein NanR in Escherichia coli. *J Bacteriol* **185**: 4806-4815.
- Kalovidouris, S. A., O. Blixt, A. Nelson, S. Vidal, W. B. Turnbull, J. C. Paulson & J. F. Stoddart, (2003) Chemically defined sialoside scaffolds for investigation of multivalent interactions with sialic acid binding proteins. *J Org Chem* **68**: 8485-8493.
- Kang, Y. K., K. D. Gibson, G. Nemethy & H. A. Scheraga, (1988) Free energies of hydration of solute molecules. 4. Revised treatment of the hydration shell model. *The Journal of Physical Chemistry* **92**: 4739-4742.
- Karlsson, K. A., (1998) Meaning and therapeutic potential of microbial recognition of host glycoconjugates. *Mol Microbiol* **29**: 1-11.
- Karpowich, N. K., H. H. Huang, P. C. Smith & J. F. Hunt, (2003) Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. *J Biol Chem* **278**: 8429-8434.
- Kay, W. W. & M. Cameron, (1978) Citrate transport in Salmonella typhimurium. *Arch Biochem Biophys* **190**: 270-280.
- 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.
- Kenoth, R., S. S. Komath & M. J. Swamy, (2003) Physicochemical and saccharide-binding studies on the galactose-specific seed lectin from Trichosanthes cucumerina. *Arch Biochem Biophys* **413**: 131-138.
- Khare, D., M. L. Oldham, C. Orelle, A. L. Davidson & J. Chen, (2009) Alternating access in maltose transporter mediated by rigid-body rotations. *Mol Cell* **33**: 528-536.
- Kilhoffer, M. C., M. Kubina, F. Travers & J. Haiech, (1992) Use of engineered proteins with internal tryptophan reporter groups and perturbation techniques to probe the mechanism of ligand-protein interactions: investigation of the mechanism of calcium binding to calmodulin. *Biochemistry* **31**: 8098-8106.
- Kirsch, R. D. & E. Joly, (1998) An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. *Nucleic Acids Res* **26**: 1848-1850.
- Knubovets, T., J. J. Osterhout & A. M. Klibanov, (1999) Structure of lysozyme dissolved in neat organic solvents as assessed by NMR and CD spectroscopies. *Biotechnol Bioeng* **63**: 242-248.
- Kornberg, H. L., L. T. Lambourne & A. A. Sproul, (2000) Facilitated diffusion of fructose via the phosphoenolpyruvate/glucose phosphotransferase system of Escherichia coli. *Proc Natl Acad Sci US A* **97**: 1808-1812.

- Larion, M., L. B. Moore, S. M. Thompson & B. G. Miller, (2007) Divergent evolution of function in the ROK sugar kinase superfamily: role of enzyme loops in substrate specificity. *Biochemistry* **46**: 13564-13572.
- Lecher, J., M. Pittelkow, S. Zobel, J. Bursy, T. Bonig, S. H. Smits, L. Schmitt & E. Bremer, (2009) The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. *J Mol Biol* **389**: 58-73.
- Lewinson, O., A. T. Lee, K. P. Locher & D. C. Rees, A distinct mechanism for the ABC transporter BtuCD-BtuF revealed by the dynamics of complex formation. *Nat Struct Mol Biol* **17**: 332-338.
- Lewis, A. L., N. Desa, E. E. Hansen, Y. A. Knirel, J. I. Gordon, P. Gagneux, V. Nizet & A. Varki, (2009) Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. *Proc Natl Acad Sci U S A* **106**: 13552-13557.
- Locher, K. P., R. B. Bass & D. C. Rees, (2003) Structural biology. Breaching the barrier. *Science* **301**: 603-604.
- Lu, G., J. M. Westbrook, 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.
- Makhatadze, G. I. & P. L. Privalov, (1993) Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration. *J Mol Biol* **232**: 639-659.
- Malakhov, M. P., L. M. Aschenbrenner, D. F. Smee, M. K. Wandersee, R. W. Sidwell, L. V. Gubareva, V. P. Mishin, F. G. Hayden, D. H. Kim, A. Ing, E. R. Campbell, M. Yu & F. Fang, (2006) Sialidase fusion protein as a novel broad-spectrum inhibitor of influenza virus infection. *Antimicrob Agents Chemother* **50**: 1470-1479.
- Mandrell, R. E., R. McLaughlin, Y. Aba Kwaik, A. Lesse, R. Yamasaki, B. Gibson, S. M. Spinola & M. A. Apicella, (1992) Lipooligosaccharides (LOS) of some *Haemophilus* species mimic human glycosphingolipids, and some LOS are sialylated. *Infect Immun* **60**: 1322-1328.
- Manson, M. D. & M. Kossmann, (1986) Mutations in tar suppress defects in maltose chemotaxis caused by specific maleE mutations. *J Bacteriol* **165**: 34-40.
- Mao, B., M. R. Pear, J. A. McCammon & F. A. Quiocho, (1982) Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J Biol Chem* **257**: 1131-1133.
- Marvin, J. S. & H. W. Hellinga, (2001) Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nat Struct Biol* **8**: 795-798.
- Millet, O., R. P. Hudson & L. E. Kay, (2003) The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. *Proc Natl Acad Sci U S A* **100**: 12700-12705.
- Morschhauser, J., H. Hoschutzky, K. Jann & J. Hacker, (1990) Functional analysis of the sialic acid-binding adhesin SfaS of pathogenic *Escherichia coli* by site-specific mutagenesis. *Infect Immun* **58**: 2133-2138.
- Moussatova, A., C. Kandt, M. L. O'Mara & D. P. Tieleman, (2008) ATP-binding cassette transporters in *Escherichia coli*. *Biochim Biophys Acta* **1778**: 1757-1771.
- Mowbray, S. L. & A. J. Bjorkman, (1999) Conformational changes of ribose-binding protein and two related repressors are tailored to fit the functional need. *J Mol Biol* **294**: 487-499.

- 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* **281**: 22212-22222.
- Mulligan, C., M. Fischer & G. H. Thomas, (2010) Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev.*
- Mulligan, C., E. R. Geertsma, E. Severi, D. J. Kelly, B. Poolman & G. H. Thomas, (2009) The substrate-binding protein imposes directionality on an electrochemical sodium gradient-driven TRAP transporter. *Proc Natl Acad Sci U S A* **106**: 1778-1783.
- Mulligan, C., D. J. Kelly & G. H. Thomas, (2007) Tripartite ATP-independent periplasmic transporters: application of a relational database for genome-wide analysis of transporter gene frequency and organization. *J Mol Microbiol Biotechnol* **12**: 218-226.
- Neidhardt, F. C., P. L. Bloch & D. F. Smith, (1974) Culture medium for enterobacteria. *J Bacteriol* **119**: 736-747.
- Newcomer, M. E., B. A. Lewis & F. A. Quiocho, (1981) The radius of gyration of L-arabinose-binding protein decreases upon binding of ligand. *J Biol Chem* **256**: 13218-13222.
- Nikaïdo, H., (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**: 593-656.
- Ogasawara, Y., T. Namai, F. Yoshino, M. C. Lee & K. Ishii, (2007) Sialic acid is an essential moiety of mucin as a hydroxyl radical scavenger. *FEBS Lett* **581**: 2473-2477.
- Okada, S., K. Ota & T. Ito, (2009) Circular permutation of ligand-binding module improves dynamic range of genetically encoded FRET-based nanosensor. *Protein Sci* **18**: 2518-2527.
- Oldham, M. L., D. Khare, F. A. Quiocho, A. L. Davidson & J. Chen, (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**: 515-521.
- Ooi, T., M. Oobatake, G. Nemethy & H. A. Scheraga, (1987) Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proc Natl Acad Sci U S A* **84**: 3086-3090.
- Orelle, C., T. Ayvaz, R. M. Everly, C. S. Klug & A. L. Davidson, (2008) Both maltose-binding protein and ATP are required for nucleotide-binding domain closure in the intact maltose ABC transporter. *Proc Natl Acad Sci U S A* **105**: 12837-12842.
- Pao, S. S., I. T. Paulsen & M. H. Saier, Jr., (1998) Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**: 1-34.
- Patzlaff, J. S., T. van der Heide & B. Poolman, (2003) The ATP/substrate stoichiometry of the ATP-binding cassette (ABC) transporter OpuA. *J Biol Chem* **278**: 29546-29551.
- Pellicer, M. T., J. Badia, J. Aguilar & L. Baldoma, (1996) *glc* locus of *Escherichia coli*: characterization of genes encoding the subunits of glycolate oxidase and the *glc* regulator protein. *J Bacteriol* **178**: 2051-2059.

- 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.
- Prakash, S., G. Cooper, S. Singhi & M. H. Saier, Jr., (2003) The ion transporter superfamily. *Biochim Biophys Acta* **1618**: 79-92.
- Privalov, P. L. & G. I. Makhatadze, (1993) Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J Mol Biol* **232**: 660-679.
- Prossnitz, E., A. Gee & G. F. Ames, (1989) Reconstitution of the histidine periplasmic transport system in membrane vesicles. Energy coupling and interaction between the binding protein and the membrane complex. *J Biol Chem* **264**: 5006-5014.
- Quioco, F. A., (1990) Atomic structures of periplasmic binding proteins and the high-affinity active transport systems in bacteria. *Philos Trans R Soc Lond B Biol Sci* **326**: 341-351; discussion 351-342.
- Quioco, F. A. & P. S. Ledvina, (1996) Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol Microbiol* **20**: 17-25.
- 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.
- Rao, V. K., G. P. Krasan, D. R. Hendrixson, S. Dawid & J. W. St Geme, 3rd, (1999) Molecular determinants of the pathogenesis of disease due to non-typable *Haemophilus influenzae*. *FEMS Microbiol Rev* **23**: 99-129.
- Rech, S., C. Wolin & R. P. Gunsalus, (1996) Properties of the periplasmic ModA molybdate-binding protein of *Escherichia coli*. *J Biol Chem* **271**: 2557-2562.
- Rizk, S. S., M. J. Cuneo & H. W. Hellinga, (2006) Identification of cognate ligands for the *Escherichia coli* phnD protein product and engineering of a reagentless fluorescent biosensor for phosphonates. *Protein Sci* **15**: 1745-1751.
- Roy, S., C. W. Douglas & G. P. Stafford, A novel sialic acid utilization and uptake system in the periodontal pathogen *Tannerella forsythia*. *J Bacteriol* **192**: 2285-2293.
- 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., (2000a) Families of transmembrane sugar transport proteins. *Mol Microbiol* **35**: 699-710.
- Saier, M. H., Jr., (2000b) A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**: 354-411.
- Saier, M. H., Jr., J. T. Beatty, A. Goffeau, K. T. Harley, W. H. Heijne, S. C. Huang, D. L. Jack, P. S. Jahn, K. Lew, J. Liu, S. S. Pao, I. T. Paulsen, T. T. Tseng & P. S. Virk, (1999) The major facilitator superfamily. *J Mol Microbiol Biotechnol* **1**: 257-279.
- Severi, E., D. W. Hood & G. H. Thomas, (2007) Sialic acid utilization by bacterial pathogens. *Microbiology* **153**: 2817-2822.
- Severi, E., A. H. Hosie, J. A. Hawkhead & G. H. Thomas, Characterization of a novel sialic acid transporter of the sodium solute symporter (SSS) family and in vivo

- comparison with known bacterial sialic acid transporters. *FEMS Microbiol Lett* **304**: 47-54.
- 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.
- Sharff, A. J., L. E. Rodseth, J. C. Spurlino & F. A. Quioco, (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.
- Shepherd, M., M. D. Heath & R. K. Poole, (2007) NikA binds heme: a new role for an *Escherichia coli* periplasmic nickel-binding protein. *Biochemistry* **46**: 5030-5037.
- Shilton, B. H., (2008) The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems. *Biochim Biophys Acta* **1778**: 1772-1780.
- Shilton, B. H., H. A. Shuman & S. L. Mowbray, (1996) Crystal structures and solution conformations of a dominant-negative mutant of *Escherichia coli* maltose-binding protein. *J Mol Biol* **264**: 364-376.
- Singer, S. J. & G. L. Nicolson, (1972) The fluid mosaic model of the structure of cell membranes. *Science* **175**: 720-731.
- Singh, S. K., C. L. Piscitelli, A. Yamashita & E. Gouaux, (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322**: 1655-1661.
- Sjoelund, V. & I. A. Kaltashov, (2007) Transporter-to-trap conversion: a disulfide bond formation in cellular retinoic acid binding protein I mutant triggered by retinoic acid binding irreversibly locks the ligand inside the protein. *Biochemistry* **46**: 13382-13390.
- Sobczak, I. & J. S. Lolkema, (2005) Structural and mechanistic diversity of secondary transporters. *Curr Opin Microbiol* **8**: 161-167.
- Sobolevsky, A. I., M. P. Rosconi & E. Gouaux, (2009) X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature* **462**: 745-756.
- Spurlino, J. C., G. Y. Lu & F. A. Quioco, (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.
- Stahl, C. L. & G. A. Sojka, (1973) Growth of *Rhodospseudomonas capsulata* on L- and D-malic acid. *Biochim Biophys Acta* **297**: 241-245.
- Steenbergen, S. M., J. L. Jirik & E. R. Vimr, (2009) YjhS (NanS) is required for *Escherichia coli* to grow on 9-O-acetylated N-acetylneuraminic acid. *J Bacteriol* **191**: 7134-7139.
- Steim, J. M., M. E. Tourtellotte, J. C. Reinert, R. N. McElhaney & R. L. Rader, (1969) Calorimetric evidence for the liquid-crystalline state of lipids in a biomembrane. *Proc Natl Acad Sci U S A* **63**: 104-109.
- Stockner, T., H. J. Vogel & D. P. Tieleman, (2005) A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys J* **89**: 3362-3371.

- 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.
- Tam, R. & M. H. Saier, Jr., (1993) A bacterial periplasmic receptor homologue with catalytic activity: cyclohexadienyl dehydratase of *Pseudomonas aeruginosa* is homologous to receptors specific for polar amino acids. *Res Microbiol* **144**: 165-169.
- Tanford, C., (1983) Translocation pathway in the catalysis of active transport. *Proc Natl Acad Sci U S A* **80**: 3701-3705.
- Tang, C., C. D. Schwieters & G. M. Clore, (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **449**: 1078-1082.
- Telmer, P. G. & B. H. Shilton, (2003) Insights into the conformational equilibria of maltose-binding protein by analysis of high affinity mutants. *J Biol Chem* **278**: 34555-34567.
- Thomas, K. J., D. B. Sherman, T. J. Amiss, S. A. Andaluz & J. B. Pitner, (2006) A long-wavelength fluorescent glucose biosensor based on bioconjugates of galactose/glucose binding protein and Nile Red derivatives. *Diabetes Technol Ther* **8**: 261-268.
- Thompson, R. B. & J. R. Lakowicz, (1984) Binding of methylumbelliferyl mannoside to concanavalin A under high pressure. *Biochim Biophys Acta* **790**: 87-90.
- 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.
- van der Heide, T. & B. Poolman, (2002) ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep* **3**: 938-943.
- van der Merwe, P. A., P. R. Crocker, M. Vinson, A. N. Barclay, R. Schauer & S. Kelm, (1996) Localization of the putative sialic acid-binding site on the immunoglobulin superfamily cell-surface molecule CD22. *J Biol Chem* **271**: 9273-9280.
- Vimr, E., C. Lichtensteiger & S. Steenbergen, (2000) Sialic acid metabolism's dual function in *Haemophilus influenzae*. *Mol Microbiol* **36**: 1113-1123.
- 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, (1985) Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli*. *J Bacteriol* **164**: 845-853.
- Vinson, M., P. A. van der Merwe, S. Kelm, A. May, E. Y. Jones & P. R. Crocker, (1996) Characterization of the sialic acid-binding site in sialoadhesin by site-directed mutagenesis. *J Biol Chem* **271**: 9267-9272.
- Waldron, K. J., S. Tottey, S. Yanagisawa, C. Dennison & N. J. Robinson, (2007) A periplasmic iron-binding protein contributes toward inward copper supply. *J Biol Chem* **282**: 3837-3846.
- 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.
- Wandersman, C., M. Schwartz & T. Ferenci, (1979) *Escherichia coli* mutants impaired in maltodextrin transport. *J Bacteriol* **140**: 1-13.

- Weber, J., S. Wilke-Mounts, S. T. Hammond & A. E. Senior, (1998) Tryptophan substitutions surrounding the nucleotide in catalytic sites of F1-ATPase. *Biochemistry* **37**: 12042-12050.
- 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.
- Wilkins, M. H., A. E. Blaurock & D. M. Engelman, (1971) Bilayer structure in membranes. *Nat New Biol* **230**: 72-76.
- Wong, P. T., E. R. Kashket & T. H. Wilson, (1970) Energy coupling in the lactose transport system of Escherichia coli. *Proc Natl Acad Sci U S A* **65**: 63-69.
- Wong, P. T. & T. H. Wilson, (1970) Counterflow of galactosides in Escherichia coli. *Biochim Biophys Acta* **196**: 336-350.
- Xiang, S. L., M. Zhong, F. C. Cai, B. Deng & X. P. Zhang, (2006) The sialic acid residue is a crucial component of C. jejuni lipooligosaccharide ganglioside mimicry in the induction Guillain-Barre syndrome. *J Neuroimmunol* **174**: 126-132.
- Yamashita, A., S. K. Singh, T. Kawate, Y. Jin & E. Gouaux, (2005) Crystal structure of a bacterial homologue of Na<sup>+</sup>/Cl<sup>-</sup>-dependent neurotransmitter transporters. *Nature* **437**: 215-223.
- Zhang, Y., P. J. Gardina, A. S. Kuebler, H. S. Kang, J. A. Christopher & M. D. Manson, (1999) Model of maltose-binding protein/chemoreceptor complex supports intrasubunit signaling mechanism. *Proc Natl Acad Sci U S A* **96**: 939-944.
- Zhang, Y., D. E. Mannering, A. L. Davidson, N. Yao & M. D. Manson, (1996) Maltose-binding protein containing an interdomain disulfide bridge confers a dominant-negative phenotype for transport and chemotaxis. *J Biol Chem* **271**: 17881-17889.
- Zheng, L., U. Baumann & J. L. Reymond, (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res* **32**: e115.
- Ziegler, C., (2004) Cantilever-based biosensors. *Anal Bioanal Chem* **379**: 946-959.

BLANK PAGE