

The chemoenzymatic synthesis of  
the rare bacterial sugar pseudaminic  
acid, and its utilisation in the study  
of a potential pseudaminidase.

Harriet Chidwick

Doctor of Philosophy

University of York

Chemistry

September 2018

## Abstract

Pseudaminic acid is a non-mammalian nonulosonic acid and a component in a number of bacterial surface structures, including *Pseudomonas aeruginosa* lipopolysaccharide and pili. It has been shown to play a role in virulence in pathogens such as influencing motility in *Campylobacter jejuni* whose flagellin is glycosylated with pseudaminic acid structures. Therefore pseudaminic acid processing enzymes have been identified as putative drug targets but are yet to be fully characterised.

Although generally harmless to healthy individuals, *Pseudomonas aeruginosa* is the most prevalent lung disease in sufferers of cystic fibrosis and is implicated in the majority of cystic fibrosis deaths. Chronic infections are associated with progression into a mucoid phenotype whereby eradication of the pathogen is almost impossible. An enzyme associated with the mucoid phenotype (PA2794) has been putatively assigned as a pseudaminidase and this project aimed to unequivocally assign this enzyme using pseudaminic acid analogues. However strategies to synthesise pseudaminic acid are currently unsuitable for large scale (< 50 mg) production and hence chemical probes for the characterisation of pseudaminic acid processing enzymes are not currently available.

It was attempted to attain a PA2794 crystal structure with pseudaminic acid in complex. However this was inconclusive in elucidating the PA2794 natural substrate and further investigations prevented due to the lack of availability of the putative ligand. Therefore efforts were turned towards the design of a strategy for the synthesis of pseudaminic acid on a large scale, to allow for the pseudaminic acid chemical probes to be developed.

The proposed synthesis utilised the *Campylobacter jejuni* pseudaminic acid biosynthetic enzymes to convert UDP-GlcNAc into pseudaminic acid in one-pot. Expression was optimised for each enzyme and modified to ensure enzyme solubility. A coupled PseB, PseC activity assay was implemented to follow conversion to the unwanted PseB by-product, and conditions adjusted to perturb this reaction. Additionally co-factor substitutes were explored in order to make large scale synthesis economically viable. A route for the large scale enzymatic production of pseudaminic acid was optimised, allowing for an economically viable, efficient and high yielding synthesis.

<b>Abstract.....</b>	<b>2</b>
<b>List of tables .....</b>	<b>7</b>
<b>List of figures.....</b>	<b>8</b>
<b>List of schemes.....</b>	<b>14</b>
<b>Acknowledgements.....</b>	<b>17</b>
<b>Author’s Declaration .....</b>	<b>18</b>
<b>1. Chapter 1 Introduction.....</b>	<b>19</b>
<b>1.1 Cell surface glycosylation .....</b>	<b>20</b>
1.1.1 The role of carbohydrates .....	20
1.1.2 Nomenclature.....	22
1.1.3 Bacterial glycans .....	25
<b>1.2 Glycosyl hydrolases.....</b>	<b>27</b>
1.2.1 An overview of glycosyl hydrolases .....	27
1.2.2 Neuraminidase structure and mechanisms.....	29
1.2.3 Human neuraminidases .....	32
1.2.4 Neuraminidases as therapeutic targets .....	34
<b>1.3 Pseudomonas aeruginosa .....</b>	<b>36</b>
1.3.1 An overview of Pseudomonas aeruginosa .....	36
1.3.2 Pseudomonas aeruginosa infection.....	37
1.3.3 Cystic Fibrosis .....	39
1.3.4 <i>Pseudomonas aeruginosa</i> infection in CF sufferers.....	41
1.3.5 Existing treatment for <i>Pseudomonas aeruginosa</i> .....	42
1.3.6 PA2794; a potential <i>Pseudomonas aeruginosa</i> therapeutic target.....	43
<b>1.4 Pseudaminic acids.....</b>	<b>44</b>
1.4.1 Pseudaminic acid structures.....	44
1.4.2 Chemical synthetic routes towards pseudaminic acid.....	47
<b>1.5 Aims.....</b>	<b>52</b>

<b>2. Chapter 2 <i>Pseudomonas aeruginosa</i> PA2794</b>	<b>54</b>
<b>2.1 Introduction</b>	<b>55</b>
2.1.1 Neuraminic acid analogues as chemical probes	55
2.1.2 PA2794 as a pseudaminidase	58
<b>2.2 Expression and purification of PA2794 and PA2794 F129A</b>	<b>62</b>
2.2.1 Transformation and expression of PA2794 and a PA2794 F129A mutant	62
2.2.2 Purification of PA2794 and PA2794 F129A	65
<b>2.3 Characterisation of PA2794 and PA2794 F129A</b>	<b>67</b>
2.3.1 PA2794 <i>apo</i> crystal structures	67
2.3.2 PA2794 ligand bound crystal structures	70
<b>2.4 Conclusions and Future work</b>	<b>73</b>
<b>3. Chapter 3 Pseudaminic acid biosynthesis</b>	<b>74</b>
<b>3.1 Introduction</b>	<b>75</b>
3.1.1 Sialic acid biosynthesis	75
3.1.2 Identification of the pseudaminic acid biosynthetic enzymes	76
3.1.3 Pseudaminic acid biosynthetic enzyme structure and function	78
3.1.4 Exploitation of the pseudaminic acid biosynthetic pathway for synthesis of pseudaminic acid analogues	83
<b>3.2 Pseudaminic acid biosynthetic enzyme production</b>	<b>85</b>
3.2.1 Optimising enzyme expression conditions	85
3.2.2 Purification of the biosynthetic enzymes	88
<b>3.3 Enzymatic synthesis of pseudaminic acid</b>	<b>90</b>
3.3.1 Enzymatic synthesis of pseudaminic acid utilising enzymes from the <i>Campylobacter jejuni</i> biosynthetic pathway	90
3.3.2 Enzymatic synthesis of pseudaminic acid utilising enzymes from the <i>Aeromonas caviae</i> biosynthetic pathway	92
<b>3.4 Conclusions and future work</b>	<b>96</b>
<b>4. Chapter 4 Optimising the PseB and PseC transformation of UDP-GlcNAc</b>	<b>97</b>
<b>4.1 PseB and PseC activity in the Pse5Ac7Ac biosynthetic pathway</b>	<b>98</b>
4.1.1 <i>Campylobacter jejuni</i> PseB; a UDP-GlcNAc 5-inverting-4,6-dehydratase	98
4.1.2 Additional PseB catalysed epimerisation	100

4.1.3	PseC; a UDP-4-keto-6-deoxy-L-IdoNAc aminotransferase .....	101
4.1.4	Campylobacter jejuni PseC.....	105
4.2	<b>Investigations into the relationship between PseB and PseC .....</b>	<b>106</b>
4.2.1	Complications whilst monitoring the coupled PseB and PseC reaction.....	106
4.2.2	Initial investigations into the coupled PseB and PseC reaction.....	107
4.2.3	The PseB and PseC coupled reaction in deuterium oxide.....	110
4.2.4	Preventing formation of the PseB by-product .....	113
4.3	<b>Conclusions and future work .....</b>	<b>117</b>
5.	<b>Chapter 5 Identification of alternative acetyl-transfer strategies.....</b>	<b>118</b>
5.1	<b>Utilisation of acetyltransferases as synthetic tools .....</b>	<b>119</b>
5.1.1	The PseH catalysed acetyl transfer step .....	119
5.1.2	Acetyl-CoA binding in GNAT enzymes.....	120
5.1.3	Acetyl-transfer methods during enzymatic syntheses.....	123
5.1.4	Introduction to the promiscuity of aminoglycoside <i>N</i> -acetyltransferases co-factor binding sites.....	125
5.2	<b>Acetylation during the chemo-enzymatic synthesis of Pse5Ac7Ac.....</b>	<b>128</b>
5.2.1	Chemical acetylation of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc .....	128
5.2.2	<i>In situ</i> regeneration of acetyl-CoA using acetyl thiocholine iodide .....	130
5.2.3	<i>N</i> -acetyl cysteamine thioacetate as a PseH co-factor substitute .....	138
5.2.4	Other alternate PseH co-factor substitutes.....	141
5.3	<b>Manipulation of the acetylation step for the synthesis of pseudaminic acid C7 derivatives.....</b>	<b>147</b>
5.3.1	Utilisation of the PseH co-factor substitutes to introduce other C7 functionality 147	
5.3.2	Synthesis of <i>N</i> -acetylcysteamine thioazidoacetate .....	148
5.3.3	Utilisation of <i>N</i> -acetylcysteamine thioazidoacetate in the enzymatic synthesis of a Pse5Ac7Az.....	149
5.3.4	Extending the utilisation of <i>N</i> -acetylcysteamine thioacyl derivatives to access other pseudaminic acid derivatives .....	151
5.4	<b>Conclusions and future work .....</b>	<b>152</b>
6.	<b>Chapter 6 Concluding remarks .....</b>	<b>153</b>
6.1	<b>Conclusions and future work .....</b>	<b>154</b>

<b>7. Chapter 7 Experimental .....</b>	<b>156</b>
<b>7.1 General methods .....</b>	<b>157</b>
<b>7.2 PA2794 and PA2794 F129A mutant.....</b>	<b>161</b>
<b>7.3 <i>Campylobacter jejuni</i> Pse5Ac7Ac biosynthetic enzymes .....</b>	<b>165</b>
<b>7.4 Chemo-enzymatic syntheses.....</b>	<b>167</b>
<b>7.5 Chemical synthesis .....</b>	<b>170</b>
<b>Appendix.....</b>	<b>172</b>
<b>1. PA2794 sequences.....</b>	<b>173</b>
<b>2. PA2794 F129A sequences.....</b>	<b>174</b>
<b>3. PA2794 and PA2794 F129A Akta trace for Ni-His trap purification.....</b>	<b>175</b>
<b>4. PA2794 and PA2794 F129A Akta trace for size exclusion purification.....</b>	<b>177</b>
<b>5. Crystallography statistics.....</b>	<b>178</b>
<b>6. <i>Campylobacter jejuni</i> biosynthetic enzymes. ....</b>	<b>179</b>
<b>7. <i>Campylobacter jejuni</i> enzymes Akta trace for Ni-His trap purification.....</b>	<b>181</b>
<b>8. LC-MS negative ESI of controls for co-factor substitute reactions.....</b>	<b>184</b>
<b>9. NMR.....</b>	<b>187</b>
<b>List of Abbreviations .....</b>	<b>188</b>
<b>References .....</b>	<b>190</b>

## List of tables

**Table 2.1** Expression trial conditions to optimise production of PA2794.

**Table 3.1** Table of the recombinant plasmids containing pseudaminic acid biosynthetic enzymes and their associated antibiotic resistance.

**Table 3.2** Expression trial conditions to optimise production of the biosynthetic enzymes.

**Table 3.3** Optimised conditions for production of pseudaminic acid biosynthetic enzymes.

**Table 3.4** The quantity of each biosynthetic enzyme purified per L of LB culture.

**Table 3.5** General reaction components and concentrations for the “one-pot”small scale enzymatic synthesis of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**) for small scale enzymatic synthesis of Pse5Ac7Ac (**3.1**)

**Table 3.6** The combination of *C. jejuni* and *A. cavaie* proposed Pse5Ac7Ac (**1.13**) biosynthetic enzymes trialled during activity experiments to elucidate the *A. cavaie* biosynthetic pathway.

**Table 4.1** LC-MS ESI [M-H]<sup>-</sup> values for the components of the *C. jejuni* PseB, PseC enzymatic synthesis

**Table 4.2** Negative ESI [M-H]<sup>-</sup> values for sugar molecules present in the deuterated PseB, PseC reaction

## List of figures

**Figure 1.1** Fischer projections of hexose examples, glucose and mannose, displaying the potential for the different arrangements of diastereomers and enantiomers.

**Figure 1.2** The different structural forms of D-glucose.

**Figure 1.3** Symbols assigned to the most abundant vertebrate monosaccharides in order to standardise depiction of glycans

**Figure 1.4** An updated symbolic representation of monosaccharides **a)** the general symbolic representation and **b)** the assignment of symbols to monosaccharides that were unknown upon publication but that have recently been characterised.

**Figure 1.5** Glycan structure examples using the CFG standardised system in either the written format or with the monosaccharides represented by symbols, with or without additional bond information for **a)** an example *N*-glycan and **b)** an example plant glycan.

**Figure 1.6** Gram positive bacteria cell wall with the thick layer of peptidoglycan surrounding a phospholipid bilayer.

**Figure 1.7** Gram negative bacteria cell wall with the thin layer of peptidoglycan between two phospholipid bilayers with an outer layer of lipopolysaccharides.

**Figure 1.8** Examples of sialic acid structures prevalent in nature with different C5 functionality.

**Figure 1.9** Different natural Neu5Ac (**1.1**) linkages;  $\alpha$ 2,3 (**1.4**),  $\alpha$ 2,6 (**1.5**), and  $\alpha$ 2,8 (**1.6**).

**Figure 1.10** Crystal structure of a typical sialidase (PDB 1EUS)<sup>1</sup> with **a)** the conserved Asp boxes (magenta) and RIP motif (cyan) highlighted in the full structure and, **b)** the active site with catalytic residues (sea green) and three interacting arginine residues (lilac) highlighted.

**Figure 1.11** A selection of neuraminidase inhibiting structures; **a)** Neuraminidase inhibitors and **b)** existing anti-viral therapeutics.

**Figure 1.12** The structurally related nonulosonic acids; Neu5Ac (**1.1**) and Pse5Ac7Ac (**1.13**).

**Figure 1.13** Examples of some pseudaminic acid derivatives.<sup>2</sup>

**Figure 1.14** Examples of pseudaminic acid cell surface structures, showing a selection of the different glycosidic linkages utilised in nature.



**Figure 2.1** Structures of Neu5Ac (1.1) chemical probes; **a**) Neu5Ac fluorophores 4-methylumbelliferyl- $\alpha$ -Neu5Ac (2.5) and *p*-nitrophenyl- $\alpha$ -Neu5Ac (2.6), and **b**) a neuraminidase inhibitor Neu5Ac2en (1.8) and covalent inactivator 2,3-difluoro-Neu5Ac (2.7)

**Figure 2.2** Nonulosonic acid structures proposed as the potential PA2794 product

**Figure 2.3** Crystal structure of PA2794 (lilac) (PDB 2W38)<sup>3</sup> overlaid with the *M. viridifaciens* neuraminidase (light blue) (PDB 1EUS)<sup>1</sup> with the Asp-box sites highlighted (purple and dark cyan).

**Figure 2.4** Crystal structure of PA2794 active site residues (purple) (PDB 2W38)<sup>3</sup> overlaid with a sialidase (cyan) in complex with Neu5Ac2en (1.8) (green) (PDB 1EUS).<sup>1</sup> Highlighting differences in **a**) the orientation of conserved neuraminidase active site residues, and **b**) the arrangement of the catalytic residues.

**Figure 2.5** *Pseudomonas aeruginosa* surface structures **a**) the O-7 antigen repeating unit and glycosylating trisaccharide of strain 1244 pili (2.8) and **b**) the structure of the O-9 antigen repeating unit (2.9).

**Figure 2.6** Proposed structures of pseudaminic acid chemical probes; **a**)  $\alpha$ - and  $\beta$ - linked fluorophores (2.10, 2.11) and **b**) the structure of the predicted pseudaminidase intermediate (2.12) and covalent inactivators (2.13, 2.14).

**Figure 2.7** SDS PAGE analysis of crude lysate samples of the PA2794 enzyme **a**) under different induction conditions after induction with 0.1 mM (A, B, E, F) or 0.5 mM (C, D, G, H) IPTG and **b**) under these conditions but with no induction with IPTG.

**Figure 2.8** DNA agarose gel following a point mutation PCR of PA2794 pEt15b recombinant plasmid.

**Figure 2.9** SDS PAGE analysis of **a**) PA2794 and **b**) PA2794 F129A after purification using a Nickel affinity column.

**Figure 2.10** SDS PAGE analysis of following purification using a Nickel affinity column and then a size exclusion column for **a**) PA2794 and **b**) PA2794 F129A.

**Figure 2.11** Image of PA2794 crystals developed in well solutions composed of **a**) Bicine (0.1 M, pH 5.0), 18 % PEG 6K and **b**) Bicine (0.1 M, pH 5.0), 24 % PEG 6K

**Figure 2.12** The fully assigned PA2794 crystal structure (lilac) resolved to 1.2 Å, with electron density also assigned to well-ordered bicine and glycerol solutes (dark purple).

**Figure 2.13** Overlay of the published PA2794 structure (PDB 2w38, lilac, dark purple) and the newly refined structure (ice blue, grey), demonstrating the similarity of **a**) their overall structure and **b**) active site side chain conformations.

**Figure 2.15** Overlay of the PA2794 *apo* structure (ice blue, grey) and the structure soaked with a potential ligand; Pse5Ac7Ac (**1.13**) (coral, gold), demonstrating the similarity of **a**) their overall structure and **b**) active site side chain conformations and the position of binding of an active site solute molecule.

**Figure 2.16** Comparison of the structures of bicine (**2.15**) and glycerol (**2.16**) with Pse5Ac7Ac (**1.13**).

**Figure 3.1** *H. pylori* PseB (PDB 2GN4) crystal structure in complex with the UDP-GlcNAc substrate (**3.1**) (green) and the tightly bound NADPH co-factor (red) and close-ups of the substrate and co-factor binding sites displaying the numerous hydrogen bonds between molecules and enzyme.

**Figure 3.2** *C. jejuni* PseH (PDB 4XPL) crystal surface structure in complex with the acetyl-CoA co-factor with the acetyl group protruding into the predicted substrate binding groove.

**Figure 3.3** SDS PAGE analysis of crude samples of the biosynthetic enzymes under different induction conditions **a**) PseB, **b**) PseC, **c**) PseH, **d**) PseG, **e**) PseI.

**Figure 3.4** Depiction of the purified *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic enzymes on a SDS PAGE gel.

**Figure 3.5** Negative ESI LC-MS of the small scale Pse5Ac7Ac (**1.13**) enzymatic reaction, **a**) negative control (reaction mixture without the first biosynthetic enzyme; PseB) after 2.5 hrs, demonstrating substrate UDP-GlcNAc (**3.1**)  $[M-H]^-$  606 Da stability and no detection of any of the biosynthetic products and **b**) analysis of the enzymatic synthesis at 2.5 hrs, demonstrating full conversion to the desired Pse5Ac7Ac (**1.13**) product  $[M-H]^-$  333 Da.

**Figure 3.6** Negative ESI LC-MS of the Pse5Ac7Ac (**1.13**) enzymatic reaction **a**) the negative control with no PseB or FlmA showing detection of only UDP-GlcNAc (**3.1**) starting material after 4 hrs, **b**) conversion to Pse5Ac7Ac (**1.13**) after 4 hrs with *C. jejuni* PseB and *A. caviae* Flm B, Flm D and NeuB and **c**) demonstrating *A. caviae* FlmA inability to turnover UDP-GlcNAc (**3.1**).

**Figure 3.7** Sequence comparison of the *H. pylori*, *C. jejuni* and *A. caviae* PseB enzymes with residues highlighted that are involved in substrate (green) and co-factor (red) hydrogen bonding.

**Figure 3.8** Negative ESI LC-MS of the small scale enzymatic reaction employing *A. cavaie* FlmA and *C. jejuni* PseC, PseH, PseG and PseI in the standard reaction mixture with the inclusion of NADPH (1.5 mM).

**Figure 4.1** The *H. pylori* PseC (PDB 2FNU) homodimer (chain A ice blue, chain B gold), with the bound co-factor and substrate indicating position of the active site at the dimer interface.

**Figure 4.2** *H. pylori* PseC crystal studies (PDB 2FN6)<sup>4</sup> **a**) the co-factor binding site in complex with the PLP co-factor (**2311**), highlighting Type 1 aminotransferase conserved residues and **b**) electron density of the PLP-enzyme internal aldimine intermediate (figure adapted from the original paper).<sup>4</sup>

**Figure 4.3** The PMP-substrate external aldimine (ball and stick model) **a**) in complex with the surrounding *H. pylori* PseC residues (cyan chainA, yellow chain B) and **b**) with the catalytic Lys183 residue.

**Figure 4.4** Alignment comparison of the *H. pylori* and *C. jejuni* PseC sequences with *H. pylori* residues shown to be proximal to the co-factor and active site highlighted (green). The conserved *C. jejuni* residues in these sites are also highlighted (cyan), along with the conserved catalytic lysine (magenta).

**Figure 4.5** *C. jejuni* PseB and PseC reaction progression **a**) after incubating for 10 mins displaying conversion to the PseC product (**3.8**), **b**) after incubating for 45 mins displaying increased conversion to the PseC product (**3.8**), and **c**) after incubating for 6 hrs displaying no further increase in conversion to the PseC product (**3.8**).

**Figure 4.6** Negative ESI analysis of formation of the *C. jejuni* PseB products in deuterated solvent; **a**) control reaction of UDP-GlcNAc (**3.1**) after 12 hrs incubation displaying no conversion **b**) 30 mins after incubating UDP-GlcNAc (**3.1**) with PseB, **c**) 120 mins after incubating UDP-GlcNAc (**3.1**) with PseB and **d**) 120 mins after addition of PseC and co-factors to the reaction mixture **c**).

**Figure 4.7**; Negative ESI analysis of formation of the *C. jejuni* PseB (**4.13-4.16**) and PseC products (**4.17**) in deuterated solvent whilst investigating the effect of increasing the PseC concentration; **a**) concentration ratio of [PseB : PseC] of 1 : 1 after 10 mins incubation, **b**) concentration ratio of [PseB : PseC] of 1 : 5 after 10 mins incubation, **c**) concentration ratio of [PseB : PseC] of 1 : 1 after 2 hrs incubation, **d**) concentration ratio of [PseB : PseC] of 1 : 5 after 2 hrs incubation.

**Figure 5.1** Crystal structure of the *C. jejuni* PseH acetyltransferase (PDB 4XPL) highlighting **a**) the canonical secondary structure of GNAT enzymes and **b**) the co-factor V-shape cleft binding site between strands  $\beta$ 4 and  $\beta$ 5 with the bound acetyl-CoA (**5.1**) co-factor.

**Figure 5.2** Breakdown of the structure of acetyl-CoA (**5.1**) and CoA (**5.2**) into substructures.

**Figure 5.3** Crystal structure of *C. jejuni* PseH (PDB 4XPL) in complex with its co-factor acetyl-CoA (**5.1**) showing hydrogen bonds between residues and the pantetheine arm and pyrophosphate group.

**Figure 5.4** Negative ESI LC-MS demonstrating **a**) Enzymatic conversion of the PseC product (**3.8**) and **b**) chemical acetylation to afford the PseH (**3.9**) product.

**Figure 5.5** Overlay of the LC-MS negative ESI m/z 631 extracted ion count traces after 4 hrs showing the effect of addition of the acetylthiocholine iodide on PseH catalysed acetyltransfer whilst using sub-stoichiometric amounts of acetyl-CoA (**5.1**).

**Figure 5.6** Negative ESI LC-MS analysis of conversion to the PseH product (**3.9**) from UDP-GlcNAc (**3.1**), investigating the use of acetylthiocholine iodide (**5.5**) as a regeneration factor with sub-stoichiometric amounts of acetyl-CoA (**5.1**) **a**) 0 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**), **b**) 0.15 mM acetyl-CoA (**5.1**) and 0 mM acetylthiocholine iodide (**5.5**), **c**) 0.15 mM acetyl-CoA (**5.1**) and 2 mM acetylthiocholine iodide (**5.5**), and **d**) 0.15 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**).

**Figure 5.7** Overlay of the LC-MS negative ESI m/z 631 extracted ion count traces after 4 hrs showing the effect of decreasing the concentration of CoA from sub-stoichiometric to catalytic amounts on PseH catalysed acetyltransfer in the presence of 20 mM regeneration factor; acetylthiocholine iodide (**5.5**).

**Figure 5.8** LC-MS negative ESI mass spectra demonstrating conversion to the acetylated PseH product (**3.9**) using 20 mM regeneration factor (**3.5**) and substoichiometric amounts of CoA (**5.2**); **a**) 0.15 mM CoA (**5.2**), **b**) 0.015 mM CoA (**5.2**), **c**) 0.0015 mM CoA (**5.2**), **d**) 0mM CoA (**4.2**).

**Figure 5.9** Overlay of the LC-MS negative ESI m/z 631 extracted ion count traces after 4 hrs showing the effect of decreasing the concentration of CoA from sub-stoichiometric to catalytic amounts on PseH catalysed acetyltransfer in the presence of 100 mM regeneration factor; acetylthiocholine iodide (**5.5**).

**Figure 5.10** LC-MS negative ESI mass spectra demonstrating conversion to the acetylated PseH product (**3.9**) using 100 mM regeneration factor (**5.5**) and substoichiometric amounts of CoA (**5.4**); **a**) 0.15 mM CoA (**5.2**), **b**) 0.015 mM CoA (**5.2**), **c**) 0.0015 mM CoA (**5.2**), **d**) 0 mM CoA (**5.2**).

**Figure 5.11** LC-MS negative ESI demonstrating successful production of Pse5Ac7Ac (**1.13**) in the “one-pot” system when employing 0.0015 mM CoA (**5.2**) and 100 mM acetyl thiocholine iodide (**5.5**).

**Figure 5.12** LC-MS negative ESI analysis of the PseH controls demonstrates that the reaction halts at the PseC product intermediate (**3.8**) in the one-pot, enzymatic synthesis of Pse5Ac7Ac (**1.13**) containing **a**) PseB, PseC, PseG, Psel (their co-factors) and PseH or **b**) PseB, PseC, PseG, Psel (their co-factors) and *N*-acetyl cysteamine thioester (**5.9**).

**Figure 5.13** LC-MS negative ESI analysis after 4 hrs, demonstrating conversion to the PseH product (**3.9**) using *N*-acetylcysteamine thioacetate (**5.9**) as a PseH co-factor substitute in a reaction consisting of UDP-GlcNAc (1 mM), PLP (1.5 mM), *L*-glutamate (10 mM), PseB, PseC and PseH in sodium phosphate buffer (50 mM, pH 7.4).

**Figure 5.14** LC-MS negative ESI analysis after 4 hrs, demonstrating conversion to the Pse5Ac7Ac (**1.13**) using *N*-acetylcysteamine thioacetate (**5.9**) as a PseH co-factor substitute in a reaction consisting of UDP-GlcNAc (1 mM), PLP (1 mM), *L*-glutamate (10 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M) and Psel (25  $\mu$ M) in sodium phosphate buffer (50 mM, pH 7.4).

**Figure 5.15** Phenolacetate (**5.19-5.21**) and phenolthioacetate (**5.22-5.24**) structures proposed as putative acetyl-transferase co-factor alternatives.

**Figure 5.16** Enzymatic synthesis of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**) in 10 % DMSO, sodium phosphate buffer (50 mM, pH 7.4) containing UDP-GlcNAc (1 mM), PLP (1 mM), *L*-glutamate (10 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M) and Psel (25  $\mu$ M).

**Figure 5.17** Negative ESI LC-MS investigating the ability of different PseH co-factors to promote conversion from UDP-GlcNAc (**3.1**) to the PseH product (**3.9**) in an enzymatic reaction; UDP-GlcNAc (1 mM), PLP (1.5 mM), *L*-glu (10 mM), PseB (25 M), PseC (125 M) and PseH (50 M) **a**) 10 mM *S*-phenyl thioacetate (**5.22**), **b**) 10 mM phenyl acetate (**5.20**), **c**) 10 mM *p*-tolyl acetate (**5.21**), **d**) 10 mM *p*-nitrophenyl acetate (**5.19**), **e**) 10 mM *S*-(*p*-tolyl)thioacetate (**5.24**) and **f**) 10 mM *S*-(*p*-tolyl)thioacetate (**5.24**).

## List of schemes

**Scheme 1.1** General reaction for the glycosyltransferase (GT) catalysed formation of a glycosidic linkage between two monosaccharides and the glycosylhydrolase (GH) catalysed cleavage of a glycosidic bond.

**Scheme 1.2** General glycosidase hydrolysis mechanisms resulting in **a)** retention or **b)** inversion of the configuration of the anomeric bond.

**Scheme 1.3** A typical retaining sialidase double displacement mechanism, using an  $\alpha$ -linked Neu5Ac (**1.1**) as an example substrate.

**Scheme 1.4** Attempted synthetic routes to Pse5Ac7Ac (**1.13**) producing a number of stereoisomeric products (adapted from **a)** Knirel and **b)** Ito).

**Scheme 1.5** Attempted synthetic routes to Pse5Ac7Ac (**1.13**) from Neu5Ac (**1.1**) via a key 5,7 bis-azide intermediate (**1.25**).

**Scheme 1.6** A novel synthetic route to pseudaminic acid derivatives producing a protected nonulosonic sugar with the desired pseudaminic acid stereochemistry.

**Scheme 2.1** Neu5Ac (**1.1**) anomeric configurations as the monosaccharide in aqueous solution and during GT catalysed glycosylation and GH catalysed cleavage.

**Scheme 2.2** Thiobarbituric assay; a coupled reaction to observe enzymatic cleavage of Neu5Ac (**1.1**). Neu5Ac (**1.1**) is oxidised so that it can act as a linker for two thiobarbituric acid (**2.3**) molecules producing a chromophore that can be detected as an indicator of linked Neu5Ac (**2.1**) cleavage.

**Scheme 3.1** The biosynthetic pathway from UDP-GlcNAc (**3.1**) to CMP-Neu5Ac (**3.5**) in bacteria and mammals.

**Scheme 3.2** The CMP-Pse5Ac7Ac (**3.16**) biosynthesis detailing the *C. jejuni* and *H. pylori* pathway (blue) with any enzymatic deviations observed in *A. caviae* (red) and *B. thuringiensis* (green) highlighted.

**Scheme 3.3** The first half-transamination reaction of PLP-dependent aminotransferases, catalysing conversion of PLP (**3.11**) to PMP (**3.12**) utilising a free amino donor (**3.14**).

**Scheme 3.4** The proposed mechanism for PseG catalysed UDP hydrolysis, utilising a histidine-activated and isoleucine-stabilised water molecule as a nucleophile to attack at the anomeric centre.

**Scheme 3.5** *C. jejuni* PseI pseudaminic acid synthetase PEP condensation mechanism.

**Scheme 3.6** A chemoenzymatic route to the PseG substrate, UDP-6-deoxy-AltNac (**3.9**).

**Scheme 3.7** One-pot enzymatic production of Pse5Ac7Ac (**1.13**) from UDP-GlcNac (**3.1**).

**Scheme 4.1** PseB catalysed oxidation, dehydration and reduction of the substrate UDP-GlcNac (**3.1**) to form the initial product UDP-4-keto-6-deoxy-L-IdoNac (**3.7**) which is in equilibrium with the hydrated form (**4.2**) in aqueous conditions.

**Scheme 4.2** *C. jejuni* interlinking schemes from UDP-GlcNac to two deoxysugar intermediates (**3.7**, **4.1**) in the biosynthetic pathways of two sugars important for protein glycosylation.

**Scheme 4.3** The proposed PseC mechanism transferring an amino group from the PMP co-factor, generated *in situ*, to the keto-sugar (**3.8**) via formation of an external aldimine (**4.10**) that crystal structures have been shown as present in the active site during the reverse reaction.

**Scheme 4.4** *C. jejuni* PseB and PseC full reaction scheme highlighting conversion to the PseB inverted-by-product (**4.1**) and the non-enzymatic production of the hydrated PseB products (**3.2**, **3.12**).

**Scheme 4.5** The coupled *C. jejuni* PseB and PseC reaction converting UDP-GlcNac (**3.1**) to the second Pse5Ac7Ac (**1.13**) biosynthetic intermediate UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNac (**3.8**).

**Scheme 4.6** Partial PseB mechanism highlighting the D<sub>2</sub>O solvent molecule utilised during the mechanism and the corresponding product detected by negative ESI LC-MS whereby all of the labile deuterium have undergone solvent exchange with the H<sub>2</sub>O/MeCN mobile phase.

**Scheme 4.7** Partial PseC mechanism in deuterium proposing formation of a stable C4 carbon-deuterium bond and the resulting sugar identified by negative ESI LC-MS.

**Scheme 5.1** *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic pathway highlighting the acetyltransfer step.

**Scheme 5.2** Acetyltransfer mechanism for synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNac (**3.9**) showing the *C. jejuni* PseH catalytic acid; Tyr128, and the proposed water mediated amine deprotonation.

**Scheme 5.3** A synthetic route of acetyltransfer to GlcN (**5.3**) to produce GlcNac (**5.4**).

**Scheme 5.4** Mechanism for the acetyl transfer from acetyl-thiocholine iodide (**5.5**) to CoA (**5.1**).

**Scheme 5.5** Acetyltransferase catalysed acetylation of GlcN-1-P utilising the natural co-factor acetyl-CoA (5.1) or *N*-acetylcysteamine thioester (5.9).

**Scheme 5.6** A chemo-enzymatic route to the Pse5Ac7Az (5.10) utilising a combination of enzymatic and chemical steps.

**Scheme 5.7** Chemical acetylation of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (3.8) to afford UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (3.9) using a silver acetate catalyst (5.15) to activate the acetic anhydride making it more susceptible to nucleophilic attack.

**Scheme 5.8** Enzymatic synthesis of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (3.8) followed by chemical acetylation to UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (3.9) using a silver acetate catalyst (4 equiv) and acetic anhydride (11.5 equiv) in MeOH, at room temperature for 3 hrs.

**Scheme 5.9** Regeneration of acetyl-CoA (5.1) *in situ* during the enzymatic synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (5.9).

**Scheme 5.10** Synthetic route to *N*-acetyl cysteamine thioacetate; addition of acetic anhydride to cysteamine HCl in H<sub>2</sub>O at pH 8.0, 0 °C, followed by stirring at pH 7.0 at 0 °C for 2 hrs.

**Scheme 5.11** The resonance forms of esters and thioesters.

**Scheme 5.12** Introduction of desired C7 functionality in pseudaminic acids utilising a co-factor substitute.

**Scheme 5.13** Three synthetic strategies for coupling azidoacetic acid and *N*-acetylcysteamine HCl to form *N*-acetylcysteamine thioazidoacetate (5.26).

**Scheme 5.14** One-pot enzymatic route to Pse5Ac7Az (5.10), utilising the *C. jejuni* Pse5Ac7Ac (1.13) biosynthetic enzymes and an azido-tagged PseH co-factor substitute (5.26).



## Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr Martin Fascione for giving me this opportunity to complete this PhD. Your kind support, words of wisdom and constant encouragement have driven me and this research to where it is today. Also thanks to Prof. Gideon Davies and Prof. Tony Wilkinson for their advice on the direction of the project and sharing their structural biology knowledge. Thanks to Martin and Alison for the enthusiastic, friendly, team atmosphere they have created, it has been inspirational to work within this lab. Thanks are also extended to every member of the Parkin/Fascione groups, who have contributed so much to making the last four years such an enjoyable experience.

I was very fortunate to begin and end this journey as part of the Fascione four; Richard, Emily and Robin thank you for all of the hugs, laughs and memories. Your individual advice and support has helped me survive this rollercoaster ride, thanks for checking that LC-MS for me, I love you guys! Thanks to Darshita, Hope and Tess for sharing their scientific knowledge over the years, and for having the patience to answer all questions. Thank you Darshita for all the advice on chemical synthesis (especially NMR) and Tess for the biological insight you've brought to the lab. Additionally I would like to thank Julia for how much you do to support every member of the lab and generally keep the whole place functioning, you are irreplaceable. Thank you for introducing and explaining to me to all of the microbiological techniques I have needed over the years.

I would like to acknowledge all of the technical staff in the Chemistry and Biology departments who have assisted me over the years. In particular Dr Ed Bergström who has been a wealth of knowledge on LC-MS and has consistently ensured a working LC-MS was available, which was essential for parts of this project. Also to Wendy Offen who imparted her wisdom of protein crystallography and assisted in collection of the data.

I would like to thank my family who have been unwavering in their love, support and confidence in my abilities throughout my life and for always picking up the phone for the many necessary "correspondence breaks". Thanks to all my friends who have contributed by making me tea, being a running buddy and making me laugh. Finally to Stephen, thank you for being my best friend and always being on my team, you've made it all easier and a lot more fun.

## Author's Declaration

I, Harriet Chidwick, declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References. Any contributions to this work other than my own are stated below.

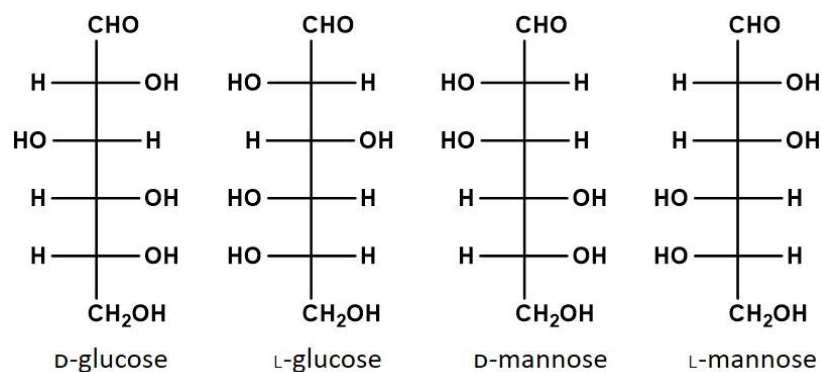
- **Chapter 2** Crystals were fished, placed in cryo protectant and flash frozen by Wendy Offen. Preliminary crystallography data (*in house*) was collected by Sam Hart (X-ray Technician).
- **Chapter 3** Initial work transforming and expressing the *Campylobacter jejuni* pseudaminic acid biosynthetic enzymes was carried out alongside Emily Flack during her Biochemistry degree. Expression trials and all data shown are my own.
- **Chapter 3** *Aeromonas caviae* pseudaminic acid biosynthetic enzymes were expressed and purified at the University of Sheffield (Joseph Ferner).
- **Chapter 5** Synthesis of N-acetylcysteamine thioformyl and use within the established system as PseH co-factor substitute was undertaken by Matthew Best (MChem).

## **Chapter 1 Introduction**

## 1.1 Cell surface glycosylation

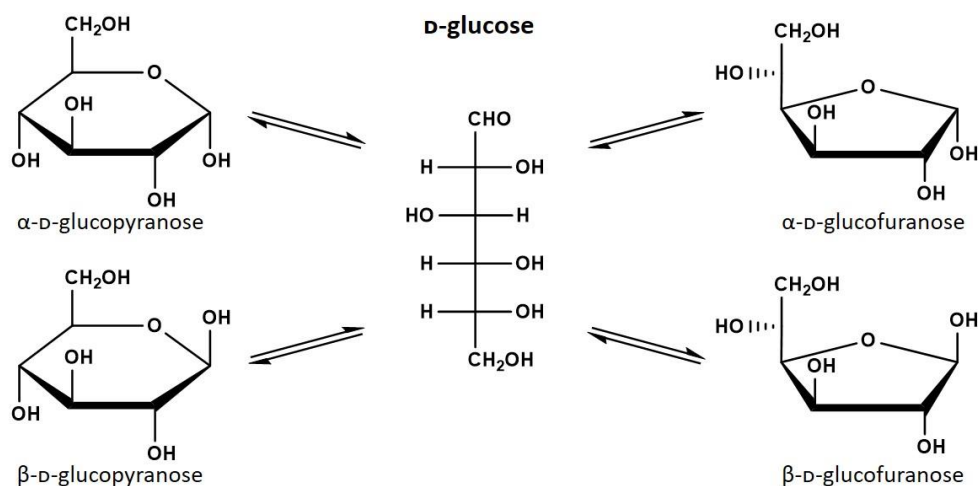
### 1.1.1 The role of carbohydrates

Carbohydrates are the most abundant group of compounds on Earth and play a vast role across a multitude of biological processes.<sup>5</sup> However, in contrast to proteins, there is no strict template for the biosynthesis of carbohydrates which results in significant heterogeneity, and an extra level of complexity in research. Historically, carbohydrates were considered as merely structural molecules or as an energy source, with significance research efforts focussed on diet and nutrition. This has somewhat limited investigations into their other biological roles in health and disease, such as their prevalence on cell surfaces and when attached to proteins (wherein they are known as glycans).<sup>6</sup> The vast complexity, diversity and heterogeneity of glycan structures<sup>7</sup> has also contributed to the lag in their functional understanding compared to other cell surface biomolecules. It has recently become clear however that a greater emphasis needs to be placed on the study of glycans in order to truly understand and manipulate biological systems. Indeed, the degree and type of glycosylation on both mammalian and bacterial cell surfaces regulates a multitude of biological processes and they are essential for facilitating cellular communication.<sup>8</sup>



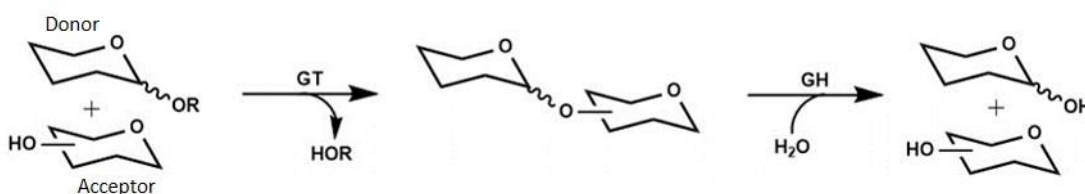
**Figure 1.1** Fischer projections of hexose examples, glucose and mannose, displaying the potential for different diastereomers and enantiomers.

The inherent complexity of carbohydrates begins at the monosaccharide building block level, with even simple hexose carbohydrate molecules (e.g. C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) having a large number of different configurational arrangements. The number of available diastereomers increases as the chain length increases, in addition to the existence of pairs of enantiomers (D and L) (**Figure 1.1**). Further complications arise due to some monosaccharides being able to exist in ring open forms and hence can interconvert between 5- and 6-membered furanose or pyranose rings as well as between different anomeric configurations ( $\alpha$  or  $\beta$ ) (**Figure 1.2**).



**Figure 1.2** The different structural forms of D-glucose.

As carbohydrates possess multiple hydroxylated stereocentres the structural diversity range of oligosaccharides increases with length. Hence upon “polymerisation” the number of both linear and branched structures accessible increases exponentially, further increasing the complexity of even simple carbohydrate studies.<sup>7</sup> In nature only a small portion of the potential oligosaccharide configurations are utilised, and their state of “polymerisation” is regulated by glycosyltransferases (GTs) and glycosylhydrolases (GHs).<sup>9</sup> GTs catalyse the biosynthesis of a glycosidic bond *via* condensation at the anomeric position of an “activated donor” and the hydroxyl of an acceptor molecule whilst GHs catalyse hydrolysis of such glycosidic bonds (**Scheme 1.1**).















**Scheme 1.1** General reaction for the glycosyltransferase (GT) catalysed formation of a glycosidic linkage between two monosaccharides and the glycosylhydrolase (GH) catalysed cleavage of a glycosidic bond.

Glycans are often covalently linked to other molecules to form a complex variety of glycoconjugates such as glycolipids, glycoproteins and glycosides. Cell surfaces can also be decorated with glycoconjugates which form a dense layer surrounding the cell called the glycocalyx.<sup>10</sup> The glycocalyx is dynamic but essential and plays a variety of physiological roles including communicating information about the cell to extracellular moieties, such as what type of cell it is, and if the cell is diseased or healthy.<sup>11</sup>

### 1.1.2 Nomenclature

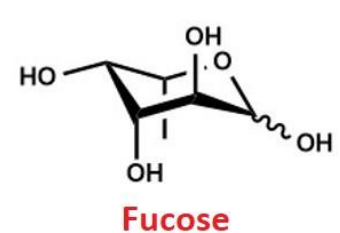
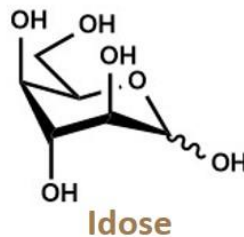
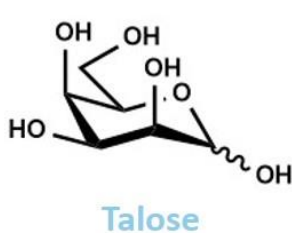
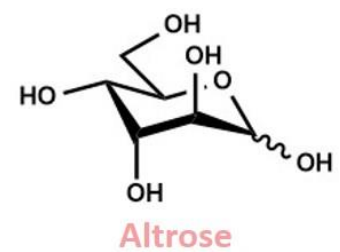
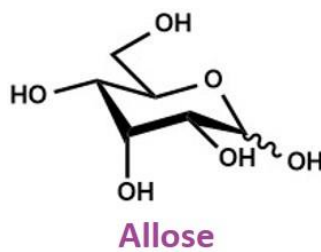
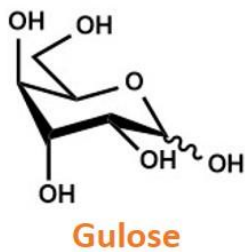
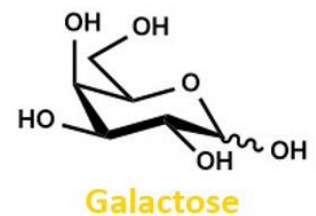
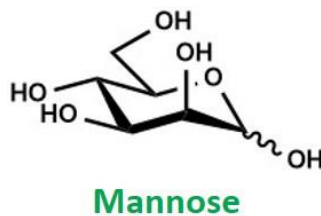
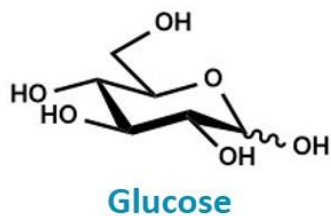
As the biological significance of carbohydrates became more apparent, a small number of monosaccharides were found to be abundant in vertebrates and hence symbol representations were assigned to enable facile depictions of these structures. Although first used sporadically, symbolic representations of glycans quickly gained popularity and it was deemed necessary to implement a standardised system. The fundamentals of such a system were initially developed in the late 1990s,<sup>5</sup> *via* extension of the monosaccharide symbols first presented by Kornfeld during the discussion of complex oligosaccharide synthesis.<sup>12</sup> Symbol representations were assigned to the monosaccharides that early investigations had found to be the most common in higher animal oligosaccharides. Basic shapes were utilised that were either blank, half-filled or fully filled to depict the monosaccharide (**Figure 1.3**).<sup>5</sup> Although this simplistic approach was successful for the facile representation of the most abundant monosaccharide building blocks, it heavily restricted the number of monosaccharides that could be represented in this way. As the diversity of characterised glycans increased and research expanded to invertebrates, a plethora of other sugars were discovered that could not be satisfactorily represented by this symbol system.

	Glucose		GlcNAc		Sialic acid, unspecified
	Fucose		GalNAc		Uronic acid, unspecified
	Galactose		HexNAc		Glucuronic acid
	Mannose		Hexose		Iduronic acid

**Figure 1.3** Symbols assigned to the most abundant vertebrate monosaccharides in order to standardise depiction of glycans

This original categorisation was thoroughly revised ten years after publication whereby the Consortium for Functional Glycomics (CFG) updated the system to include colour and a wider number of monosaccharides.<sup>13</sup> This was quickly criticised for still being rather limiting considering the increasing number of important monosaccharides being studied and was reorganised in 2015 to give the system that is universally adopted today.<sup>14</sup> The introduction of colour and a systemisation of symbols allowed for representation of a much greater number of monosaccharides (**Figure 1.4a**). Sugars of the same functionality (hexoses, hexosamines, acidic sugars etc.) were designated the same shape with differences in their stereochemistry reflected by a change in colour (e.g. galactose stereochemistry is yellow, mannose is green etc.).

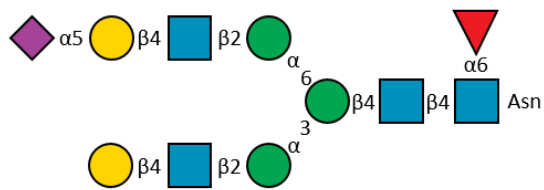
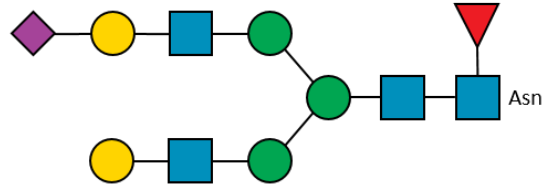
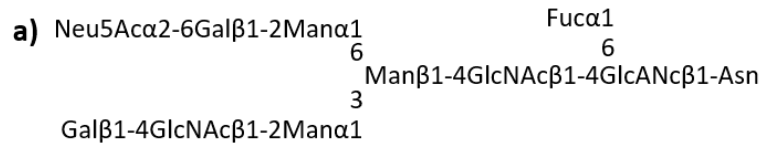
a) ○	Hexose	△	Deoxyhexose	◇	Deoxynonulosonate
□	HexNAc	▲	DeoxyhexNAc	◊	Di-deoxynonulosonate
▧	Hexosamine	◻	Di-deoxyhexose	⬡	Unknown
◊	Hexuronate	☆	Pentose	◑	Assigned



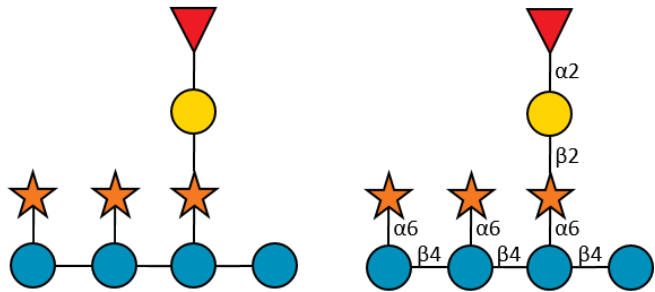
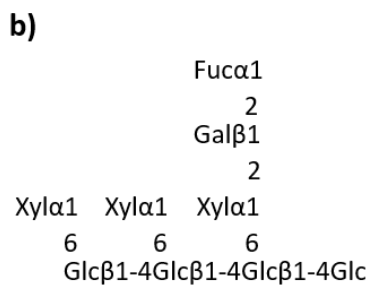
b) ●	Bac	●	LDmanHep	●	Kdo	●	Dha
●	DDmanHep	●	MurNAc	●	MurNGc	●	Mur

**Figure 1.4** An updated symbolic representation of monosaccharides **a)** the general symbolic representation devised in 2015 and **b)** the assignment of symbols to monosaccharides that were unknown upon publication but that have recently been characterised.

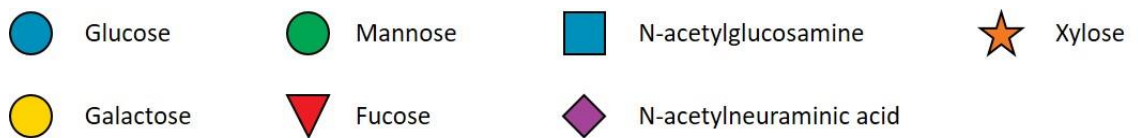
Ongoing modifications are being made to the CFG symbols as required to expand the different monosaccharides represented by symbols that were originally annotated as unknown, such as to include rare non-mammalian monosaccharides (**Figure 1.4b**). Additionally as this system was designed for representing complex glycans an outline was also proposed for detailing the type and position of bonding between sugars. The orientation of the symbols relative to each other can be used to depict the binding site or writing within the bond can be used (**Figure 1.5**).<sup>14</sup>



N-glycan



Plant glycan



**Figure 1.5** Glycan structure examples using the CFG standardised system in either the written format or with the monosaccharides represented by symbols, with or without additional bond information for **a)** an example *N*-glycan and **b)** an example plant glycan.



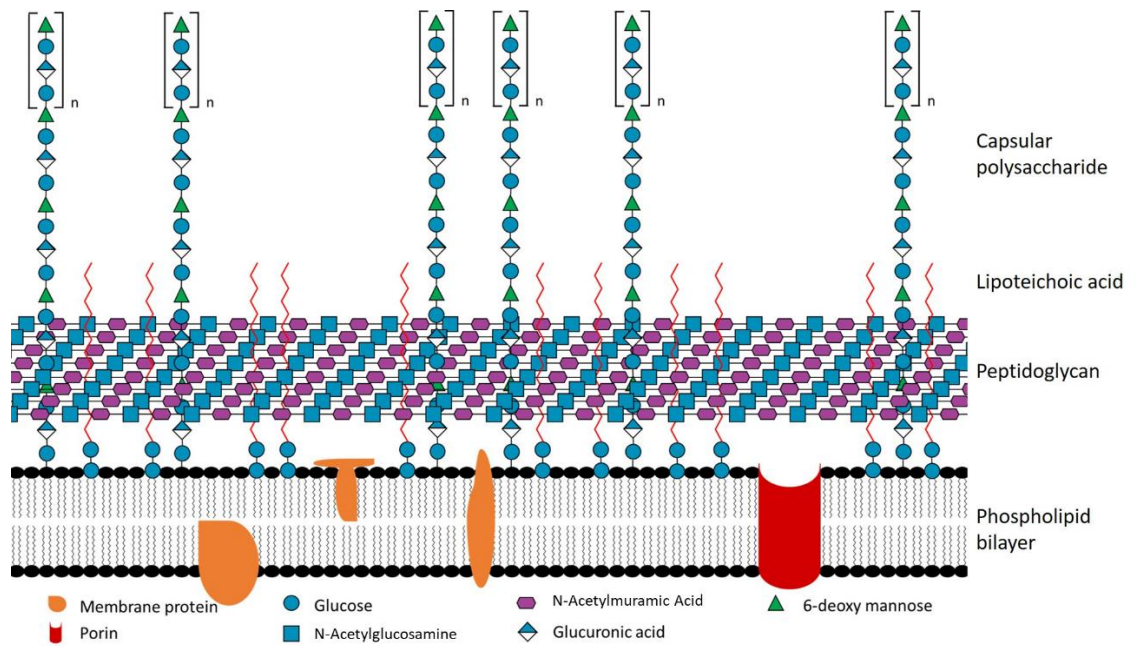
### 1.1.3 Bacterial glycans

Bacteria also utilise glycans in a number of different cell surface structures for protection, recognition and to regulate permeability. The most exposed layer of bacterial cells are decorated with carbohydrates in different forms; the capsular polysaccharide (CPS), lipopolysaccharide (LPS), peptidoglycan and on flagellin and pili. Therefore these carbohydrate structures have extracellular interactions that are essential for pathogenicity and virulence.<sup>15</sup>

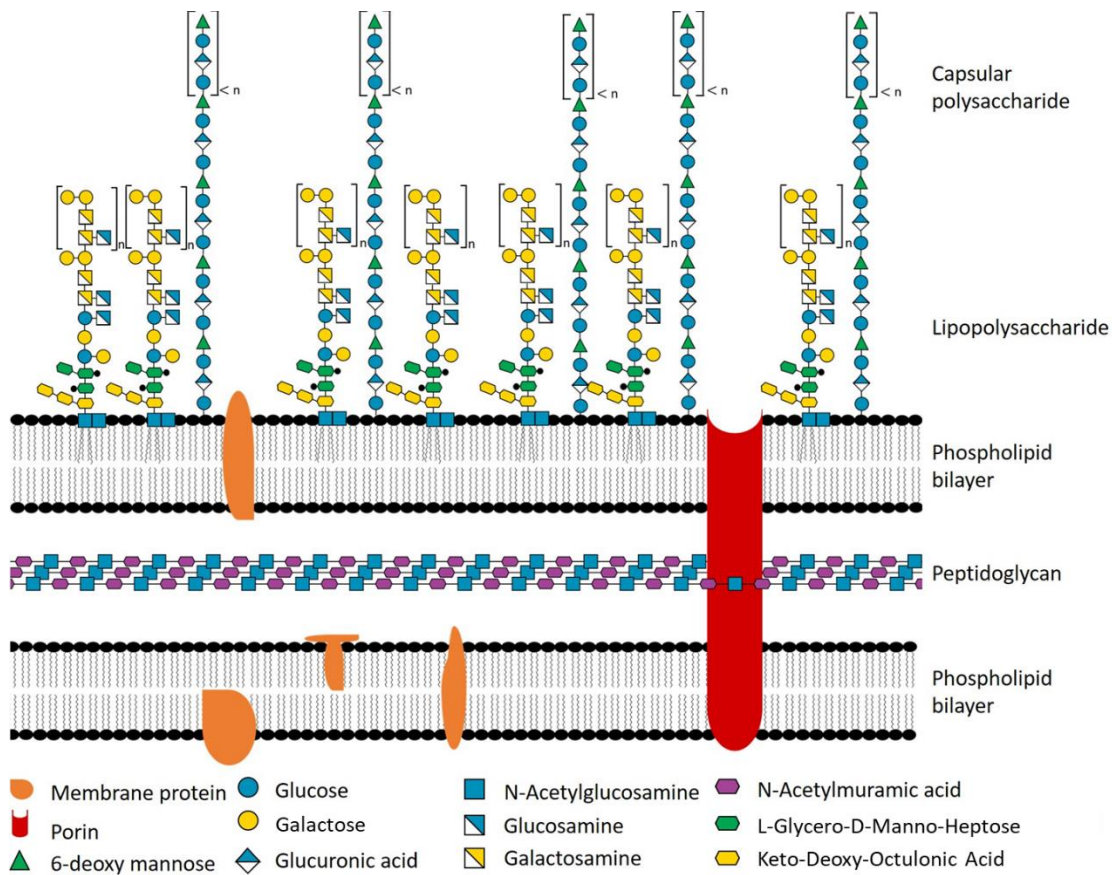
According to their different glycocalyx structures bacteria can be categorised into two groups; Gram positive or Gram negative, and are characterised as such based on whether they retain the Gram stain.<sup>16</sup> Gram positive bacteria have a thick peptidoglycan layer and hence retain the crystal violet stain to a much higher degree than Gram negative bacteria that have a thinner, less accessible peptidoglycan layer. To confirm the presence of Gram negative bacteria, after washing away the crystal violet, a positively charged secondary stain, such as safranin, is applied and binds to cell membranes. This stain will also bind to Gram positive bacteria but be overpowered by the crystal violet stain allowing for the two groups of bacteria to be distinguished by a difference in colour.

Gram positive bacteria (**Figure 1.6**) have a phospholipid bilayer cell membrane littered with permeability mediating proteins and receptor proteins which can span the bilayer. They also have a thick peptidoglycan cell wall that can account for 90 % of their dry weight,<sup>17</sup> composed of polysaccharide chains of N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) cross linked with peptides to form a lattice.<sup>18</sup> This cell wall further mediates permeability and aids in preventing the cell from bursting due to osmosis. Teichoic acids span the peptidoglycan layer and consist of poly(alditolphosphates), which can be anchored to the membrane through disaccharides modified with lipids (lipoteichoic acid). The teichoic acids can bind to host receptors and initiate immune responses however the exact molecular interactions for binding are currently unknown.<sup>19</sup>

Gram negative bacteria also have a phospholipid bilayer cell membrane and peptidoglycan layer<sup>20</sup> (**Figure 1.7**) however the latter is much thinner than in Gram positive bacteria. This decrease of structural rigidity from the peptidoglycan is compensated for by a secondary phospholipid bilayer that can be heavily functionalised with lipopolysaccharides (LPS). The LPS has roles in cell signalling,<sup>21</sup> bacterial toxicity<sup>22</sup> and protection<sup>23</sup> and is exposed on the cell surface in non-capsulated Gram negative strains. Interactions between LPS molecules and cations allows for aggregation<sup>24</sup> and prevents diffusion of small hydrophobic compounds to the cell surface. Upon infection the LPS allows for bacterial survival in harsh conditions<sup>25</sup> and LPS-defective mutants have been shown to be eradicated by a number of antibiotics.<sup>26</sup>



**Figure 1.6** Gram positive bacteria cell wall with the thick layer of peptidoglycan surrounding a phospholipid bilayer.



**Figure 1.7** Gram negative bacteria cell wall with the thin layer of peptidoglycan between two phospholipid bilayers with an outer layer of lipopolysaccharides.

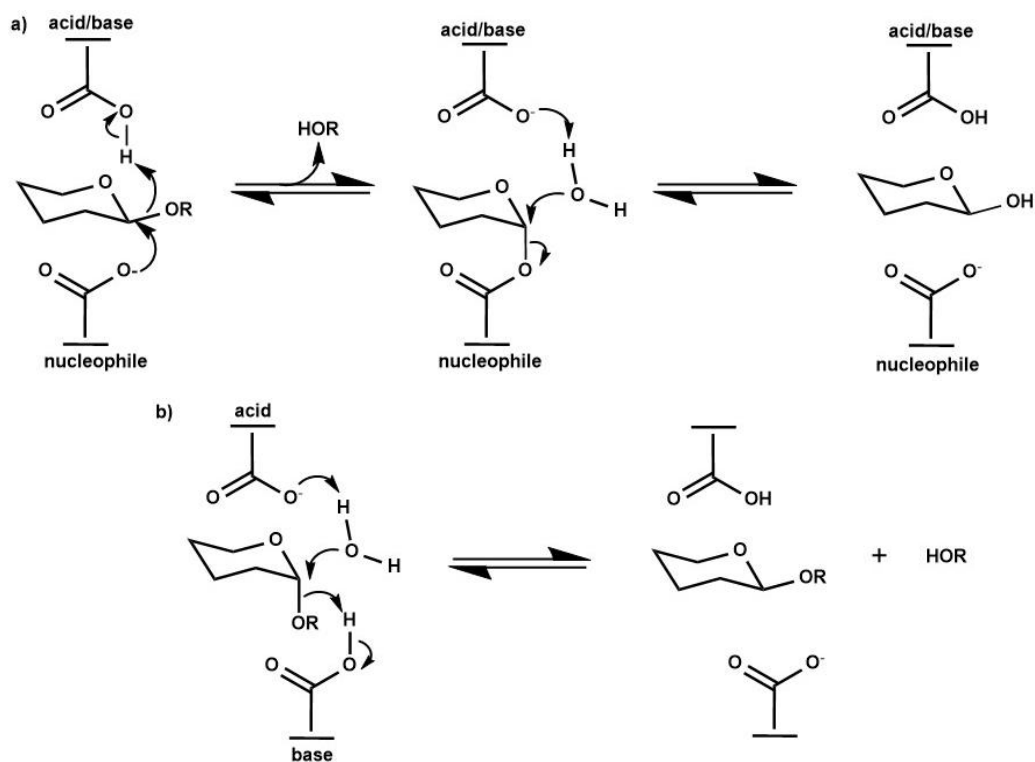
## 1.2 Glycosyl hydrolases

### 1.2.1 An overview of glycosyl hydrolases

The ever-emerging biological importance of glycans and the vast array of glycoconjugates has highlighted a plethora of enzymes that are required for the synthesis and regulation of such structures. One class of glycosyl processing enzymes are the glycosyl hydrolases (GHs) which catalyse hydrolysis of a glycosidic bond between two carbohydrates or between a carbohydrate and another moiety.<sup>6</sup> Glycosyl cleavage is an important biological process greatly influencing cellular regulation processes such as differentiation, adhesion, development, immune responses, and infection.

A recent surge in the assignment of GH encoding genes necessitated the requirement for a classification system and led to the formation of the CAZy (Carbohydrate Active Enzymes) database.<sup>27</sup> This database categorises GHs into different families based on their sequence similarity. Identification and characterisation of GHs has promoted their utilisation in a multitude of practical applications such as; therapeutic targets, biomass conversion, and even as synthetic catalysts through transglycosylation (reverse hydrolysis). The diversity of substrates and hydrolysis processes required is reflected in the number of GH structures, which display specificity to the glycosyl donor.<sup>28</sup>

GHs can also be categorised into retaining GHs or inverting GHs based on the stereochemistry at the anomeric centre of their products relative to the stereochemistry of the substrate.<sup>29</sup> Hydrolysis with retention generally occurs *via* a double displacement mechanism with the two catalytic residues consistently being held 5.5 Å apart.<sup>29</sup> Glutamic acid or Aspartic acid are generally employed as the acidic residue that assists departure of the alkoxide leaving group whilst a nucleophilic residue (commonly a carboxyl group) attacks the anomeric centre. An active site water molecule is then activated by the general base to attack the anomeric centre, releasing the hydrolysed product with retention of stereochemistry (**Scheme 1.2a**). When the catalytic residues are held further apart,<sup>29</sup> a concerted line of attack occurs with direct displacement of the leaving group with a water molecule. An acidic residue activates a water molecule at the anomeric carbon whilst a basic residue donates a proton to promote removal of the leaving group giving rise to inversion at the anomeric carbon (**Scheme 1.2b**).



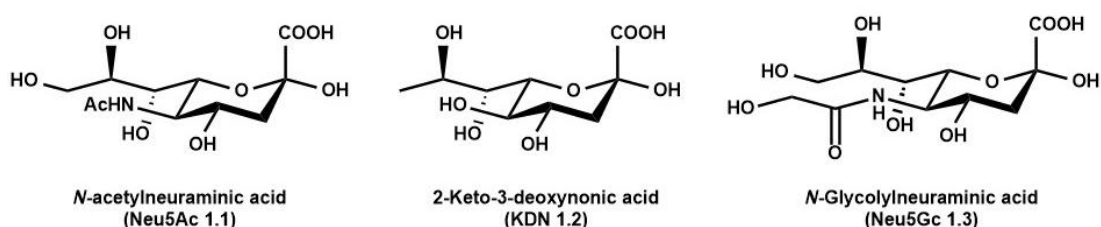
**Scheme 1.2** General glycosidase hydrolysis mechanisms resulting in **a)** retention or **b)** inversion of the configuration of the anomeric bond.

GHs have been identified as particularly useful for biotechnological applications as they catalyse the degradation and recycling process of cellular structures. GHs have been explored as catalysts for the production of bioethanol from biomass as a potential substitute to gasoline<sup>30</sup> and can also provide a route to synthetically challenging small molecules, whereby a careful concoction can catalyse selective degradation of abundant materials into the desired small molecules. These molecules can then be further processed into materials of high value from cheap readily available reagents.<sup>31</sup>

Additionally some GHs have lower substrate specificity and can be used more crudely in place of non-selective chemical degradation processes. For example cellulase degradation activity of a number of xylanases has been explored for utilisation during the paper recycling process as an alternative to the harsh chemical conditions otherwise required.<sup>32</sup> Instead of screening an extensive list of different GH enzymes, detailed structural and mechanistic studies of a few xylanases provided insight into their substrate specificity and highlighted mutations that could be made to alter this specificity as desired as well as enhancing activity.<sup>33</sup>

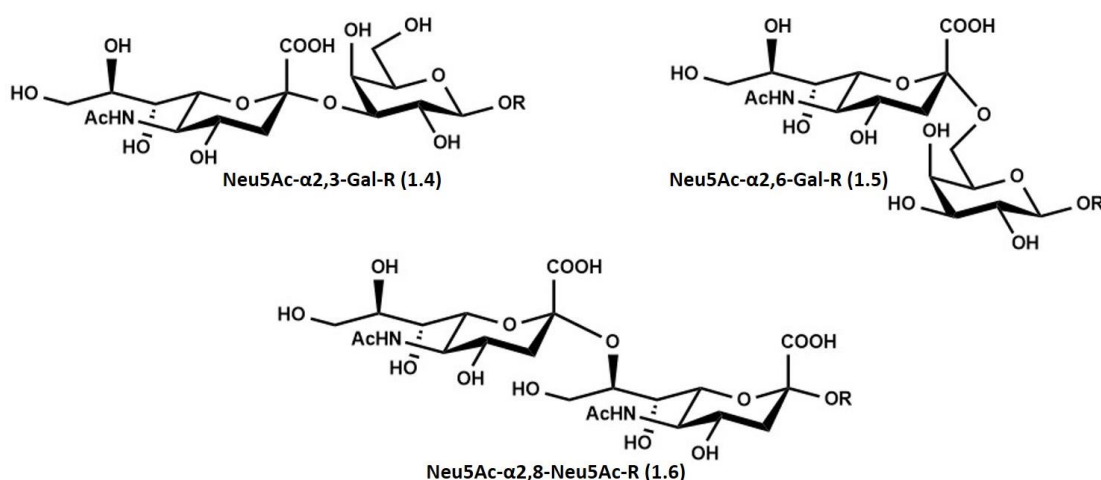
### 1.2.2 Neuraminidase structure and mechanisms

Nonulosonic acids, or sialic acids as they are commonly referred to, (**Figure 1.8**) are often displayed as the terminal sugar on cell surface structures expressed in higher eukaryotes, viruses and bacteria.<sup>13</sup> Sialic acids have importance in a range of biological functions including embryonic development,<sup>34</sup> signalling<sup>35</sup> and adhesion.<sup>36</sup> In nature they are present as a diversely functionalised class of nine carbon containing carbohydrates, the most prevalent of which is *N*-acetylneuraminic acid (Neu5Ac **1.1**).<sup>37</sup> Diversity arises mainly from the incorporation of different functionality at C4, C5, C7, C8 and C9,<sup>38</sup> for example Neu5Ac (**1.1**) has a C5 acetamido group whereas in 2-Keto-3-deoxynonic acid (KDN **1.2**) it is replaced by a hydroxyl and *N*-glycolylneuraminic acid (Neu5Gc **1.3**) has a modified C5 acetamido group with the addition of an hydroxyl (**Figure 1.8**).



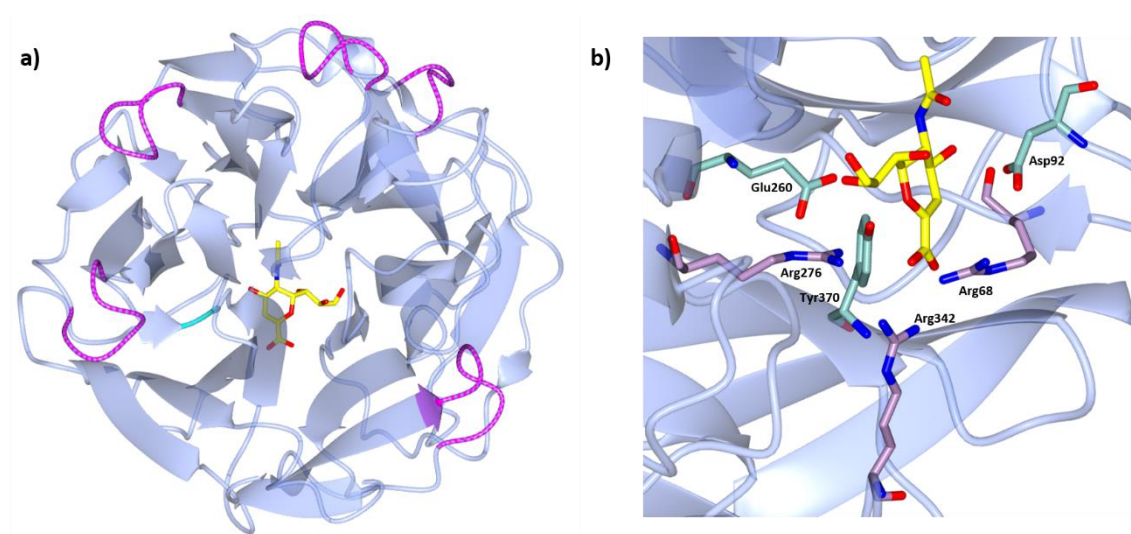
**Figure 1.8** Examples of sialic acid structures prevalent in nature with different C5 functionality.

Sialic acids, such as Neu5Ac (**1.1**), are routinely found linked in the alpha conformation predominantly *via*  $\alpha$ 2-3,  $\alpha$ 2-6 or  $\alpha$ 2-8 linkages (**1.4-1.6**)<sup>39</sup> (**Figure 1.9**). Sialidases/neuraminidases are utilised to specifically cleave the glycosidic bond between sialic acids and other moieties. These enzymes make up an extremely important class of glycosidases and mammalian, viral and bacterial neuraminidases have been discovered for cleavage of both native and non-self sialic acids highlighting their importance in symbiosis and infection.



**Figure 1.9** Different natural Neu5Ac (**1.1**) linkages;  $\alpha$ 2,3 (**1.4**),  $\alpha$ 2,6 (**1.5**), and  $\alpha$ 2,8 (**1.6**).

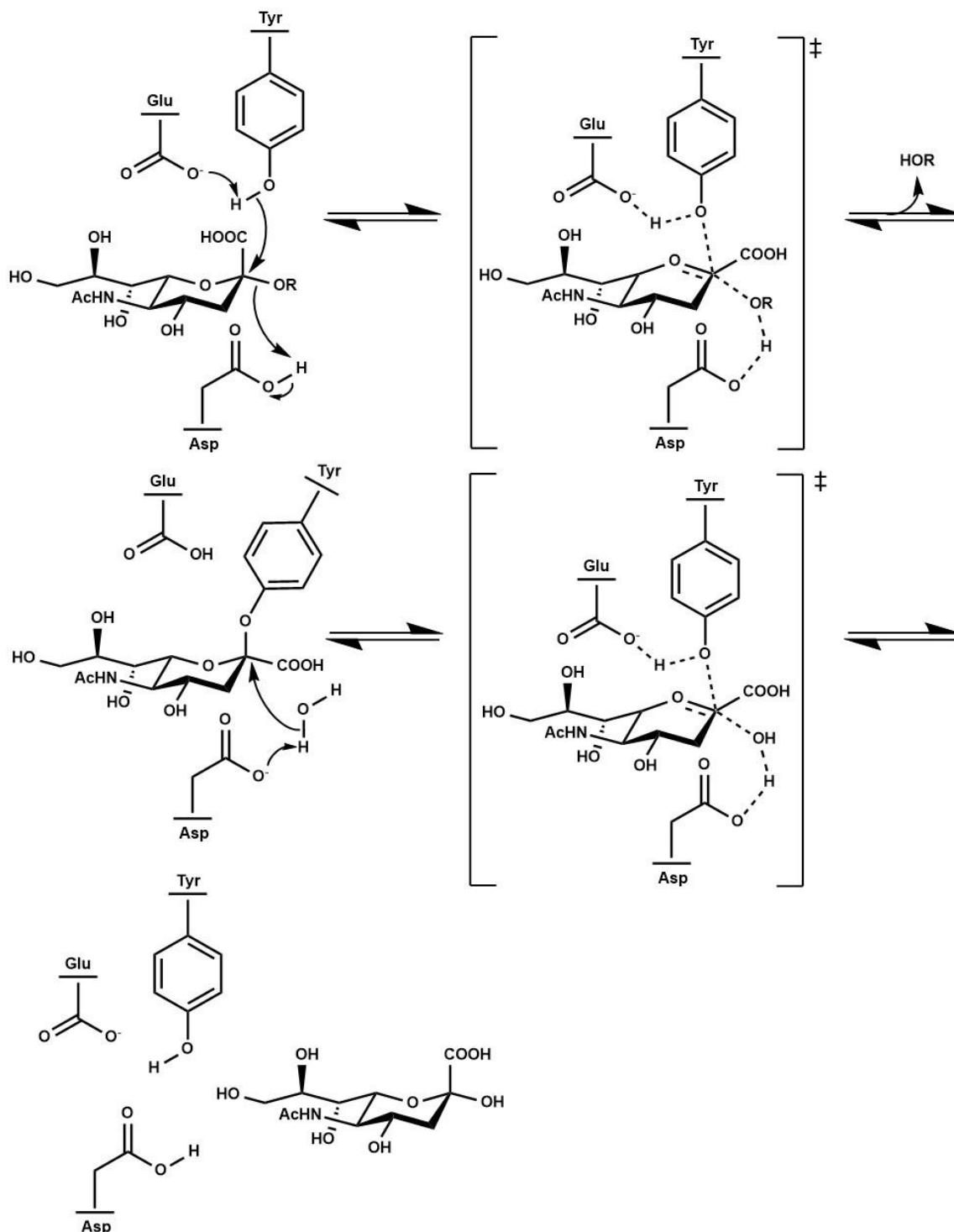
Based on their sequence homology neuraminidases are classified into five different families GH33, GH34, GH83 (*exo*-sialidases) or GH58 (*endo*-sialidases) using the CAZy system.<sup>9</sup> To date, all *exo*-sialidases investigated have been found to hydrolyse sialic acid with a retention of stereochemistry.<sup>40</sup> However recent research investigating an *endo*-sialidase from *Escherichia coli* (*E. coli*) found that this enzyme hydrolysed sialic acid polymers with an inversion of stereochemistry.<sup>41</sup> Although they belong to a number of different GH families the catalytic domain of *exo*-sialidases has been found to be the highly conserved six-bladed  $\beta$ -propeller (**Figure 1.10a**).<sup>42</sup> Although neuraminidase sequences have a low overall homology, there are a number of characteristic motifs that are conserved across all neuraminidases.<sup>43</sup> For example all neuraminidases encode between three and five Asp boxes<sup>43</sup> – a sequence motif that gives rise to  $\beta$ -hairpin folds in topologically equivalent locations. The Asp boxes are thought to be too remote from the active site to be directly involved in the cleaving of glycosidic bonds, however they are important for the protein 3-D structure (**Figure 1.10a**).<sup>1</sup> A second motif is also conserved across neuraminidases; a RIP/RLP sequence with the arginine configured so that it protrudes into the active site, aiding substrate binding *via* interactions with the sialic acid carboxylate group (**Figure 1.10a**).<sup>43</sup> This arginine makes up one of three that are all situated in the active site to stabilise the sialic acid carboxylate group (**Figure 1.10b**).<sup>44</sup>



**Figure 1.10** Crystal structure of a typical sialidase (PDB 1EUS)<sup>1</sup> with **a)** the conserved Asp boxes (magenta) and RIP motif (cyan) highlighted in the full structure and, **b)** the active site with catalytic residues (sea green) and three interacting arginine residues (lilac) highlighted.

Neuraminidases cleave Neu5Ac (**1.1**) with retention of the anomeric stereochemistry following the general double displacement mechanism described previously (**Scheme 1.3**). There is a conserved active site aspartic acid that is typical of GHs that first acts as an acid and then as a base as the reaction proceeds.<sup>45</sup> However, neuraminidases are unique amongst GHs as the

enzymes do not employ a carboxylic acid based residue as the nucleophile but instead use a conserved glutamate-activated tyrosine (**Figure 1.10b**).<sup>46</sup> It has been proposed that the use of tyrosine reduces Coulombic repulsion that occurs between the catalytic residue and the carboxylate group of the sialic acid.<sup>47</sup>



**Scheme 1.3** A typical retaining sialidase double displacement mechanism, using an  $\alpha$ -linked Neu5Ac (**1.1**) as an example substrate.

### 1.2.3 Human neuraminidases

Humans express four different neuraminidases (NEU1, NEU2, NEU3, and NEU4) which are essential for development and abnormal functioning of these enzymes has been linked to genetic diseases and cancer.<sup>48</sup> The regulation of Neu5Ac (**1.1**) on human cells is important for many biological processes including; cell differentiation,<sup>49</sup> antigen masking and apoptosis.<sup>50</sup> Neuraminidases can also modulate the recycling of glycans preventing the need for recurrent *de novo* biosyntheses of certain glycoconjugates.<sup>51</sup>

The requirement of four different human sialidases highlights the importance and widely distributed uses of cellular sialylation in humans. Each of the sialidases differs in activity conditions, substrate specificity and cellular localisation; for example, NEU1, NEU3 and NEU4 display optimal activity at acidic pH whereby NEU2 cleaves Neu5Ac (**1.1**) at near neutral.<sup>52</sup> NEU3 displays the highest substrate specificity with Neu5Ac (**1.1**) cleavage limited to specific gangliosides<sup>53</sup> whereby the others display activity to a variety of glycoconjugates. NEU2 is dispersed within the cytoplasm,<sup>54</sup> whereas NEU1 and NEU3 are predominantly localised at the cell edges in lysosomes<sup>55</sup> or associated with the plasma membrane<sup>56</sup> respectively. NEU4 has been found in a number of organelles and research shows it to predominantly expressed in the liver, with detectable levels in the kidney, heart and brain.<sup>57</sup>

The impact of Neu5Ac (**1.1**) cleavage can be tentatively summarised, however, this process is involved in complex pathways where the chronology of cause and effects are difficult to assign and the pathway molecular mechanisms are not fully understood. The majority of research has been undertaken on NEU1 as this enzyme is the most highly expressed and functioning. However research surrounding NEU3 is slightly less complex as it is more specific, and only directly contributes towards the physiological function, and catabolism, of gangliosides.<sup>53</sup> Gangliosidic Neu5Ac (**1.1**) cleavage can still affect a cascade of many different biological processes, for example, Neu5Ac (**1.1**) hydrolysis is involved in mediating release of ceramide and sphingosine<sup>58</sup> which can act as signalling and regulatory molecules, as well as being recycled for glycolipid biosynthesis.

At the lysosome, NEU1 catalysed cleavage of Neu5Ac (**1.1**) signals for further glycoconjugate degradation. NEU1 is implicated in the complex machinery for regulating lysosomal exocytosis,<sup>59</sup> involving lysosomal docking to the plasma membrane. This mechanism has an important role in membrane repair as well as pathogenic removal through the cellular immune response.<sup>60</sup> NEU1 is also integral to regulation of Neu5Ac (**1.1**) on cell surface glycoconjugates. In this way it has also been shown to be an important signalling regulator for biological processes such as cell adhesion,<sup>61</sup> elastogenesis<sup>62</sup> and phagocytosis.<sup>63</sup>



NEU1 has also been implicated in the immune response with inhibition of NEU1 resulting in a reduced ability for macrophages to produce cytokines and engulf bacteria. This is partially attributed to NEU1 (and NEU3) catalysed desialylation of monocytes allowing them to be more sensitive to bacterial LPS, inducing production of specific cytokines.<sup>64</sup> Additionally it has been shown that NEU1 expression is significantly increased during lymphocyte activation with a higher concentration of NEU1 displayed on the cell surfaces. This increase in NEU1, in turn, has been shown to increase production of interleukin-4 which can be critical for immunity against some viral infections.<sup>65</sup>

Currently, NEU1 is the only human neuraminidase whose mutations have been directly linked with a genetic disease-sialidosis. This is a lysosomal storage disease caused by excessive storage of glycoconjugates in the lysosome due to a lack of desialylation signalling further degradation of these molecules.<sup>55</sup> Type I sialidosis doesn't usually present until a patient's second or third decade but can quickly progress and can cause myoclonus, seizures and visual defects associated with "cherry-red spot".<sup>66</sup> Treatment is patient centred with emphasis placed on managing the symptoms rather than regulation of sialylation as therapeutics for this do not currently exist.<sup>67</sup> Generally Type II sialidosis is more severe than Type I and is characterised by the early onset, mucopolysaccharidosis-like phenotype and, in most cases, a drastically reduced life expectancy.<sup>68</sup> For example, severe pre-birth deficiencies in NEU1 resulting in Type II sialidosis can cause sufferers to be still born or die soon after birth with excessive sialylation leading to a number of health complications such as an enlarged spleen and liver, fluid accumulation and abnormal bone development.<sup>69</sup>

Considering the impact of glycoconjugates in the hallmarks of cancer, it is unsurprising that tumorous cells display an altered cell surface fingerprint compared to non-cancerous cells.<sup>70</sup> One of these fundamental differences, is an overall increase in the amount of Neu5Ac (**1.1**) present on cancer cell surface glycoconjugates. Not only does the hypersialylation increase siglec binding hence aiding in immune response evasion but it can also enhance tumour progression by regulating a number of mechanisms.<sup>71</sup> Neuraminidase expression levels in tumorous cells have also been found to be abnormal<sup>52</sup> and influence the cells ability for increased cell proliferation and cell death resistance.<sup>72</sup> For example, NEU3 in cancer colon has an up-regulated expression; increasing ganglioside de-sialylation and revealing glycolipid structures that are important for reducing apoptosis.<sup>73</sup>

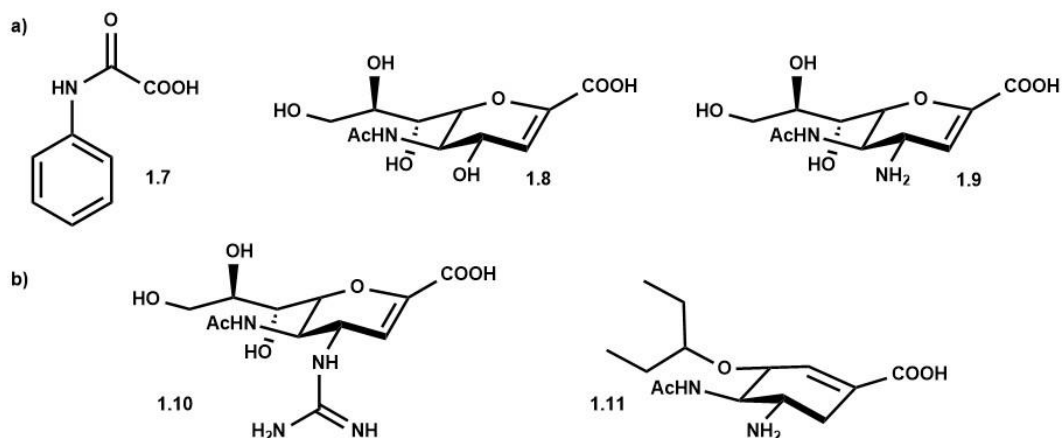
#### 1.2.4 Neuraminidases as therapeutic targets

Bacteria and viruses also utilise neuraminidases for their own biological processes and to aid in the invasion of hosts. In particular neuraminidase activity has been found to enable increased binding between bacteria and host epithelial cells, although the molecular interactions of the resulting binding are unknown.<sup>74</sup> For example, pre-treatment of human cells with a viral neuraminidase has been shown to allow for an increased adhesion by a number of different bacterial cells including *Streptococcus pneumoniae* (*S. pneumoniae*).<sup>75</sup> As well as displaying importance during initial colonisation, *S. pneumoniae* nanA was found to be important for the longevity of infection both in the respiratory tract and blood samples.<sup>76</sup> The cell surface neuraminidase in *Capnocytophaga canimorsus* (*C. canimorsus*) has also been shown to be essential for bacterial persistence by de-glycosylating host Neu5Ac (**1.1**) to reveal glycoproteins that could act as a source of carbon and nitrogen to sustain the bacteria.<sup>77</sup>

One of the most exclusively studied types of neuraminidase includes those displayed on the outer surface of the influenza virus. This enzyme adopts the canonical neuraminidase fold (six-bladed  $\beta$ -propeller) and has the strictly conserved neuraminidase catalytic residues, however homology across the rest of the sequence with other neuraminidases is low.<sup>78</sup> To initiate infection, the virus needs to be able to pass along the respiratory tract, however the host has evolved a protective barrier in the form of a mucus layer containing Neu5Ac (**1.1**) glycoproteins. Another viral surface protein, hemagglutinin, binds to host glycoproteins to initiate infection and then neuraminidase activity is required at this stage to allow the virus to progress along the epithelial cells.<sup>79</sup> The neuraminidase is also employed following viral replication to release the virions from the viral surface to prevent viral aggregation and promote dispersion.<sup>80</sup> The neuraminidase activity has been shown to be essential for effective influenza viral infections and hence has been a target for numerous therapeutics.

To aid drug design Lipinski suggested four objectives that if a molecule adhered to, it should be well absorbed by the body; < 5 hydrogen bond donors, < 10 hydrogen bond acceptors, molecular weight < 500 g mol<sup>-1</sup> and calculated log P < 5.<sup>81</sup> It was deemed that if a molecule met at least three of these requirements then it could be an orally administered drug based on its predicted solubility and permeability,<sup>82</sup> however many exceptions to the rule have occurred. Following these guidelines, potential Neu5Ac-based neuraminidase inhibitors are attractive as Neu5Ac (**1.1**) meets the requirements; has 9 hydrogen bond acceptors, a molecular weight of 309 g mol<sup>-1</sup> and calculated log P of -3.5. Therefore putative inhibitor analogues display potential to be well-absorbed, orally administered drugs.

*N*-substituted oxamic acids, such as *N*-phenyloxamic acid (**1.7**) (**Figure 1.11**), were identified as prudent neuraminidase inhibitors in the 1960s.<sup>83</sup> However these molecules lost all activity during *in vivo* experiments and displayed a lack of enzymatic specificity so were deemed unsuitable as therapeutics.<sup>84</sup> Subsequent research focussed on developing inhibitors that had higher selectivity for neuraminidases and, preferentially targeting the viral, and/or bacterial enzymes over the host (mammalian) neuraminidases.



**Figure 1.11** A selection of neuraminidase inhibiting structures; **a)** Neuraminidase inhibitors and **b)** existing anti-viral therapeutics.

Insight into the double displacement neuraminidase mechanism with the predicted formation of oxocarbenium ion transition states<sup>29</sup> (**Scheme 1.3**) allowed for the mechanism based design of a transition state analogue; 2-deoxy-2,3-dehydro-*N*-acetylneuraminic acid (Neu5Ac2en) (**1.8**) as a potential inhibitor (**Figure 1.11a**).<sup>84</sup> Although inhibition ability was improved from previous candidates, this was true for mammalian and pathogenic neuraminidases alike and hence fine-tuning of this molecule was required in order to gain the desired neuraminidase selectivity.<sup>85</sup> Crystal structures of the influenza neuraminidase<sup>86</sup> enabled structure based drug design and comparison of neuraminidase active sites allowed for discovery of selective inhibitors of pathogenic neuraminidases.<sup>85</sup> For example, 4-amino-Neu5Ac2en (**1.9**) binds with a stronger affinity to influenza A neuraminidase compared to neuraminidases from other species (**Figure 1.11a**).<sup>85</sup> A number of drugs based on the Neu5Ac2en (**1.8**) inhibitor structure, are commercially available such as Zanamivir<sup>87</sup> (**1.10**) which simply has the C4 hydroxyl substituted with a guanidine group, and Oseltamivir<sup>88</sup> (active metabolite **1.11**) whose structure varies more greatly from Neu5Ac2en (**1.8**) (**Figure 1.11b**). Oseltamivir (active metabolite **1.11**) importantly still mimics the neuraminidase transition state through its  $sp^2$  appended carboxyl and a C5 acetamido group.

## 1.3 *Pseudomonas aeruginosa*

### 1.3.1 An overview of *Pseudomonas aeruginosa*

*Pseudomonas aeruginosa* (*P. aeruginosa*) is a versatile, gram negative, opportunistic bacteria. It is ubiquitous inhabiting terrestrial, aquatic, animal, human and plant environments alike and thrives on moist surfaces and predominantly invokes illness in immuno-compromised humans or those with loss of a physical barrier to infection e.g burns or lesions. *P. aeruginosa* colonisation can result in a number of different infections such as urinary tract infections,<sup>89</sup> respiratory system infections,<sup>90</sup> and a variety of systemic infections depending on the site of affliction.<sup>91</sup> This bacteria is generally considered an aerobic pathogen,<sup>92</sup> however reports suggest that it can adapt to microaerobic or anaerobic growth if required.<sup>93</sup> For the majority of healthy individuals *P. aeruginosa* cannot establish an infection and the immune system can overcome minor infections. However for immunocompromised individuals, particularly those in hospital, this infection can manifest and cause considerable illness. *P. aeruginosa* has also been shown to affect sufferers of cystic fibrosis (CF) with the bacteria displaying many features that contribute to significant health issues and even mortality.

*Pseudomonas* was first recorded in the 1860s when Fordos investigated a blue-green stain on surgical dressings and numerous other hospitalised articles<sup>94</sup> and Lucke discovered its association with a rod-shaped organism.<sup>95</sup> This pigment was later assigned as pyocyanin, a toxin released by bacteria such as *P. aeruginosa*.<sup>96</sup> Although full pathogenesis is still yet to be detailed it is thought to be multifactorial and a number of virulence factors have been identified: production of extracellular proteases for adherence and invasion of host cells, binding to sugars on host epithelial cells, toxin secretion, and biofilm formation. Toxin production is thought to stimulate the symptoms displayed by patients including fever, low blood pressure, reduced urine production, and confusion. Site specific symptoms can also occur such as the development of pneumonia in infected lungs.

Research suggests that *P. aeruginosa* survival centres around the bacteria having two different phases of infection; an initial acute phase whereby the bacteria become established, before major modifications and diversification of these colonies to produce a chronic infection.<sup>97</sup> This complicates treatment designs as acute infection is difficult to detect using available methods and upon entering the chronic phase the bacteria are much less susceptible to antibiotic attack.

### 1.3.2 Pseudomonas aeruginosa infection

*P. aeruginosa* infections are almost always associated with hosts who have a compromised immune defense, for example patients undergoing chemotherapy.<sup>98</sup> Although it can colonise multiple environmental niches, *P. aeruginosa* most commonly manifests in the lungs.<sup>99</sup> During acute infection, *P. aeruginosa* have high expression of virulence factors such as the flagella and secretion of toxins,<sup>100</sup> for example, production of pyocyanin protects the bacteria from the immune system by triggering neutrophil apoptosis.<sup>101</sup> Isolates display a metabolic and mutational rate similar to that of *P. aeruginosa* isolates from inanimates and are more sensitive to antibiotics and immune defences compared to the chronic phase.<sup>102</sup> It is apparent that during acute infection the bacteria's main purpose is to establish an infection of the lungs and to regulate gene expression to adapt to their new environment.

The infected environment creates a pressure over *P. aeruginosa* to select for an adapted population leading to diverse colonies which vary greatly from the bacteria involved in early onset. These typically have a reduction of virulence factors, are mucoid and mutations are selected to reduce quorum sensing<sup>103</sup> and increase alginate biosynthesis.<sup>104</sup> It is these adaptations to the environment (further discussed below) that promotes the permanent survival of *P. aeruginosa*. The extreme diversification of the bacteria during chronic infection renders it virtually impossible to eradicate, as a concoction of complex mechanisms would be required to eliminate all of the bacteria.

**Biofilm production;** during chronic infection there is an increase in the production of an extracellular matrix predominantly composed of the negatively charged polysaccharide - alginate.<sup>105</sup> This is achieved via a mutation of *mucA*, a gene that expresses a cytoplasmic membrane bound protein which acts as an anti  $\sigma$  factor,<sup>106</sup> blocking transcription of *algT* which encodes a protein required for expression of the alginate operon.<sup>107</sup> Mutations occur preventing *mucA* from binding allowing for an increased expression of  $\sigma^{22}$  and consequently *algT* and alginate.<sup>108</sup> Subsequently, the formation of biofilms provides a mucus layer around the bacteria, protecting it from antibiotic attack and immune defences.<sup>109</sup> Biofilms also allow for greater bacterial diversification which further increases multi-antibiotic resistance. They also aid in dispersal events of non-mucoid colonies resulting in the spread of *P. aeruginosa* to new respiratory niches.<sup>110</sup>

**Reduction of virulence factors;** chronic *P. aeruginosa* isolates are often characterised as lacking in pili and flagella thus allowing them to evade some host immune defences.<sup>111</sup> Additionally, these appendages are of little use after biofilm production whereby the matrix prevents

significant movement and so mutations reducing their production prevents a mass energy waste.<sup>112</sup>

**Mutators;** alongside variability gained from acute infections diversity is also achieved in chronic infections by mutators which occur across chronic populations of *P. aeruginosa*.<sup>112</sup> Mutators can mutate up to 1000 times faster than non-mutators and as such promote survival against unpredictable stress factors that often occur in the cystic fibrosis lung environment.<sup>113</sup> Additionally the ability of *P. aeruginosa* to hypermutate promotes the occurrence of sub-populations with different phenotypes thus allowing for a potential increase in multi-antibiotic resistance.<sup>114</sup>

**Altered metabolic pathways;** although *P. aeruginosa* maintains core metabolism pathways throughout infection, isolates from the chronic stage also displayed an adaptation strategy. Induction of substrate-specific metabolic pathways occurs allowing for growth in limited nutrient environments.<sup>115</sup> During the chronic phase, carbon metabolism is regulated by catabolite repression control, whereby short chain fatty acids, amino acids and polyamines are preferentially metabolised over carbohydrates.<sup>116</sup> In cystic fibrosis the sputum often contains increased amounts of DNA, lipids, mucin and amino acids that can be used to support growth after bacterial manipulation. *P. aeruginosa* has been shown to up-take these nutrients from the sputum and although their uses are unclear, high levels of auxotrophs are often detected in isolates.<sup>117</sup>

**Production of energy under microaerobic or anaerobic conditions;** *P. aeruginosa* is classified as an aerobic bacteria, employing oxygen as the terminal electron acceptor.<sup>118</sup> However, during chronic infection the environment in the lungs can be microaerobic or, as biofilms develop and thicken, completely anaerobic.<sup>100</sup> *P. aeruginosa* is able to survive under these conditions by upregulating different metabolic pathways and regulators and employing a number of different energy producing strategies.<sup>119</sup> For example under anaerobic conditions *P. aeruginosa*:

- Up-regulates production of the nonspecific OprF porin to allow for diffusion of nitrate and nitrite present in the mucus into the cell.<sup>120</sup> In low oxygen environments these compounds can act as alternate electron acceptors thus allowing for energy production via denitrification.
- Increases expression of proteins involved in arginine metabolism to exploit an increased up-take of arginine as a carbon and energy source.<sup>121</sup>
- Can use pyruvate as an energy source in niches that are deficient in the compounds described above that are otherwise employed for energy production.<sup>122</sup>

### 1.3.3 Cystic Fibrosis

Cystic fibrosis (CF) affects around 100,000 people worldwide and in the U.K alone, every week five babies are born with this disease. Sufferers of CF have a highly decreased life expectancy with 39 years being the average age of death across Europe (2008-2009).<sup>123</sup> CF is caused by mutations in the recessive *cystic fibrosis transmembrane conductance regulator (CFTR)* gene which has been shown to encode a protein chloride channel.<sup>124</sup> This channel is located on the apical membrane of epithelial cells thus mutations can disrupt transport of chloride ion into and out of cells.<sup>125</sup> Thus causing these tissues to carry defects in the absorptive and secretory processes producing an abnormal environment across a number of organs<sup>126</sup> and leads to a reduction of efficacy. This is particularly significant in the lungs and pancreas where there is an accumulation of a more viscous mucus due to osmotic abnormalities, leading to a reduction in the delivery of oxygen and digestive enzymes to the blood and digestive system respectively.<sup>127</sup>

Early research regarding CF focussed on identification of the gene responsible for the disease,<sup>128</sup> the protein it encodes<sup>129</sup> and the mutations that give rise to the CF phenotype.<sup>130</sup> Four transcribed sequences thought to contain the CF locus were genetically analysed and a locus on human chromosome 7 was found to contain the CFTR nucleotide sequence.<sup>131</sup> Transcripts of this gene were more prevalent in organs that are directly affected by CF such as the lung and sweat glands but there was no significant difference between the number of transcripts in CF tissues compared to controls. The complete cDNA sequence was identified and an open reading frame encoding a peptide chain of 1480 amino acids, including the initial methionine codon, was suggested.<sup>129</sup> This polypeptide displays two membrane-spanning domains, two nucleotide-binding domains and a regulatory domain. It was found to contain characteristics homologous to P-glycoprotein<sup>132</sup> and thus it was predicted to be an ion transport channel.

Expression of CFTR into cells that did not natively contain it resulted in an increase in chloride conductance when coordinated by ATP<sup>133</sup> and provided evidence for CFTR classification as a chloride channel mediated by cAMP dependent phosphorylation of the regulatory domain. Comparisons of the nucleotide sequence of the CF open reading frame and that from unaffected individuals led to the identification of a three base pair deletion resulting in a loss of the phenylalanine 508 residue ( $\Delta$ Phe-508).<sup>126</sup> Phe508 lies in a sequence highly conserved across similar proteins of the nucleotide-binding domain and its mutation has been shown to result in a loss of secondary structure and hence the majority of  $\Delta$ Phe-508CFTR remains in the endoplasmic reticulum.<sup>134</sup>

Subsequent research concluded that  $\Delta$ Phe-508CFTR is not the only mutation that gives rise to CF and multiple different mutations can occur.<sup>135</sup> However this is by far the most prevalent mutation in cystic fibrosis patients; accounting for about 70 % of cases.<sup>130</sup> These mutations can be categorised according to how they disrupt the normal functioning of CFTR; Class I severely reduces protein production,<sup>136</sup> Class II causes defective folding,<sup>137</sup> Class III produces channels that have defective regulation,<sup>138</sup> Class IV generates CFTR that can be membrane bound but has defective pore properties.<sup>139</sup> Class I and Class II cause the most severe symptoms and Class II is the most common as it hosts the  $\Delta$ Phe-508CFTR mutation.

Up until recently most therapeutics for CF sufferers have targeted management of the symptoms instead of attempting to address the causes of these symptoms. For example, treatments have focussed on attempting to remove lung mucus using a combination of enzymes such as pulmozyme<sup>®140</sup> or hypertonic saline to break down the mucus,<sup>141</sup> bronchodilator drugs<sup>142</sup> to relax airways and physiotherapy treatments involving massaging of the back and chest. Additionally, other enzymatic treatments are prescribed to target affected organs such as the pancreas to aid in regaining proficient absorption of necessary nutrients.<sup>143</sup> However, although these methods can extend life expectancy and reduce discomfort in CF patients, mortality and quality of life is still vastly reduced for sufferers of the disease.

Recent research has instead focussed on directly tackling the genetic mutation and attempting to prevent the detrimental effect this displays on sufferers organs, and in particular their lungs. *In vitro* gene therapy trials proved promising, successfully delivering CFTR cDNA to the airway epithelium and correcting the chloride transport defect.<sup>144</sup> However, this achievement was difficult to replicate *in vivo* with debates occurring over the type of cells to target and the optimal vector to use.

Genome editing has also been explored as a potential treatment of CF whereby a zinc finger nuclease binds to the gene locus nearby the target mutation enabling an endonuclease to cleave the DNA allowing for recombination with a non-mutated DNA template.<sup>145</sup> “Potentiators” have also been developed to regulate the function of the mutated chloride ion channels. This class of therapeutics is designed to increase the time that CFTR channels remain open therefore allowing for the sufficient movement of chloride ions across the membrane.<sup>146</sup> In order to reduce mortality rates in CF sufferers, referrals for lung transplants are considered for patients with advanced CF and results have been encouraging however this is not a viable option in many cases and thus only examined in extremes.<sup>147</sup>



### 1.3.4 *Pseudomonas aeruginosa* infection in CF sufferers

Alongside the issues associated with mutation of the CFTR gene, chronic bacterial infections can also occur that are normally eradicated in non-cystic fibrosis individuals. CFTR has been shown to act as a receptor for *P. aeruginosa* binding it to airway walls to aid removal.<sup>148</sup> However, in its mutated form it is hypothesised that this CFTR-bacterial interaction is disrupted so *P. aeruginosa* removal is consequently reduced. *P. aeruginosa* is the most common lung disease associated with CF and is directly responsible for decreasing the median survival of CF sufferers by over ten years.<sup>149</sup> A number of *P. aeruginosa* strains have been genetically sequenced and analysis showed each to contain a core genome that has about a 70 % sequence identity across strains.<sup>150</sup> Of these, the PA14 strain is the most common worldwide<sup>151</sup> and has high virulence caused by pathogenicity islands that are lacking in less virulent strains such as PAO1.<sup>152</sup> The *P. aeruginosa* Liverpool epidemic strain (LESB58) is the strain most frequently associated with lung infection in CF sufferers<sup>153</sup> and displays a wider antibiotic resistance, enhanced virulence and morbidity compared to other strains such as PAO1.<sup>154</sup> Although *P. aeruginosa* chronic infection has historically been limited to CF sufferers and immunocompromised individuals, it has recently been reported that this LESB58 strain was able to infect and establish chronic colonisation in healthy individuals.<sup>153</sup>

Accumulation of the natural *P. aeruginosa* virulence factors selected for by the pressures of their environment and the issues associated with CF contributes highly to patient mortality and quality of life. For example *P. aeruginosa* alginate adds to the already unusually viscous mucus in the sputum of CF sufferers thus advancing respiratory insufficiency.<sup>155</sup> An additional complication for the lungs of CF sufferers infected with *P. aeruginosa* is the host immune response producing a combination of defences that flood to the lungs. A vicious circle of events exists whereby infection signals for an over-whelming immune response, which produces inflammation and secretion of proteinase enzymes<sup>156</sup> which in turn stimulates mucus production and promotes the *P. aeruginosa* bacterial infection. It is not exactly clear how the defective chloride ion transport associated with CF leads to a heightened susceptibility and longevity of *P. aeruginosa* infection although hypotheses based around the altered lung surface fluid do exist. For example, it has been found that bacterial activity is present in all lung surface fluid but most humans can prevent long-term infection using host immune defences. However, a higher sodium chloride concentration in the lung fluid of cystic fibrosis sufferers results in them lacking the ability to kill some bacteria such as *P. aeruginosa*.<sup>157</sup> It has also been suggested that a lack of diffusion of the surface fluid which normally lines the epithelial cells damages the cilia and hence they lose bacteria removal efficacy.<sup>158</sup>

### 1.3.5 Existing treatment for *Pseudomonas aeruginosa*

As a result of the challenges associated with *P. aeruginosa* infections there has been an increased focus on therapeutics that directly target elimination of it from CF sufferers lungs. Although acute infection is not associated with an immediate decline in lung function, if left untreated the bacteria evolve into the highly damaging chronic phase.<sup>159</sup> Therefore if *P. aeruginosa* infection is detected in the acute phase, aggressive treatment with antimicrobial therapy is used which often results in effective eradication due to the lack of protection available to the bacteria in this phase.<sup>160</sup> *P. aeruginosa* treatment employs a combination of a  $\beta$ -lactam to block synthesis of the bacterial cell wall<sup>161</sup> and an aminoglycoside antibiotic that selectively disturb bacterial protein synthesis.<sup>162</sup>

Although somewhat effective for acute infections, neither of these mechanisms can target *P. aeruginosa* in the mucoid form;  $\beta$ -lactams are only active against rapidly dividing bacteria whereas growth slows in the chronic phase<sup>93</sup> and aminoglycosides require internalisation by bacterial transport mechanisms which have a highly reduced activity during the anaerobic growth that often occurs in the chronic phase.<sup>93</sup> Therefore for such treatments to be successful it is paramount that detection techniques with an increased sensitivity are developed and alternative methods investigated to target *P. aeruginosa* in its chronic phase.

Vaccines for CF sufferers have been proposed as another viable strategy to protect this vulnerable group from *P. aeruginosa* infections as they don't rely on early detection of the infection. Nasal and oral vaccines directed toward OprF and OprI proteins have undergone drug trials in healthy individuals and specific antibodies were produced but only partial protection was achieved.<sup>163</sup> Other research has focussed on directing drugs to target specific *P. aeruginosa* virulence factors. For example, other potential vaccines are under investigation that stimulate the production of flagella binding antibodies.<sup>164</sup>

Another obvious target for therapeutic disruption is biofilm formation as this occurs in conjunction with transition to the chronic phase<sup>165</sup> and is heavily involved in bacterial protection<sup>166</sup> and enabling differentiation.<sup>167</sup> This formation of biofilms, which occurs in some strains during established infections,<sup>168</sup> acts as a physical barrier rendering the bacterium virtually impermeable to antibiotics<sup>169</sup> and creates an internal environment. Therefore therapeutics that prevent *P. aeruginosa* residing in a mucoid state could be essential for enabling treatment of infected individuals, either individually or as part of a multi-drug therapy.

### 1.3.6 PA2794; a potential *Pseudomonas aeruginosa* therapeutic target

Genetic analysis of various strains of *P. aeruginosa* from CF patient isolates and non-CF patient isolates revealed a major distinction to be the expression of *nan1* (coding for a protein termed PA2794). Bacterium collected from CF patients displayed a significantly higher prevalence of *nan1* expression compared to the other isolates<sup>170</sup> and a *nan1* knockout was found to be unable to colonise the respiratory tract.<sup>171</sup> Therefore it was suggested that this gene was influential in the ability of some *P. aeruginosa* strains to colonise the lungs of CF patients. One such role of *nan1* was proposed to be in adhesion as a correlation was found between *P. aeruginosa* isolates with the *nan1* gene and an increased ability to adhere to buccal epithelial cells.<sup>172</sup> Another potential function was identified when knock-out study of the PA2794 gene locus resulted in a reduction in the propensity of the bacteria to form biofilms.<sup>173</sup> This role was further confirmed by an increase in biofilm production following up regulation of the PA2794 locus.<sup>171</sup> However, it is unclear how the predicted function of the PA2794 enzyme relates to its apparent importance to the *P. aeruginosa* pathogenic phenotype.

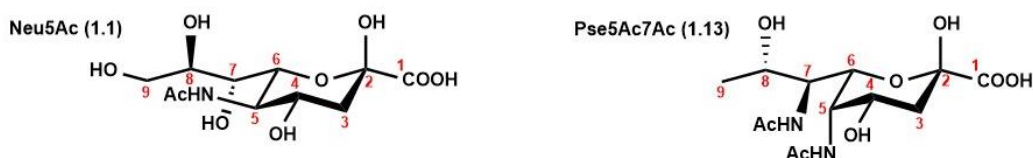
Early investigations of PA2794 resulted in its assignment as a GH and specifically a neuraminidase based on its sequence containing the typical bacterial sialidase motifs.<sup>174</sup> However the importance of Neu5Ac (**1.1**) cleavage in relation to biofilm formation was disputed as Neu5Ac (**1.1**) was known not to be a component of *P. aeruginosa* biofilm.<sup>175</sup> Therefore it was hypothesised that the role of PA2794 in infection could be to catalyse the cleavage of Neu5Ac (**1.1**) on host cells. It was suggested that the extracellular PA2794 enzyme could be involved in the cleavage of mammalian Neu5Ac (**1.1**) exposing a glycan receptor on epithelial cells which the bacteria could bind to during initial infection.<sup>176</sup> However, previous studies have demonstrated that *P. aeruginosa* isolates from cystic fibrosis differed to isolates from other environments in that they did not anchor to the epithelial cells,<sup>177</sup> rendering cleavage of Neu5Ac (**1.1**) for *P. aeruginosa* host binding as unbeneficial for infection. Therefore it was suggested that the natural substrate may be a different glycosidically linked nonulosonic sugar, such as pseudaminic acid (Pse5Ac7Ac, **1.13**).

Although the natural substrate has not been identified, research suggests that this gene may be influential in the propensity for *P. aeruginosa* to colonise the lungs of CF patients, especially during the mucoid phase whereby current treatments are largely ineffective.<sup>93</sup> Therefore this enzyme presents as an ideal drug target and full enzymatic characterisation would aid in rational drug design.

## 1.4 Pseudaminic acids

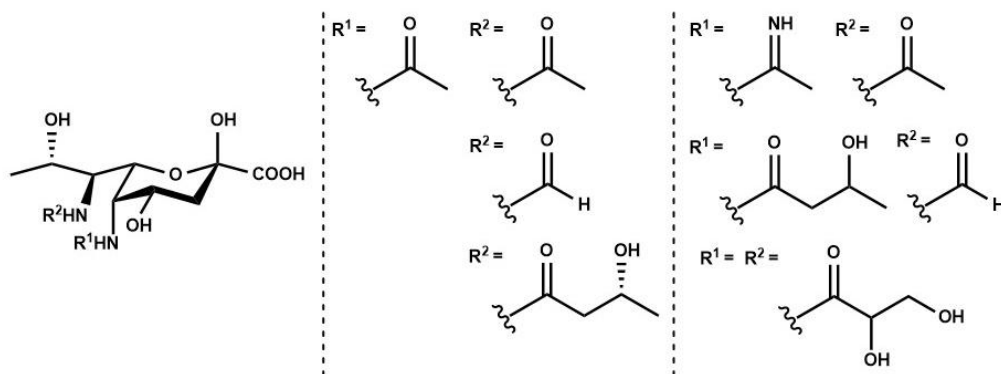
### 1.4.1 Pseudaminic acid structures

Nonulosonic acids can differentiate in their stereochemistry as well as functionality. Pseudaminic acids are rare, non-mammalian nonulosonic acids that belong to this class of carbohydrates, and are epimeric at the C5, C7 and C8 positions when compared to Neu5Ac (**1.1**) stereochemistry (**Figure 1.12**). The general term pseudaminic acid usually relates to the Pse5Ac7Ac (**1.13**) structure which has the most similar functionality to Neu5Ac (**1.1**), except that an acetamido group replaces a hydroxyl at C7 and C9 is de-hydroxylated (**Figure 1.12**). Pse5Ac7Ac (**1.13**) was first discovered as a component of *P. aeruginosa* and *Shigella Boydii* (*S. Boydii*) LPS and NMR used to assign the configuration of this new carbohydrate as  $\iota$ -glycero- $\iota$ -manno.<sup>178</sup> This carbohydrate has since been identified as biosynthesised by a number of gram negative bacteria, including pathogens *Aeromonas caviae* (*A. caviae*),<sup>179</sup> *Helicobacter pylori* (*H. pylori*)<sup>180</sup> and *Campylobacter jejuni* (*C. jejuni*)<sup>181</sup> and more recently it has also been discovered in a gram positive bacteria; *Bacillus thuringiensis* (*B. thuringiensis*).<sup>182</sup>



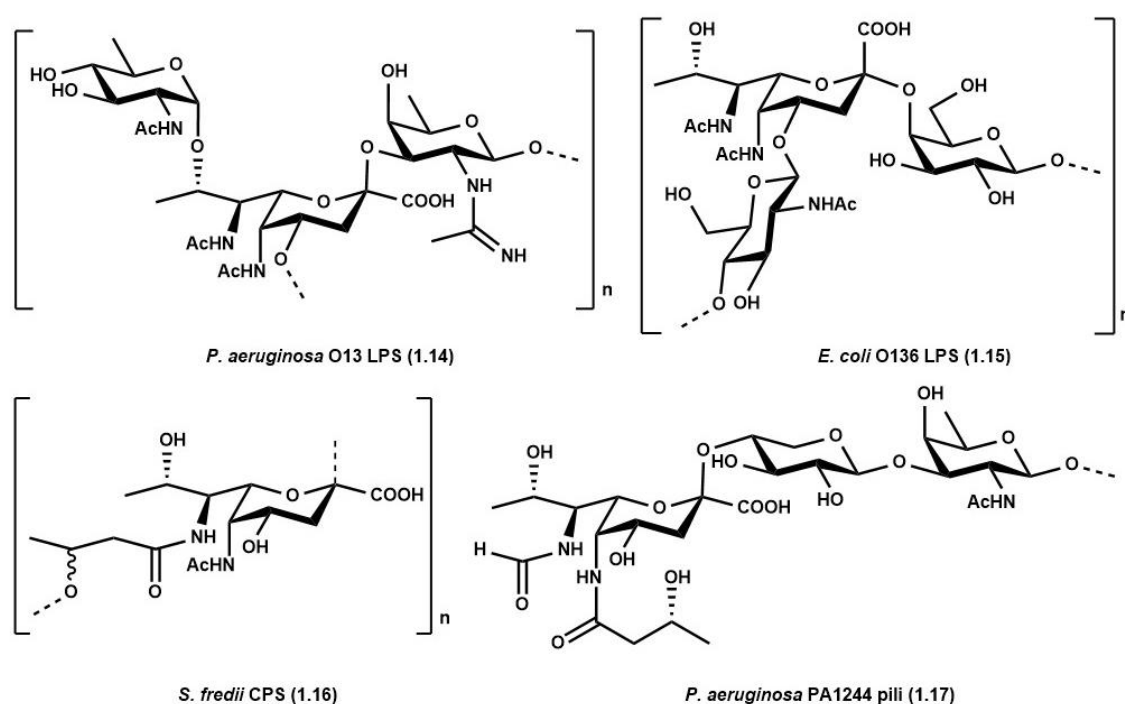
**Figure 1.12** The structurally related nonulosonic acids; Neu5Ac (**1.1**) and Pse5Ac7Ac (**1.13**).

During the initial identification of Pse5Ac7Ac (**1.13**), a derivative was also tentatively assigned with NMR peaks observed consistent with an *N*-(3-hydroxybutyryl) group at the C5 position.<sup>178</sup> Subsequent research on other pseudaminic acid structures has highlighted that derivatisation of the C5 and C7 acetamido groups is common (**Figure 1.13**).<sup>2</sup> For example, *C. jejuni* flagellin protein is extensively glycosylated with pseudaminic acid in the form of Pse5Ac7Ac (**1.13**), Pse5Am7Ac and minor amounts of other derivatives.<sup>183</sup>



**Figure 1.13** Examples of some pseudaminic acid derivatives.<sup>2</sup>

Pse5Ac7Ac (**1.13**) and derivatives are most common on bacterial cell surfaces with reports of LPS and CPS containing pseudaminic acids as well as pseudaminic acid glycosylated flagella and pili. As well as differentiation occurring from the different functionalities present on the C5 and C7 acetamido groups<sup>2</sup> (**Figure 1.13**) structural variations occur *via* the numerous glycosidic linkages available in Pse5Ac7Ac (**1.13**) containing carbohydrates. Unusually for sialic acids, there is evidence for Pse5Ac7Ac (**1.13**) linked in both the  $\alpha$ 2- and  $\beta$ 2- conformation, for example *P. aeruginosa* O13 LPS incorporates a Pse5Ac7Ac( $\alpha$ 2-3)<sub>L</sub>-FucAm bond (**1.14**)<sup>184</sup> whereas *E. coli* O136 LPS has a Pse5Ac7Ac( $\beta$ 2-4)<sub>D</sub>-Gal (**1.15**) (**Figure 1.14**).<sup>185</sup> Additionally Pse5Ac7Ac (**1.13**) is more commonly found linked within the glycan in LPS and CPS structures rather than at the terminal position which is more typical of sialic acids.



**Figure 1.14** Examples of pseudaminic acid cell surface structures, showing a selection of the different glycosidic linkages utilised in nature.

As well as glycosidic bonds forming with hydroxyls in the sugar ring of Pse5Ac7Ac (**1.13**), certain derivatives allow for binding to the functional groups. For example, the *Sinorhizobium fredii* (*S. fredii*) CPS is a Pse5Ac7hydroxybutyryl homo-polysaccharide (**1.16**) with the  $\alpha$  glycosidic linkage between the anomeric position and the hydroxyl group on the C5 *N*-hydroxybutyryl (**Figure 1.14**).<sup>186</sup> Investigations into *P. aeruginosa* PA1244 pili highlighted a pseudaminic acid derivative as the terminal sugar of a trisaccharide (**1.17**) with each pilin monomer found to be glycosylated with one trisaccharide (**Figure 1.14**).<sup>187</sup> This trisaccharide (**1.17**) has also been shown to be the

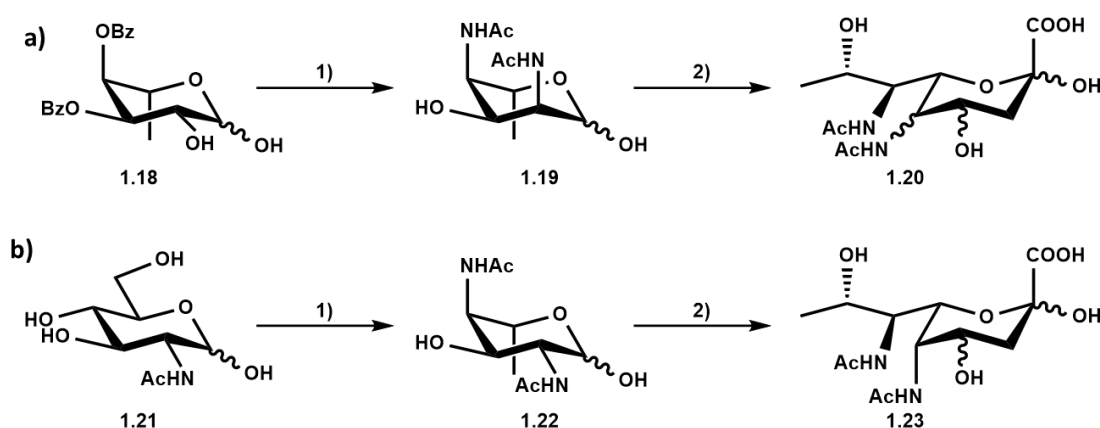
repeating unit in the LPS of *P. aeruginosa* belonging to the O7 serotype.<sup>188</sup> The incorporation of pseudaminic acid derivatives are utilised in many *P. aeruginosa* strains however heterogeneity exists within both the pseudaminic acid derivative utilised and the glycan sequence and structure.

The positioning of pseudaminic acid derivatives on bacterial cell surface structures such as flagella, pili, CPS and LPS, suggests a potential pathogenic role.<sup>189</sup> Therefore research focussed on confirming whether incorporation of Pse5Ac7Ac (**1.13**) (and derivatives) offers a selective advantage has shown that Pse5Ac7Ac (**1.13**) can have a direct role in the functioning of virulence factors.<sup>2</sup> For example, *C. jejuni* flagellin proteins (such as FlaA1) are glycosylated with Pse5Ac7Ac (**1.13**) monosaccharides (or a derivative) at 19 sites.<sup>183</sup> Mutations to *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic genes resulted in incorrectly assembled flagellar and reduced motility.<sup>190</sup> Similarly *H. pylori* FlaA and FlaB flagellin proteins have been found to be exclusively glycosylated with Pse5Ac7Ac (**1.13**) and prevention of flagellin glycosylation resulted in no detectable flagella and non-motile bacterium.<sup>180</sup> As motility has previously been shown to be a key factor in the ability of these two bacteria to establish an infection,<sup>191-192</sup> it can be suggested that therapeutics targeting the Pse5Ac7Ac (**1.13**) flagellin glycosylation process could reduce the virulence of such bacteria.

Furthermore it has been proposed that Pse5Ac7Ac (**1.13**) has a role in bacterial evasion from the host immune system as it has structural similarities with Neu5Ac (**1.1**) (a prevalent host sugar). Therefore glycosylation with this sugar may prevent the pathogenic bacteria from being recognised as a foreign body. For example, many pathogens manipulate host immune system by binding to receptors promoting interleukin-10 induction which can suppress immune responses.<sup>193</sup> *C. jejuni* FlaA1 was found to be important for binding to siglec-10 (a host immune system glycan receptor) and when mutations prevented glycosylation with Pse5Ac7Ac (**1.13**) there was a significant decrease in the induction of interleukin-10. Thus suggesting that Pse5Ac7Ac (**1.13**) can be recognised by siglec-10 and can mediate immune responses.<sup>194</sup>

### 1.4.2 Chemical synthetic routes towards pseudaminic acid

Although Pse5Ac7Ac (**1.13**) is structurally similar to the ubiquitous, commercially available sugar Neu5Ac (**1.1**) (**Figure 1.16**), synthesis of Pse5Ac7Ac (**1.13**) has proven to be much more challenging. The main complications can be attributed to the epimeric stereochemistry at the C5 position and the different desired functionality and stereochemistry on the propyl chain. A number of syntheses have been attempted but an economically viable, facile, high yielding synthesis is yet to be developed which would allow for large scale synthesis of this rare sugar. Chemical syntheses of Pse5Ac7Ac (**1.13**) have tended to adopt one of two approaches; synthesis of the biosynthetic precursor (or stereoisomers thereof) followed by a condensation reaction to attach the functionalised propyl chain, or *via* manipulation of structurally similar nonulosonic acids to gain the desired pseudaminic acid functionality and stereochemistry.



**Scheme 1.4** Attempted synthetic routes to Pse5Ac7Ac (**1.13**) producing a number of stereoisomeric products adapted from **a)** Knirel<sup>195</sup> and **b)** Ito<sup>196</sup>. Reagents and conditions:

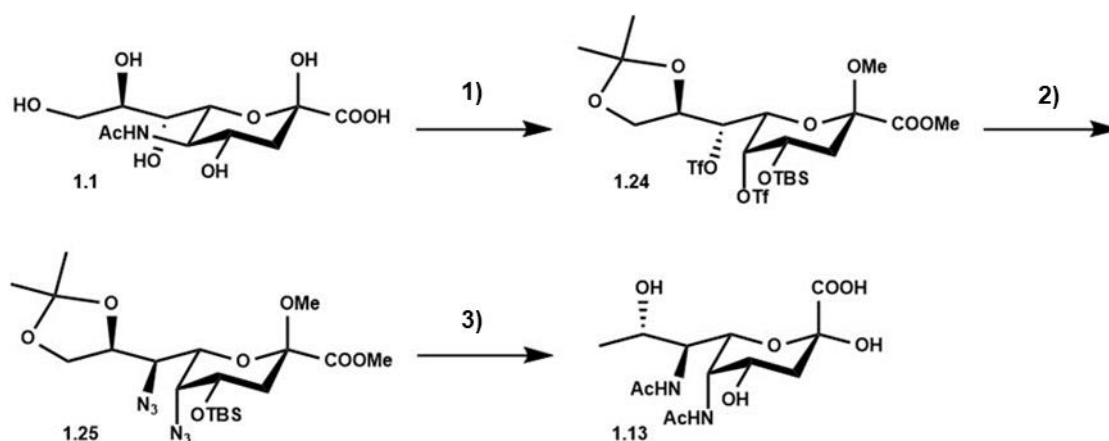
**a) 1)** Seventeen steps; (i) Bu<sub>2</sub>SnO, benzene, reflux, then BnBr, Bu<sub>4</sub>NBr, benzene, reflux. (ii) MeONa, MeOH. (iii) 2,2-dimethoxypropane, PTSA, acetone, rt. (iv) oxalyl chloride, DMSO, *i*-Pr<sub>2</sub>NEt, CH<sub>2</sub>Cl<sub>2</sub>, -60 °C. (v) NaBH<sub>4</sub>, aq EtOH, rt. (vi) 80 % aq AcOH, 40 °C. (vii) Bu<sub>2</sub>SnO, benzene, reflux. (viii) BzCl, benzene, 0 °C-rt. (ix) Tf<sub>2</sub>O, pyridine, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C. (x) Bu<sub>4</sub>NN<sub>3</sub>, toluene, 60 C. (xi) NaN<sub>3</sub>, DMF, dibenzo-18-crown-6, rt. (xii) NaN<sub>3</sub>, DMF, r.t. (xiii) LiAlH<sub>4</sub>, THF, 0 °C-r.t. (xiv) Ac<sub>2</sub>O, MeOH, r.t. (xv) MsCl, pyridine, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C-r.t. (xvi) AcONa, aq 2-methoxyethanol, reflux. (xvii) H<sub>2</sub>, Pd(OH)<sub>2</sub>/C, aq MeOH, 35 °C.

**2)** One step; (i) oxalacetic acid, Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>, pH 10.5, r.t.

**b) 1)** Nine steps; (i) BaO, Ba(OH)<sub>2</sub>, DMF, benzyl bromide then formic acid, rt, 18 hrs. (ii) I<sub>2</sub>, PPh<sub>3</sub>, imidazole, THF, 0 °C, 2 hrs. (iii) TIPSOTf, 2,6-lutidine, CH<sub>2</sub>Cl<sub>2</sub>, 12 hrs. (iv) *t*-BuOK, THF, 70 °C, 10 hrs, then TBAF, THF, 2 hrs. (v) H<sub>2</sub>, RhCl(PPh<sub>3</sub>)<sub>3</sub>, benzene, EtOH, 3 hrs. (vi) Dess-Martin periodinane, NaHCO<sub>3</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 17 hrs. (vii) MeONH<sub>2</sub>.HCl, NaHCO<sub>3</sub>, MeOH, 65 °C, 17 hrs. (viii) Sml<sub>2</sub>, MeOH, THF, 12 hrs, then Ac<sub>2</sub>O, pyridine, 6 hrs. (ix) H<sub>2</sub>, Pd(OH)<sub>2</sub>, EtOH, 4 hrs.

**2)** Three steps; (i) In, 0.1 N HCl-EtOH (1:6) 40 °C, 12 hrs. (ii) O<sub>3</sub>, MeOH, -78 °C, then 30 % H<sub>2</sub>O<sub>2</sub>, H<sub>2</sub>O, HCO<sub>2</sub>H, 90 min. (iii) TEA-H<sub>2</sub>O (1:3), 0 °C, 2 hrs.

Knirel *et al* detailed the first synthesis of Pse5Ac7Ac (**1.13**) in 2001 using the first approach detailed above (**Scheme 1.4a**).<sup>195</sup> Although this was a significant breakthrough, issues were highlighted, such as the lengthy 17-step synthesis of the precursor sugar (**1.19**) for which many steps only gave poor yields and required expensive and toxic reagents. The ultimate disadvantage of this synthesis was that it was impossible to control stereochemistry of the product during conversion from the key intermediate (**1.19**) and oxalacetic acid, thus only forming Pse5Ac7Ac (**1.13**) as the minor product. Ito *et al* followed a similar approach and documented a novel synthesis to the biosynthetic precursor (**1.22**) with the aim that the alternative stereochemistry of this sugar would improve the stereoselectivity of the condensation reaction.<sup>196</sup> Unfortunately two stereoisomers (**1.23**) prevailed and optimisation could still only select for the desired conformation marginally more than the undesired (**Scheme 1.4b**). Additionally this synthesis also suffered from requiring toxic and costly reagents with steps that only produced low yields of desired product.



**Scheme 1.5** Attempted synthetic routes to Pse5Ac7Ac (**1.13**) from Neu5Ac (**1.1**) via a key 5,7 bis-azide intermediate (**1.25**). Reagents and conditions:

**1)** Six steps; (i) NaNO<sub>2</sub>, Ac<sub>2</sub>OH-AcOH (2:1), 0 °C, 1 hr, then 50 °C, 6 hrs. (ii) MeONa, MeOH. (iii) 2,2-dimethoxypropane, PTSA, acetone, r.t. (iv) Imidazole, TBDMS-Cl, DMF, r.t, 16 hrs. (v) Ac<sub>2</sub>O, pyridine, r.t, overnight. (vi) Tf<sub>2</sub>O, pyridine, CHCl<sub>2</sub>, -78 °C, 10 mins, then 0 °C, 1 hr and 0 °C-r.t 5 hrs.

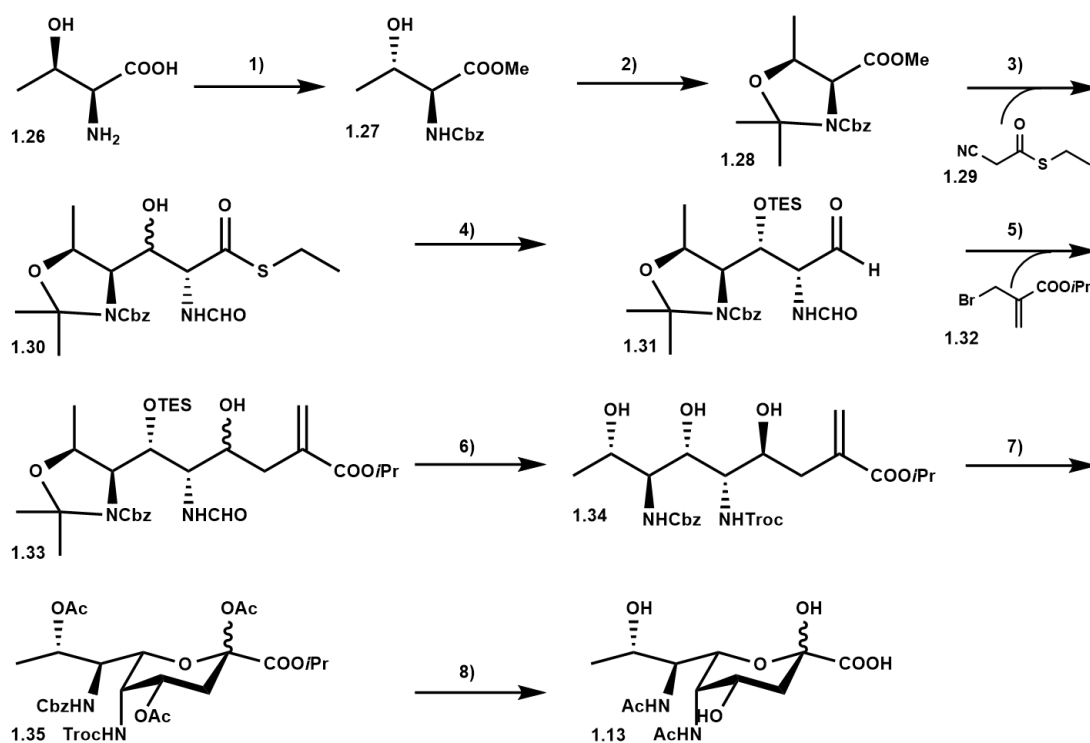
**2)** One step; NaN<sub>3</sub>, DMF, Ar (g), 0 °C, then 4 °C, 24 hrs.

**3)** Thirteen steps; (i) 50 % aq. TFA, CHCl<sub>2</sub> 0-r.t, 2 hrs. (ii) TBDMS-Cl, imidazole, DMF, r.t, 16 hrs. (iii) Dess-Martin periodinane, CHCl<sub>2</sub>, r.t, 2 hrs. (iv) BH<sub>3</sub>.THF, THF, r.t, 16 hrs. (v) Ac<sub>2</sub>O, pyridine, DMAP, r.t, 12 hrs. (vi) *p*-TSOH.H<sub>2</sub>O, Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, MeOH, r.t, 2 hrs. (vii) Ac<sub>2</sub>O, pyridine, r.t, 12 hrs. (viii) TFA, THF-H<sub>2</sub>O (4:1), r.t, 30 mins. (ix) NaOMe, MeOH, r.t, 2 hrs. (x) I<sub>2</sub>, PPh<sub>3</sub>, imidazole, THF, 60 °C, 2 hrs. (xi) *i*Pr<sub>2</sub>EtN, Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, MeOH, r.t, 16 hrs. (xii) aq. NaOH (1 M), 40 °C, 1 hr. (xiii) Dowex-50WX8(H<sup>+</sup>), 80 °C, 36 hrs.



A second approach has been explored by Kiefel<sup>197</sup> who noted that a protected 5,7 bis-azide (**1.25**) with the desired stereochemistry could be produced from the readily available Neu5Ac (**1.1**) to form a key intermediate in the pathway to Pse5Ac7Ac (**1.13**) (**Scheme 1.5**). However even synthesis of this intermediate required vast optimisation with each step requiring multiple attempts in order to produce satisfactory yields.<sup>197</sup> Further issues developed later in the pathway when upon reduction of the azides, undesired products were predominantly formed with little desired Pse5Ac7Ac (**1.13**) under a number of different conditions.<sup>198</sup> Eventually a synthesis to Pse5Ac7Ac (**1.13**) was reported in 17 steps using the optimised synthesis of the bis-azide (**1.25**) (**Scheme 1.5**).<sup>198</sup> All of these chemical syntheses suffer from being unable to control stereoselectivity, use of expensive and/or toxic reagents and ultimately have only been able to produce minor amounts of Pse5Ac7Ac (**1.13**) in very low yields. Additionally, major modifications of these syntheses would be required in order to introduce different functionalities at the desired sites to mimic the natural derivatives or produce chemical probes.

A new chemical route has recently been developed involving two chain elongations to produce pseudaminic acid structures, including those with different acetamido functionalities, which can undergo glycosylation.<sup>199</sup> Protected *L-allo*-threonine (**1.27**) was synthesised from the low-cost *L*-threonine starting material (**1.26**), protected (**1.28**) and reacted with a glycine thioester isonitrile (**1.29**) in an aldol-type reaction<sup>200</sup> to produce two stereoisomers of a di-amino skeleton (**1.30**) (**Scheme 1.6**). To drive the stereoselectivity of the coupling reaction to that of the desired product a large screen of *L-allo*-threonine (**1.20**) protecting groups, solvents and catalysts was completed and achieved a maximal selectivity of 5:1 desired:undesired.<sup>199</sup> After conversion of the thioester (**1.30**) to an aldehyde (**1.31**), the second elongation step was performed utilising an indium-mediated Barbier-type allylation<sup>201</sup> (**1.33**) and resulting isomers purified. Partial deprotection followed by alkene cleavage and acetylation led to the formation of a cyclised sugar (**1.35**) with the same stereochemistry as Pse5Ac7Ac (**1.13**) (**Scheme 1.6**). This product was then used in the desired glycosylation as well as being used to yield Pse5Ac7Ac (**1.13**) *via* removal of the protecting groups.



**Scheme 1.6** A novel synthetic route to pseudaminic acid derivatives producing a protected nonulosonic sugar with the desired pseudaminic acid stereochemistry. Reagents and conditions:

- 1)** Six steps; (i)  $\text{SOCl}_2$ , MeOH, reflux. (ii) AcCl, Et<sub>3</sub>N,  $\text{CHCl}_2$ . (iii)  $\text{SOCl}_2$ . (iv) 10 % aq. HCl, reflux. (v) CbzCl,  $\text{Na}_2\text{CO}_3$ , H<sub>2</sub>O. (vi) MeI,  $\text{KHCO}_3$ , DMF.
- 2)** One step; DMP,  $\text{BF}_3 \cdot \text{OEt}_2$ , DCM, 12 hrs.
- 3)** Four steps; (i)  $\text{NaBH}_4$ , CaCl<sub>2</sub>, EtOH-THF, 24 hrs. (ii) BAIB, TEMPO,  $\text{CH}_2\text{Cl}_2$ , 0 °C-r.t., 10 hrs. (iii) LiOTf, iPr<sub>2</sub>Net, DCE-DMF, r.t., 3 hrs. (iv) THF-H<sub>2</sub>O, reflux, 10 hrs.
- 4)** One step; Et<sub>3</sub>SiH, Pd/C, THF, 3 hrs.
- 5)** One step; Indium powder, NH<sub>4</sub>Cl, EtOH, 2 hrs.
- 6)** Six steps; (i) Dess-Martin periodinane,  $\text{CHCl}_2$ , 0 °C, 2 hrs. (ii) TBAF, HOAc, THF, 1 hr. (iii)  $\text{NaBH}(\text{OAc})_3$ , HOAc, MeCN, -40 to -20 °C, 10 hrs. (iv) HOAc, H<sub>2</sub>O, 50 °C, 20 hrs. (v) 3 % aq. HCl in MeOH, 0 °C to r.t., 8 hrs. (vi) TrocCl, 0.5 M  $\text{Na}_2\text{CO}_3$ , MeCN, 2hrs.
- 7)** Two steps; (i)  $\text{O}_3$ ,  $\text{CHCl}_2$ , -78 °C, 30 mins, then Me<sub>2</sub>S. (ii) Ac<sub>2</sub>O, pyridine, DMAP.
- 8)** Six steps; (i) TolSH,  $\text{BF}_3 \cdot \text{OEt}_2$ ,  $\text{CHCl}_2$ , 16 hrs. (ii) BnOH, TolSCL, AgOTf, AW-300 MS, -78 °C. (iii) Zn (s), Ac<sub>2</sub>O, HOAc, 40 °C, 3 hrs. (iv) Pd/C, H<sub>2</sub>, NH<sub>4</sub>OAc, MeOH- $\text{CH}_2\text{Cl}_2$ , 1 hr. (v) aq. LiOH, MeOH-THF (4:1), r.t., 24 hrs. (vi) Pd/C, H<sub>2</sub>, MeOH-H<sub>2</sub>O, 12 hrs.

This work not only constitutes a significant improvement in the synthesis of Pse5Ac7Ac (**1.13**); higher yields, lower cost of reagents, higher control of stereochemistry, but allows for other natural pseudaminic acid structures to be synthesised which has long since been a desired synthetic characteristic, but previously unachievable. Despite these advantages, this chemical synthesis to Pse5Ac7Ac (**1.13**) is still lengthy, time-consuming and requires high skill for optimal production. Furthermore, it would be deemed unsuitable for large scale synthesis of Pse5Ac7Ac (**1.13**) due to low atom economy and yield. Notably, the highest yield of Pse5Ac7Ac (**1.13**) recorded was 11 % from any chemical method discussed and only produced 15 mg Pse5Ac7Ac (**1.13**).<sup>199</sup>

## 1.5 Aims

Following the discovery of the first example of the *N*-linked glycosylation biosynthetic pathway in 2002,<sup>202</sup> the study of protein glycosylation in bacteria has undergone an exponential research boom which has revealed a plethora of new carbohydrate processing enzymes which are potential antibacterial therapeutic targets. Enzymes active on non-mammalian glycans hold particular promise for biochemical characterisation and inhibition as the workload of designing small molecules with specificity for only a bacterial enzyme with no off target binding to human enzymes is greatly increased. One such sugar is pseudaminic acid which has been identified on a number of different pathogens and is associated with virulence,<sup>2</sup> for example as a constituent of *Pseudomonas aeruginosa* LPS and pili.<sup>178, 187</sup> Specifically, the uncharacterised enzyme, PA2794 from *Pseudomonas aeruginosa* has been shown to be involved in biofilm formation<sup>173</sup> and has been identified as a putative pseudaminidase.<sup>3</sup>

Although the proposed function has yet to be unequivocally defined, transition state mimics, akin to those used as inhibitors of neuraminidases, but based on the pseudaminic acid scaffold, have been proposed as potential inhibitors of PA2794.<sup>3</sup> These may have therapeutically beneficial effects on biofilm formation in *P. aeruginosa*, with specific application in the treatment of cystic fibrosis patients colonised by this archetypal biofilm bacteria.<sup>173</sup>

Primary aims of this thesis were therefore to **1)** definitively assign the function of PA2794 and elucidate the mechanistic and structural details of its proposed activity on pseudaminic acid. It follows that any structural insight gained could be used to guide rational drug design efforts in a longer term goal. However the characterisation of PA2794 and other potential processing enzymes would require access to a number of pseudaminic acid based chemical probes and analogues which have not previously been synthesised. Therefore another primary aim of the project was **2)** to develop a strategy for the production of the pseudaminic acid derivative, Pse5Ac7Ac (**1.13**) on a large scale (>100 mg) in order to supply the quantity of material required for future studies into the synthetic derivatisation into chemical based enzymatic probes.

As detailed in section **1.4.2**, the existing synthetic routes to Pse5Ac7Ac (**1.13**) have a number of shortcomings, and as such were deemed unfitting for development in this project. Instead we envisaged that the Pse5Ac7Ac (**1.13**) enzymatic pathway could be recapitulated *in vitro* to ensure production of the sugar on a large scale in a simple and economically viable process. Although previous research had demonstrated that the Pse5Ac7Ac (**1.13**) biosynthetic enzymes could be utilised in a one-pot system to synthesise Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**),<sup>203</sup> the method has yet to be utilised widely in the literature, and is also unsuitable for large scale production of Pse5Ac7Ac (**1.13**). This is likely due to limitations arising from enzyme insolubility, unwanted by-product formation, the excessive cost of required co-factors and the challenges of final purification of Pse5Ac7Ac (**1.13**).

Therefore, herein we sought to optimise this biosynthetic strategy, specifically focussing on the development of an augmented chemoenzymatic route, where;

- i) the production of biosynthetic enzymes was optimised,
- ii) the ratio of enzyme concentration was optimised through biochemical exploration, and
- iii) the cost, scalability and flexibility of pseudaminic acid production were improved through co-factor replacement.

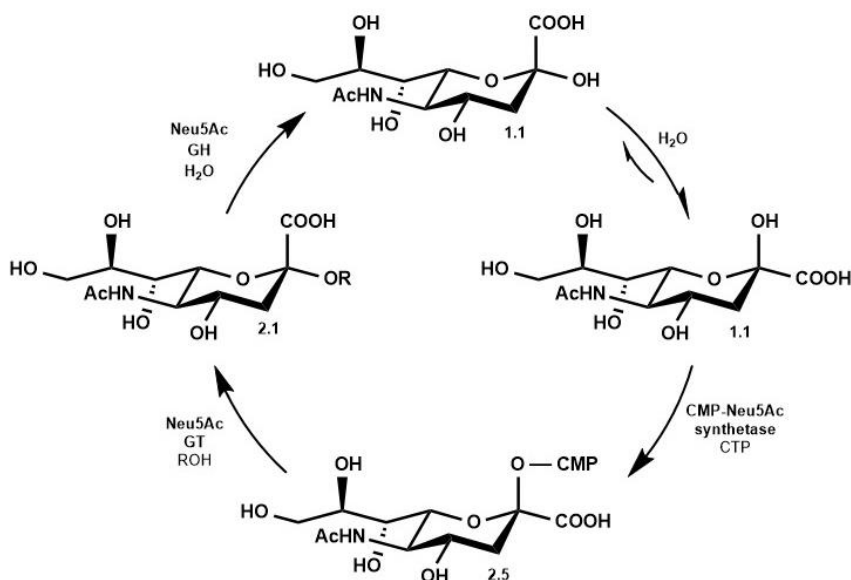
## **Chapter 2 *Pseudomonas aeruginosa* PA2794**

## 2.1 Introduction

### 2.1.1 Neuraminic acid analogues as chemical probes

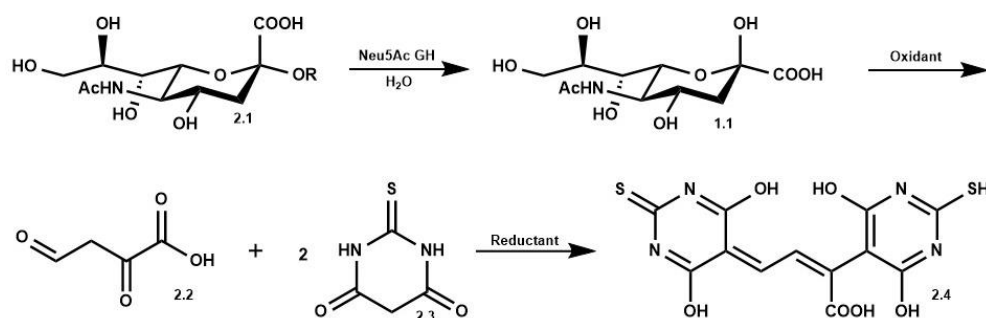
Characterising carbohydrate processing enzymes holds a high degree of interest due to the ubiquitous and complex nature of sugar molecules and their role in mediating disease and infection.<sup>13</sup> In order to characterise and categorise such enzymes it is desired to; observe activity, elucidate mechanisms and investigate essential enzymatic components. Such information is often acquired utilising a combination of structural analysis<sup>204</sup> and *in vitro* biochemical assays,<sup>205</sup> which require isolated, purified enzyme and analogues of the natural substrate/product.<sup>206</sup>

Neuraminidases have been extensively studied due to their importance in controlling the level of Neu5Ac (**1.1**) glycosylation which is essential for normal cell functioning<sup>48</sup> and interactions with invading cells during infection.<sup>74</sup> For example, microorganisms have been shown to cleave host Neu5Ac (**1.1**) from cell surface structures to enable adhesion to the remaining cell surface structure<sup>207</sup> and have their own sialic acid glycosylation which can aid in preventing activation of immune responses.<sup>208</sup> As Neu5Ac (**1.1**) is a ubiquitous terminal saccharide in glycoproteins and glycolipids in the cell membrane,<sup>209</sup> thorough investigations of neuraminidases have been utilised to understand the specifics of each enzyme. This detailed characterisation using a number of different Neu5Ac (**1.1**) analogues has enabled rational drug design and allowed for selective targeting of the desired neuraminidase.<sup>210</sup>



**Scheme 2.1** Neu5Ac (**1.1**) anomeric configurations as the monosaccharide in aqueous solution and during GT catalysed glycosylation and GH catalysed cleavage.

Neu5Ac (**1.1**) is exclusively bound as the  $\alpha$ -anomer in glycans (**5.1**), and all *exo*-neuraminidases have been found to release Neu5Ac (**1.1**) with retention of the  $\alpha$ -configuration.<sup>211</sup> However in solution, mutarotation to the  $\beta$ -anomer quickly occurs as this is the more thermodynamically favourable configuration (**Scheme 2.1**). This  $\beta$ -configuration is retained during the enzyme catalysed formation of the nucleotide activated sugar<sup>212</sup> and hence sialyltransferases are utilised to invert the stereochemistry to the  $\alpha$ -anomer (**2.1**) during the enzyme catalysed glycosidic bond formation (**Scheme 2.1**).<sup>213</sup> As a result all GH and GT Neu5Ac (**1.1**) chemical probes have been designed to mimic the anomeric configuration and any commercially available probes are exclusively  $\alpha$ -linked. A number of methods exist for quantifying neuraminidase activity and either utilise coupled reactions or modified substrates to increase sensitivity. Neuraminidase activity kits are commercially available and quantify release of Neu5Ac (**1.1**) by addition of reagents that react with free Neu5Ac (**1.1**) to convert it into colorimetric products. For example, the thiobarbituric assay relies on oxidation of free Neu5Ac (**1.1**) to form  $\beta$ -formylpyruvic acid (**2.2**) which can then react with added thiobarbituric acid (**2.3**) to produce a chromophore (**2.4**) (**Scheme 2.2**).<sup>214</sup> Therefore enzymatic activity for Neu5Ac (**1.1**) cleavage can be calculated based on detection of the chromophore (Abs<sub>532</sub>) which is more sensitive than detecting Neu5Ac (**1.1**).

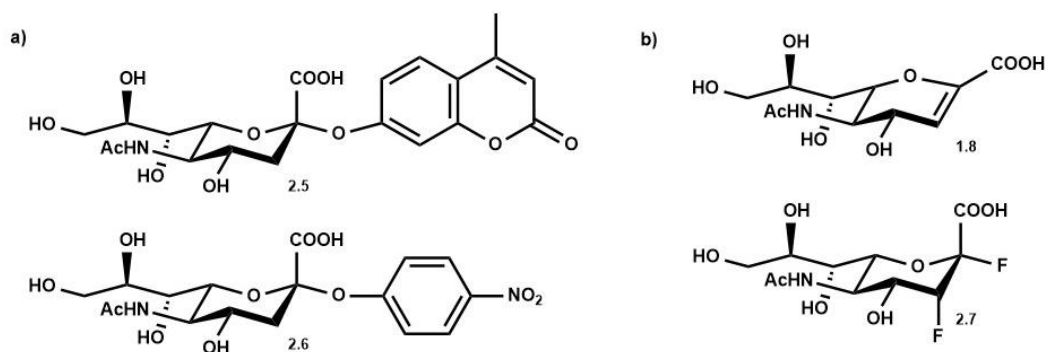


**Scheme 2.2** Thiobarbituric assay; a coupled reaction to observe enzymatic cleavage of Neu5Ac (**1.1**). Neu5Ac (**1.1**) is oxidised so that it can act as a linker for two thiobarbituric acid (**2.3**) molecules producing a chromophore that can be detected as an indicator of linked Neu5Ac (**2.1**) cleavage.

However the most common method for neuraminidase activity quantification utilises Neu5Ac (**1.1**) glycosidically linked to a fluorophore such as 4-methylumbelliferone (**2.5**)<sup>215</sup> or para-nitrophenol (**2.6**) and measures the release of said fluorophore (**Figure 2.1a**). These assays have been shown to be both more sensitive and specific than the coupled assays. Additionally unlike coupled reactions, they don't require the use of multiple steps or addition of other reagents to favour a secondary reaction to convert Neu5Ac (**1.1**) into a more detectable structure. However although it can be normalised, it doesn't make use of the natural substrates and hence the absolute activity of the enzyme of interest cannot be calculated. For example, some neuraminidase active sites may be able to better accommodate the fluorophore, affecting the



observed activity of the enzyme. Therefore complementary assays are often used, alongside structural observations in order to fully understand and compare the observed neuraminidase activity with fluorometric assays.



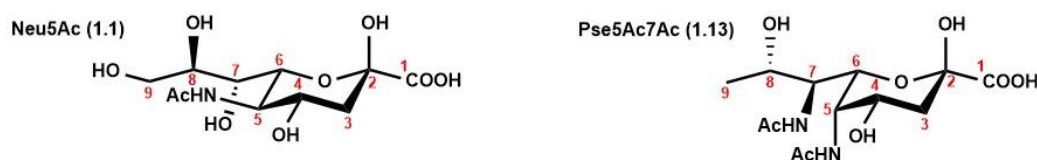
**Figure 2.1** Structures of Neu5Ac (**1.1**) chemical probes; **a)** Neu5Ac fluorophores 4-methylumbelliferyl- $\alpha$ -Neu5Ac (**2.5**) and *p*-nitrophenyl- $\alpha$ -Neu5Ac (**2.6**), and **b)** a neuraminidase inhibitor Neu5Ac2en (**1.8**) and covalent inactivator 2,3-difluoro-Neu5Ac (**2.7**).

Alongside the proposal of Neu5Ac2en (**1.8**) as a neuraminidase inhibitor, mechanistic and structural insight into neuraminidases allowed for predictions of structures that could act as neuraminidase covalent inactivators. In turn these inactivators have been used to further probe the active site for extra details, such as for identification of catalytic residues and for observation of the sugar conformation in the enzyme-bound intermediate (**Scheme 1.3, Page 33**). Knowledge of the retaining mechanism led to the design of molecules that were proposed to increase the rate of formation of the enzyme-bound intermediate but decrease the rate of the deglycosylation step, to prevent release of Neu5Ac (**1.1**). Fluorinated Neu5Ac analogues were identified as ideal candidates as they could drastically change the reactivity of the sugar molecule without incurring potential steric issues. In order to attenuate release of Neu5Ac (**1.1**), structures were designed that would de-stabilise formation of the oxocarbenium ion transition state such as introduction of an electronegative group at the C3 position. As both the glycosylation and deglycosylation steps occur *via* an oxocarbenium ion transition state both steps of the reaction would be slowed. Therefore to compensate for this in the glycosylation step, addition of a good leaving group but poor nucleophile at C2, was suggested, to promote attack of the nucleophilic residue and increase formation of the enzyme-bound intermediate. Incorporation of fluorine at the C2 and C3 positions in Neu5Ac (**2.7**) has been shown to increase the first step in the neuraminidase mechanism and drastically reduce the second step; trapping the sugar in the enzyme bound state (**Scheme 1.3**). Hence enzyme crystallographic studies in complex with 2,3-difluoro-Neu5Ac (**2.7**) have aided in identification of the nucleophilic catalytic residue and prediction of active site residues involved in intermediate interactions (**Figure 2.1b**).

### 2.1.2 PA2794 as a pseudaminidase

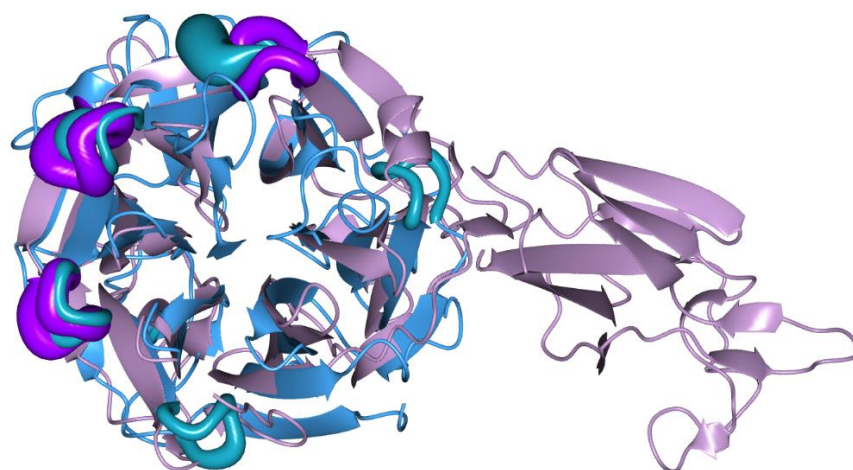
Although a plethora of research exists surrounding the biosynthesis of Pse5Ac7Ac (**1.13**),<sup>2, 216</sup> investigations into pseudaminic acid processing enzymes are very rare. This is mainly attributed to the lack of access to pseudaminic acid analogues required for probing enzymatic reactions. However a motility-associated factor protein, maf1, that is considered a transferase enzyme, has been investigated in *A. caviae* for its ability to transfer pseudaminic acid onto flagellin.<sup>217</sup> Additionally, *A. baumannii* KpsS1 has been identified as putative pseudaminyl transferase based on its sequence homology to a *E. coli* Kdo transferase.<sup>218</sup> A glycosidase from *P. aeruginosa* PA2794, has also been tentatively assigned as a pseudaminidase based on its sequence homology to neuraminidases and information gained from the crystal structure.<sup>3</sup>

Biochemical assays in an attempt to identify the PA2794 natural substrate have thus far been inconclusive. *In vitro* activity assays monitoring the release of Neu5Ac (**1.1**) often suggest very muted activity and different studies are contradictory in their results. For example, early research reported detection of PA2794 cleaved Neu5Ac (**1.1**) in a fluorogenic assay and thiobarbituric assay using a partially purified enzyme,<sup>219</sup> however, subsequent fluorogenic assays using pure protein have failed to reproduce such results.<sup>3</sup> Therefore other structurally related nonulosonic acid structures, such as Pse5Ac7Ac (**1.13**), have been considered for their potential as the natural product (**Figure 2.2**).



**Figure 2.2** Nonulosonic acid structures proposed as the potential PA2794 product.

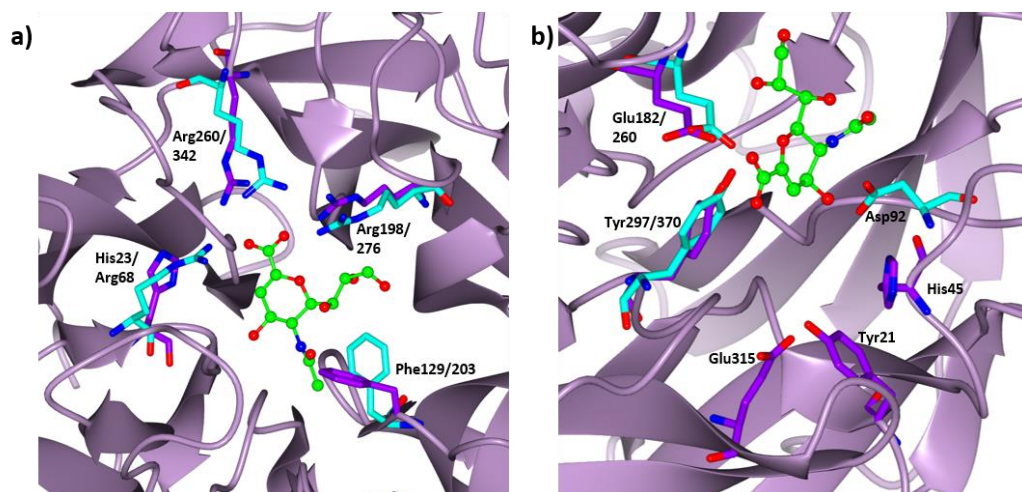
In order to gain further information on the natural substrate of PA2794, X-ray crystallography studies were performed. The X-ray crystal structure of PA2794 highlighted that the enzyme displays the neuraminidase characteristic six-bladed  $\beta$ -propeller domain<sup>174</sup> which, in PA2794, is linked to an immunoglobulin-like domain (**Figure 2.3**).<sup>3</sup> The PA2794 C-terminus domain has a similar structure to the *M. viridifaciens* neuraminidase (R.M.S 2.32 Å for 282 C $^{\alpha}$  atoms) although they do not display high sequence homology.<sup>3</sup> As with neuraminidases the three PA2794 Asp-boxes are situated between the blades and overlap significantly with three of the Asp-boxes from the *M. viridifaciens* neuraminidase (**Figure 2.3**). The structural conservation of the Asp boxes highlights importance of this moiety in forming the overall fold of the catalytic domain.



**Figure 2.3** Overall crystal structure of PA2794 (lilac) (PDB 2W38)<sup>3</sup> overlaid with the *M. viridifaciens* neuraminidase (light blue) (PDB 1EUS)<sup>1</sup> with the characteristic Asp-box sites highlighted (purple-PA2794 and dark cyan-neuraminidase).

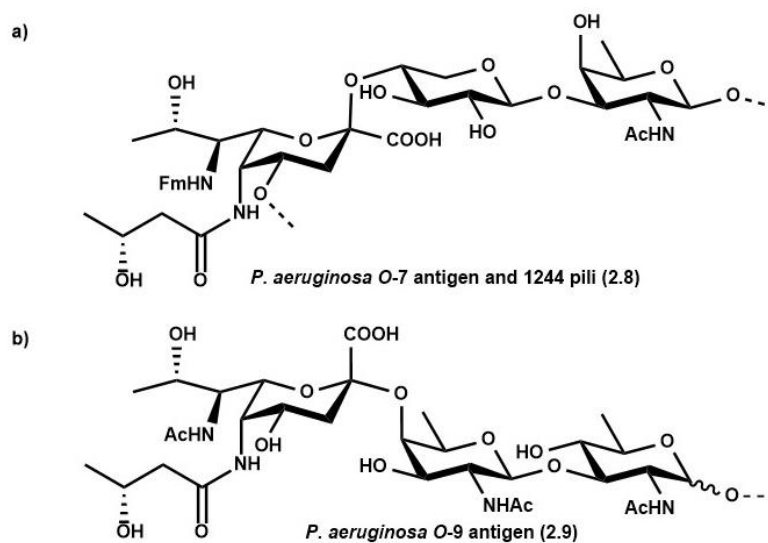
Bacterial *exo*-sialidase active site residues are generally well conserved,<sup>174</sup> with invariant catalytic residues and active site side chain conformations, however the PA2794 active site deviates from some of these features. In particular, PA2794 orientates a phenylalanine ring with an alternate conformation to archetypal sialidases such as a sialidase from *Micromonospora viridifaciens* (*M. viridifaciens*).<sup>1</sup> During *in silico* docking studies, the Phe129 ring conformation was shown to sterically clash with the C5 *N*-acetyl in Neu5Ac2en (**1.8**) but accommodate the stereoisomeric C5 *N*-acetyl of Pse5Ac7Ac (**1.13**) (**Figure 2.4a**).<sup>3</sup> An arginine triad exists in the active site of all characterised neuraminidases that interacts with the carboxylate group. PA2794 displays arginine residues at sites equivalent to the second and third arginine residues in neuraminidases but a histidine residue occupies the site of the first arginine (**Figure 2.4a**).<sup>3</sup>

The tyrosine-glutamine pair that are canonical in neuraminidases were identified as the most likely residues to act as a catalytic nucleophile in PA2794 (Tyr297-Glu182). Comparison of the side chain configurations of these residues with the *M. viridifaciens* neuraminidase shows that they are held in very similar conformations with access to the predicted substrate binding site (**Figure 2.4b**). PA2794 displays a loop in the active site that is homologous to neuraminidases and commonly consists of an aspartic acid that acts as the catalytic acid/base residue. However in PA2794 at the equivalent site in the loop there is no aspartic acid residue. Instead, the predicted acid/base catalytic residue is His45 which can exist as both an acid and base through a charge relay system with a nearby Tyr and Glu (**Figure 2.4b**).<sup>3</sup>



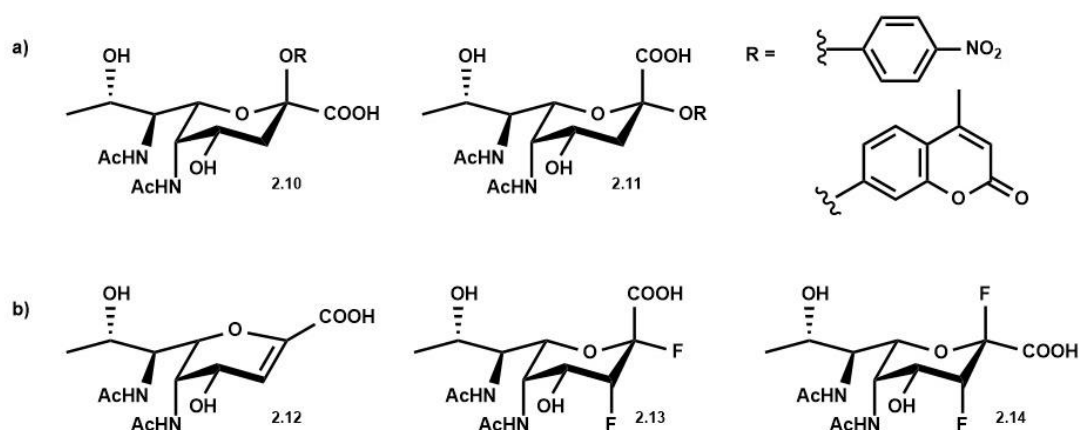
**Figure 2.4** Crystal structure of PA2794 active site residues (purple) (PDB 2W38)<sup>3</sup> overlaid with a sialidase (cyan) in complex with Neu5Ac2en (**1.8**) (green, ball and stick model) (PDB 1EUS).<sup>1</sup> Highlighting differences in **a)** the orientation of conserved neuraminidase active site residues, and **b)** the arrangement of the catalytic residues.

These structural differences imply that the PA2794 enzyme is a nonulosonic acid GH, but not a neuraminidase. In particular the orientation of the Phe129 ring indicates that the enzyme is a pseudaminidase based on the Pse5Ac7Ac (**1.13**) stereochemistry at the C5 position. Additionally this prediction is rationalised as Pse5Ac7Ac (**1.13**) is a natural *P. aeruginosa* nonulosonic acid and expression of a pseudaminidase could aid in control of the level of glycosylation of the LPS and pili.<sup>178, 184, 188</sup> Investigation of the structure of *P. aeruginosa* O-7 and O-9 antigens in 1984 identified a trisaccharide repeating unit with a novel sugar structure; Pse5Ac7Ac (**1.13**) in the terminal position.<sup>178</sup>



**Figure 2.5** *Pseudomonas aeruginosa* surface structures **a)** the O-7 antigen repeating unit and glycosylating trisaccharide of strain 1244 pili (**2.8**) and **b)** the structure of the O-9 antigen repeating unit (**2.9**).

Further research has elucidated the precise structure of these antigens in each serotype showing that Pse5Ac7Ac (**1.13**) is actually most commonly found as  $\alpha$ -linked Pse5RHb7Fo (**2.8**) in *O*-7 antigens (**Figure 2.5a**) and  $\beta$ -linked Pse5Ac7RHb (**2.9**) in *O*-9 antigens (**Figure 2.5b**), with some structures also acetylated at the C4 hydroxyl. The *O*-7  $\alpha$ -linked trisaccharide has also been found to glycosylate *P. aeruginosa* 1244 pilin (**Figure 2.5b**)<sup>187</sup> and stimulate an immune-response in mice models that targets the *P. aeruginosa* *O*-7 antigen.<sup>220</sup> Based on the prediction that a pseudaminic acid glycoside is the natural substrate for PA2794, it is unsurprising that attempts to characterise this enzyme with neuraminic acid glycosides were unsuccessful. As the proposed mechanism<sup>3</sup> is consistent with that of the retaining *exo*-sialidases,<sup>40</sup> it was suggested that effective pseudaminidase probes (**2.10-2.14**) could be designed based on the Neu5Ac analogues (**1.8, 2.5-2.7**) (**Figure 2.6**). For example, it was hypothesised that pseudaminidase activity could be measured *via* monitoring release of a fluorophore from an  $\alpha$ - or  $\beta$ - linked Pse5Ac7Ac glycoside (**2.10, 2.11**) (**Figure 2.6a**). Additionally it was noted that Pse5Ac7ac2en (**2.12**) should behave as an inhibitor and that di-fluorinatedPse5Ac7Ac (**2.13, 2.14**) (**Figure 2.6b**) have potential as covalent inactivators for the reasons discussed above with Neu5Ac analogues.



**Figure 2.6** Proposed structures of pseudaminic acid chemical probes; **a)**  $\alpha$ - and  $\beta$ - linked fluorophores (**2.10, 2.11**) and **b)** the structure of the predicted pseudaminidase intermediate (**2.12**) and covalent inactivators (**2.13, 2.14**).

Unfortunately, as previously discussed, even the synthesis of the most simple pseudaminic acid; Pse5Ac7Ac (**1.13**) is challenging and currently there are no efficient strategies to synthesise Pse5Ac7Ac (**1.13**) on a large scale (> 100 mg) which would allow for further modifications to form the desired chemical probes. Although full characterisation is desired it was proposed that preliminary investigations of PA2794 could still aid in rational drug design by elucidating the natural substrate. Therefore initial investigations focussed on attempting to attain structural evidence for the natural substrate utilising crystal structures in complex with Pse5Ac7Ac (**1.13**).

## 2.2 Expression and purification of PA2794 and PA2794 F129A

### 2.2.1 Transformation and expression of PA2794 and a PA2794 F129A mutant

A recombinant plasmid based on pET15b was designed containing a T7 promoter site upstream of a number of sequences coding for restriction enzymes and a hexa-his tag coding sequence upstream of a thrombin cleavage site. To afford PA2794 with a *N*-terminal hexa-His tag, the plasmid was designed to incorporate the *PA2794* gene sequence into the plasmid between *Nde*I and *Bam*HI restriction enzyme sites using codons optimised for *E. coli* (**Appendix 1**). The resulting recombinant plasmid (purchased from GenScript) was transformed *via* heat shock into chemically competent *E. coli* BL21 DE3 cells and deemed successful by a display of ampicillin resistance during growth on LB agar.

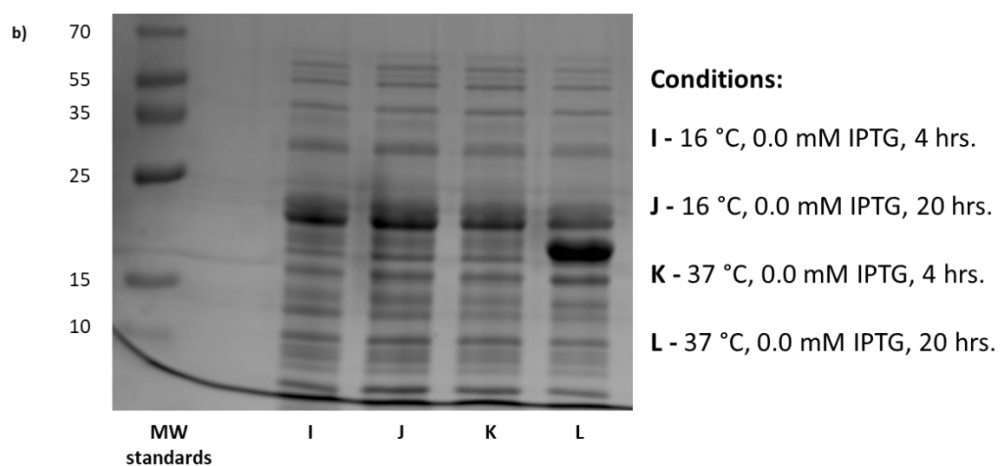
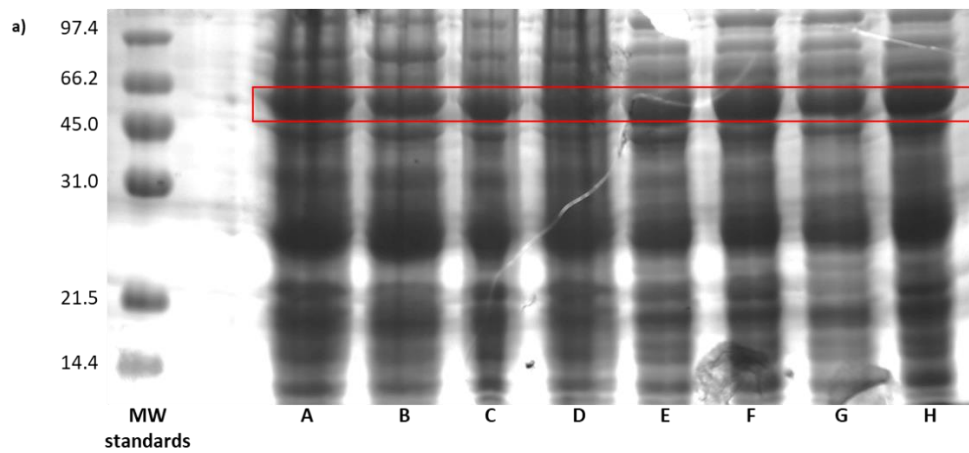
Incorporation of the PA2794 recombinant plasmid did not appear to be detrimental to bacterial growth which was monitored prior to induction of the protein of interest. Although a protocol for expression and purification of PA2794 has previously been determined, incubation temperature and time after induction were investigated in order to optimise expression levels (**Table 2.1**). Prior to induction, a culture sample was taken from each condition and treated exactly the same but without the addition of IPTG to provide an indication of the expression levels of *E. coli* proteins. Expression levels were determined using SDS PAGE analysis of crude lysates from small scale grows up under these induction conditions (**Figure 2.7**).

Samples that had been induced with IPTG displayed an over-expressed protein between the marker 45.0 kDa and 66.2 kDa bands (**Figure 2.7a**) which was much less prominent in samples that had not been induced (**Figure 2.7b**). Therefore this band was tentatively assigned as corresponding to the desired protein PA2794 (47.1 kDa). There were a number of other significant proteins detected by SDS-PAGE, such as one at ~ 25 kDa, which were thought to potentially be truncated PA2794 protein and could require further investigation of expression conditions. However as these were also present in the non-induced samples they were assigned as highly expressed *E. coli* proteins and not over-expressed truncated PA2794 (**Figure 2.7**).

Upon varying the concentration of IPTG there was no significant difference observed in the level of expression between induction with 0.1 mM and 0.5 mM IPTG. However, it was apparent that expression levels could be slightly increased by inducing at a temperature of 37 °C compared to 16 °C and by leaving the cells to grow for longer periods of time (**Figure 2.7a**). Therefore, in subsequent grow ups, expression conditions involved incubation of cells at 37 °C for 18 hrs following induction with IPTG.

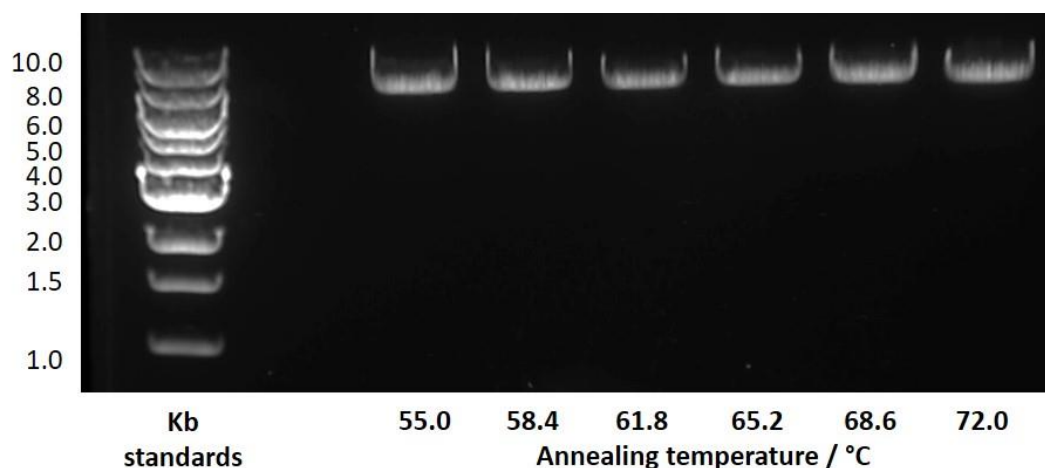
**Table 2.1** Expression trial conditions to optimise production of PA2794.

Condition	A	B	C	D	E	F	G	H
Post induction incubation time	4 hrs				18 hrs			
IPTG concentration	0.1 mM		0.5 mM		0.1 mM		0.5 mM	
Post induction incubation temp	16 °C	37 °C	16 °C	37 °C	16 °C	37 °C	16 °C	37 °C



**Figure 2.7** SDS PAGE analysis of crude lysate samples of the PA2794 enzyme **a)** under different induction conditions after induction with 0.1 mM (**A, B, E, F**) or 0.5 mM (**C, D, G, H**) IPTG and **b)** under these conditions but with no induction with IPTG.

Mutation of the PA2794 phenylalanine 129 to an alanine residue was carried out on this recombinant plasmid to afford an enzyme that has previously been suggested to increase PA2794 activity with Neu5Ac (1.1) analogues. The forward primer was designed to replace the phenylalanine129 codon (TCC) with the most similar codon for an alanine (GCC). PCR utilising a Phusion site-directed mutagenesis kit was carried out with a number of different annealing temperatures (55 °C-72 °C) and the DNA amplification analysed on an acrylamide gel. The presence of a single band on the gel demonstrated successful amplification of a specific length of DNA. However the DNA band appeared slightly higher on the gel than expected under all conditions, suggesting the amplified DNA may have more base pairs than the desired recombinant plasmid (7.2 Kilobases) (**Figure 2.8**). As it was very unlikely that a larger piece of DNA had been amplified with no amplification of the desired plasmid, it was instead suggested that this was the correct stretch of DNA and the apparent increase in number of base pairs was attributed to analysis error. For example, it is hard to distinguish between the standards with large numbers of kilobases as the resolution between them is low. Additionally it was noted that the PCR samples were in a different buffer to that of the standards and hence this may affect their diffusion through the gel. In order to fully assess the success of the PCR reaction the DNA was extracted from the gel and transformed *via* heat shock into chemically competent *E.coli* BL21 DE3 cells.



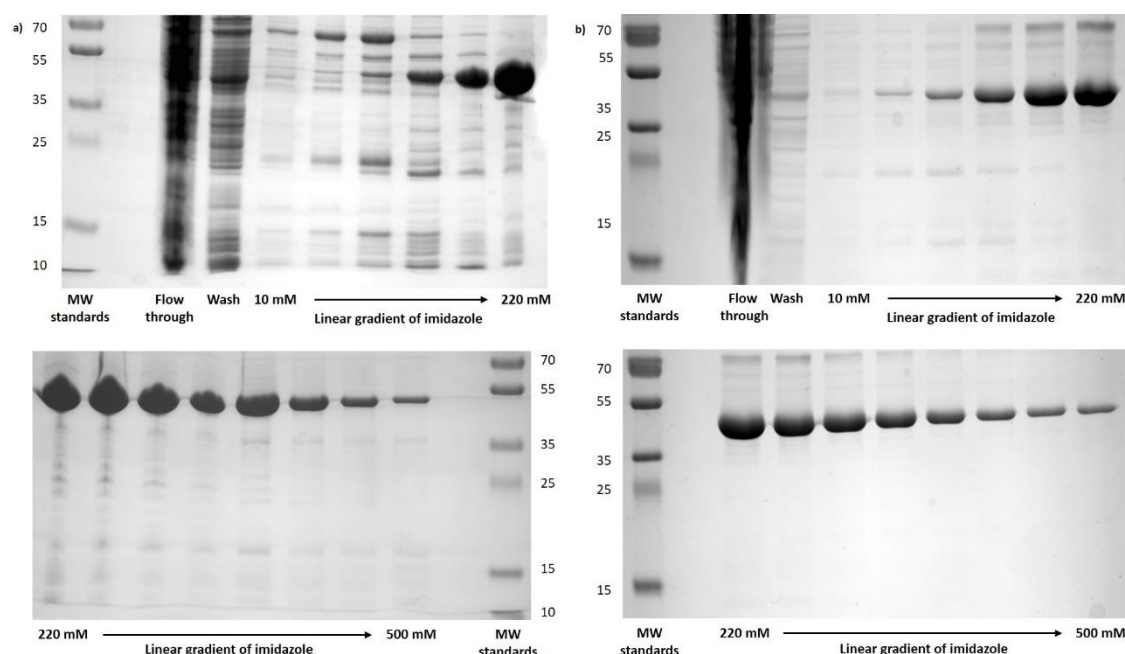
**Figure 2.8** DNA agarose gel following a point mutation PCR of PA2794 pEt15b recombinant plasmid.

The growth of these colonies on ampicillin containing LB agar suggested successful amplification and transformation of the recombinant plasmid and sequencing of a number of colonies (**Appendix 2**) showed incorporation of the F129A mutation in all cases. Prior to DNA extraction the individual colonies were cultured overnight (10 mL LB containing ampicillin 100  $\mu\text{g mL}^{-1}$ ) and a small sample inoculated on LB agar containing ampicillin 100  $\mu\text{g mL}^{-1}$  to ensure that colonies were available containing the DNA that had been sequenced.



### 2.2.2 Purification of PA2794 and PA2794 F129A

Following induction of the desired enzyme (PA2794 and PA2794 F129A) in *E. coli* BL21 DE3 cells on a large scale (3 L LB), cell pellets were resuspended and lysed in Tris-HCl buffer (50 mM, pH 7.5), 10 % glycerol containing 10 mM imidazole, Benzonase (25 U/L grow up) and protease inhibitor tablet (used as instructed). Initial purification from other buffer soluble proteins was attempted by loading onto a Nickel affinity column and applying an increasing imidazole concentration gradient in Tris-HCl buffer (50 mM, pH 7.5) and 10 % glycerol. Presence of a UV Abs<sub>280</sub> peak during washing with higher concentrations of imidazole (after loading and washing the column with low concentrations of imidazole) confirmed over-expression of the His<sub>6</sub>-tagged PA2794 protein (**Appendix 3**). The eluted protein corresponding to the predominant UV Abs<sub>280</sub> peaks was confirmed as the desired PA2794 or PA2794 F129A enzyme by SDS-PAGE analysis which displayed over-expressed protein with the expected molecular weight (**Figure 2.9a,b**).

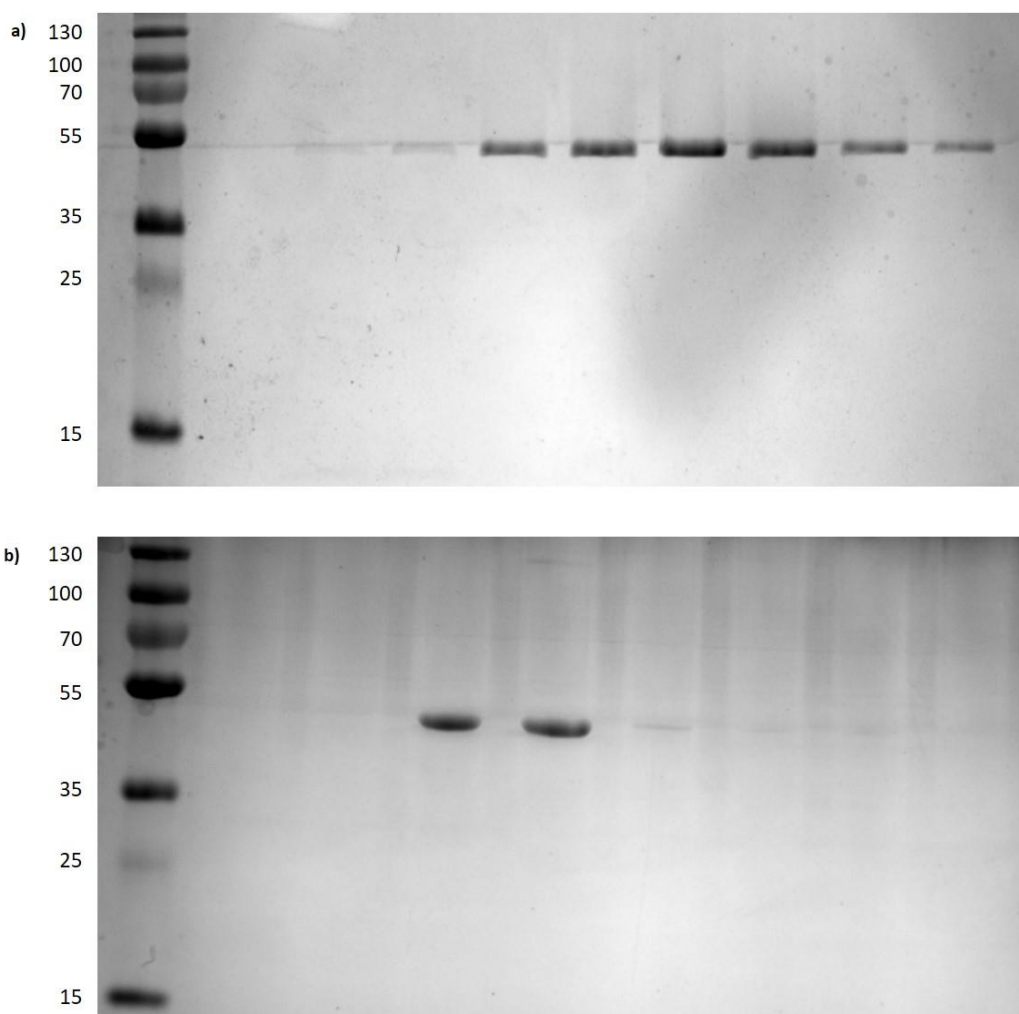


**Figure 2.9** SDS PAGE analysis of **a)** PA2794 and **b)** PA2794 F129A after purification using a Nickel affinity column.

Although this method was successful in purifying PA2794 and PA2794 F129A from the majority of the other proteins, impurities are still evident in both samples, especially for enzyme collected at the lower concentrations of imidazole. It was proposed that in subsequent purifications this could be improved upon by increasing the gradient of imidazole less steeply or by introducing step-wise increments in imidazole concentration instead of application of a linear gradient. However size exclusion was utilised to improve the purity of the PA2794 and PA2794 F129A samples. Elution of the column in Tris-HCl buffer (50 mM, pH 7.5) and 10 % glycerol, resulted in

the desired highly pure PA2794, or PA2794 F129A, as observed by one symmetrical peak in the gel filtration UV trace (**Appendix 4**) and only one protein band in the SDS PAGE analysis of the fractions whereby an increase in Abs<sub>280</sub> was detected (**Figure 2.10a,b**).

It was apparent from the SDS PAGE gels that considerably less PA2794 F129A enzyme was expressed compared to the native PA2794 enzyme. However satisfactory yields of 9 mg per L of LB grow up for PA2794 and 4 mg per L of LB culture for PA2794 F129A were extracted. Following purification, enzymes were concentrated to 3.5 mg mL<sup>-1</sup> in Tris-HCl buffer (50 mM, pH 7.5), 10 % glycerol and flash frozen prior to storage at -80 °C.

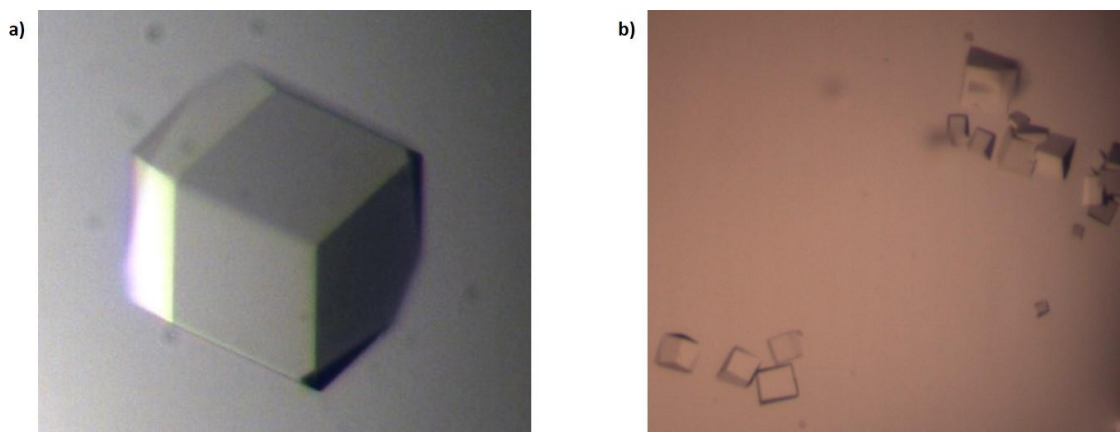


**Figure 2.10** SDS PAGE analysis of **a)** PA2794 and **b)** PA2794 F129A following purification with a size exclusion column (120 mL, HiLoad 16/600) after a Nickel affinity column.

## 2.3 Characterisation of PA2794 and PA2794 F129A

### 2.3.1 PA2794 *apo* crystal structures

Crystallisation conditions for PA2794 had been previously published and were utilised as the initial conditions for investigations. Sitting drop trays containing well solutions of Bicine (0.1 M, pH 4.5, 5.0 or 5.5) and PEG 6K (14 % - 28 %) and drops consisting of 0.3  $\mu\text{L}$  or 0.5  $\mu\text{L}$  PA2794 (3.5  $\text{mg mL}^{-1}$ ) and 0.5  $\mu\text{L}$  well solution were constructed. Crystals were observed developing across a number of conditions and all formed were cubic and well defined after 10 days. A number of drops only contained a single crystal with the maximum recorded size of  $\sim 0.1 \text{ mm} \times 0.1 \text{ mm} \times 0.1 \text{ mm}$  (**Figure 2.11a**). However numerous small crystals developed in the drops containing well solutions with higher PEG 6 K concentrations (24 % - 28 %) which were deemed too small for using in crystallography analysis (**Figure 2.11b**). Generally to produce crystals suitable for X-ray analysis, optimal conditions were; a well solution of Bicine (0.1 M, pH 4.5), PEG 6K 14% -22% or Bicine (0.1 M, pH 5.0), PEG 6K 16% - 22%, and a drop consisting of an equal ratio of mother liquor and PA2794 enzyme (3.5  $\text{mg mL}^{-1}$ ).

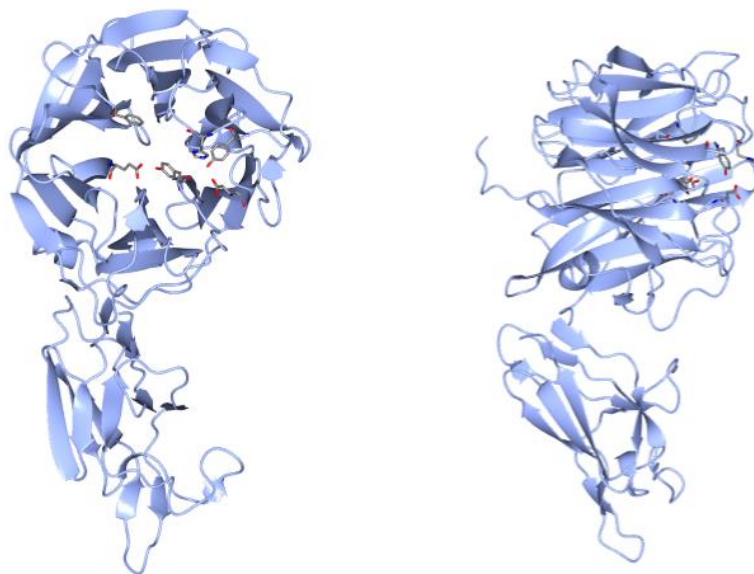


**Figure 2.11** Image of PA2794 crystals developed in well solutions composed of **a)** Bicine (0.1 M, pH 5.0), 18 % PEG 6K and **b)** Bicine (0.1 M, pH 5.0), 24 % PEG 6K

An individual crystal was removed from the drop, placed in a solution containing the well solutes and cryoprotectant (20 % glycerol) and flash frozen. In order to determine if the crystals contained regularly orientated proteins and would provide high resolution electron density, *in-house* diffraction tests were performed. Crystals displaying a regular diffraction pattern to a resolution of 2.3-2.5  $\text{\AA}$  were selected as suitable for further analysis and stored in liquid nitrogen prior to data collection.

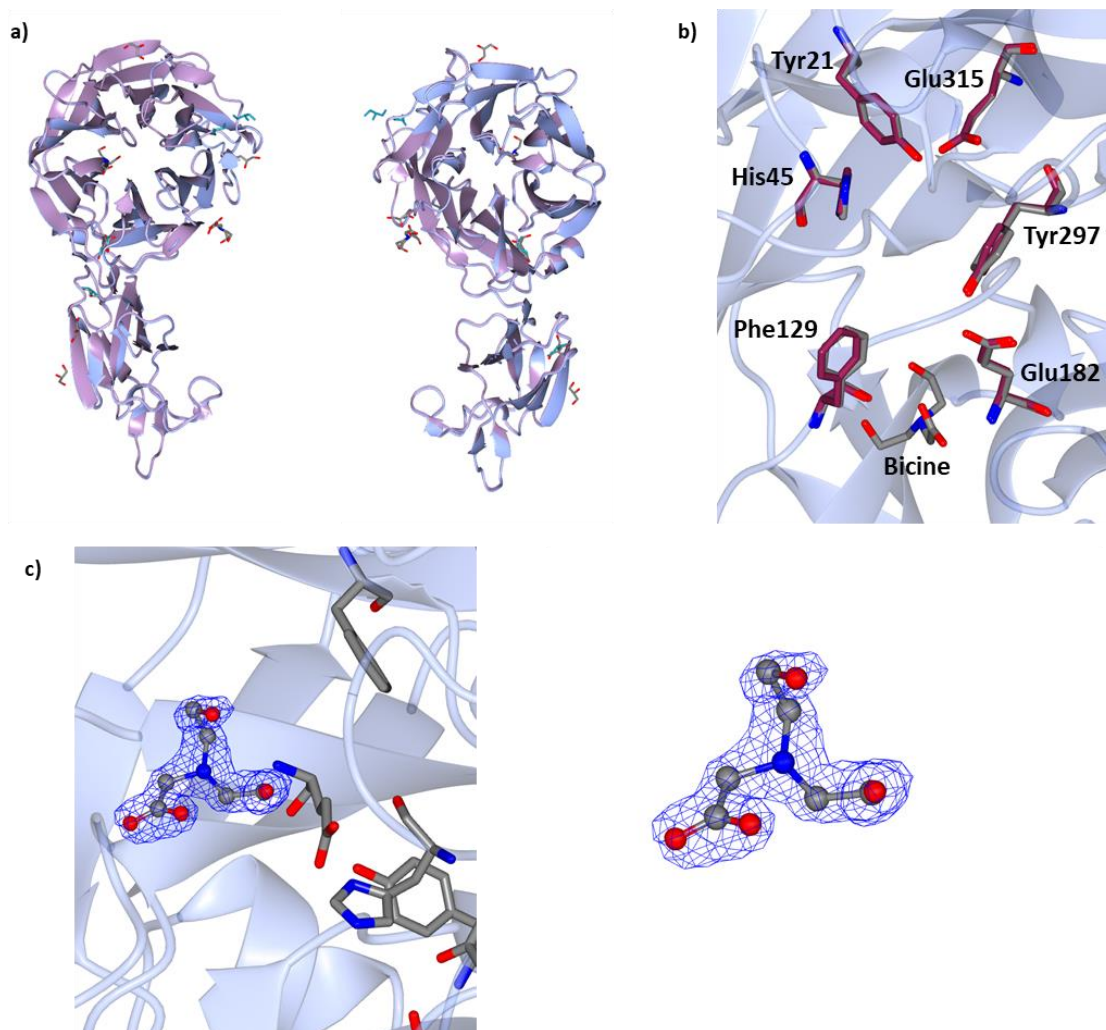
After the collection of X-ray crystallography data (Diamond synchrotron, Oxford), molecular replacement using the previously published crystal data was employed to solve the electron density and provide a base structure for refinement (CCP4I2). Refinement of the PA2794 *apo*

structure was carried out (RefMac5, Coot) and the electron density was assigned to all of the 438 residues with a resolution of 1.2 Å and the majority was in strong agreement with the previously published structures. However this structure was determined at a higher resolution than the published data (1.6 Å) and allowed for the extension of the *N*-terminal to assign the full enzymatic sequence (**Figure 2.12**). The PA2794 crystal had one monomer per asymmetric unit and assigned to the *P2*<sub>1</sub>3 space group with the dimensions  $a = b = c = 126.6$  Å,  $\alpha = \beta = \gamma = 90^\circ$  (with data collection statistics reported in **Appendix 5**).



**Figure 2.12** The fully assigned PA2794 crystal structure (ice blue) resolved to 1.2 Å, with the active site residues highlighted (grey).

Overlaying this refined structure with that of the previously published structure shows a strong agreement of the overall structure (R.M.S 0.19 Å for 432 C $^{\alpha}$  atoms) (**Figure 2.13a**) including the conformations of residues in the active site (**Figure 2.13b**). Importantly the proposed catalytic residues are displayed in the same orientation and hence support the mechanism previously proposed by Xu *et al*<sup>3</sup> (**Figure 2.13b**). This mechanism suggests a His-Tyr-Glu charge relay system could be utilised as the acid/base catalyst, if there was a rotation of the loop to allow for His45 to be orientated closer to the sugar. Unfortunately no further evidence was gained for this putative acid/base catalytic residue, with the proposed His45, Tyr21, Glu315 charge relay system occupying the same orientation as the published structure.<sup>3</sup> Additionally the conformation of the Phe129 ring overlaps with the previously published (**Figure 2.13b**), with no electron density corresponding to other conformations supporting the notion that this enzyme would not be able to accommodate the transition states of nonulosonic acids with an equatorial C5 group when in the <sup>2</sup>C<sub>5</sub> conformation.



**Figure 2.13** Overlay of the published PA2794 structure (PDB 2w38, lilac, dark purple) and the newly refined structure (ice blue, grey), demonstrating the similarity of **a)** their overall structure, **b)** active site side chain conformations **c)** electron density observed in the active site, fitted to a bicine molecule.

Careful allocation of the electron density to the *apo* structure displays a bicine molecule at the edge of the active site which was not assigned in the published structure (**Figure 2.13c**). It was proposed that even though this molecule did not fully occupy the active site it may still compete with potential ligands and perturb crystallisation of PA2794 in complex with Pse5Ac7Ac (**1.13**). Additionally both structures have a number of glycerol molecules associated with the enzyme however, these were identified in different areas in the different structures indicative that many residue sequences can interact with these molecules.

### 2.3.2 PA2794 ligand bound crystal structures

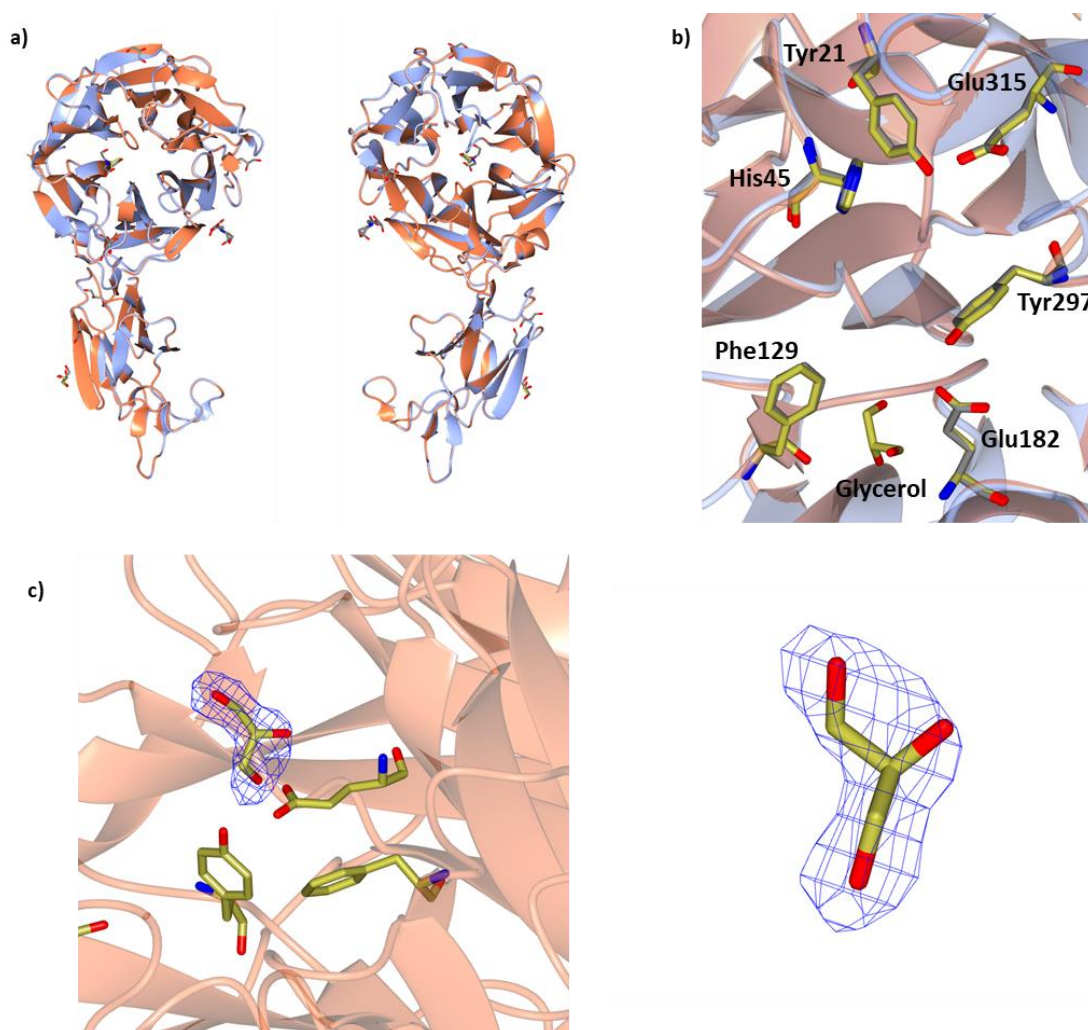
Based on the existing structural data, a glycosidically linked pseudaminic acid derivative is still the main candidate for the PA2794 natural substrate. However structural evidence for this hypothesis is lacking as a crystal structure has not been solved complexed with a ligand in the active site. In order to rectify this, crystallisation conditions were screened in an attempt to collect diffraction data of PA2794 in complex with Pse5Ac7Ac (**1.13**). Due to the potential competition from bicine molecules that were observed in the active site of an *apo* structure, alternative crystallography conditions were utilised for the following data collection.

Sitting drop trays containing well solutions of imidazole (0.1 mM, pH 7.5, 8.0 or 8.5) and PEG 8K (4 % - 18 %) were constructed. Drops consisting of 0.3  $\mu\text{L}$  or 0.5  $\mu\text{L}$  PA2794 (3.5 mg mL<sup>-1</sup>) and 0.5  $\mu\text{L}$  well solution were added alongside the corresponding well solution. Cubic crystals developed under some of these new conditions however these crystals were smaller than the previously used conditions, even after storage for two weeks.

Solid Pse5Ac7Ac (**1.13**) was introduced into drops containing crystals and monitored by eye for diffusion throughout the drop. Addition of Pse5Ac7Ac (**1.13**) caused the immediate disintegration of the crystals, suggesting that it was altering the composition of the drop and having a detrimental effect on the integrity of the crystalline protein. It was proposed that addition of 1 mM Pse5Ac7Ac (**1.13**) in mother liquor may invoke less drastic changes in the drop composition and thus this alternate method for Pse5Ac7Ac (**1.13**) addition was attempted. Addition of Pse5Ac7Ac (**1.13**) in this way did not appear to affect the crystals and crystals were fished after an hour, held in a drop of the mother liquor with 20 % glycerol and flash frozen. Electron density data for this crystal was collected as before; with the ligand checked *in house* then sent to the Diamond Synchrotron for data collection. Molecular replacement of the resulting data was carried out using the *apo* structure from this research (CCP4i2), refinement was carried out (RefMac5, Coot) and the graphics developed (CCP4MG).

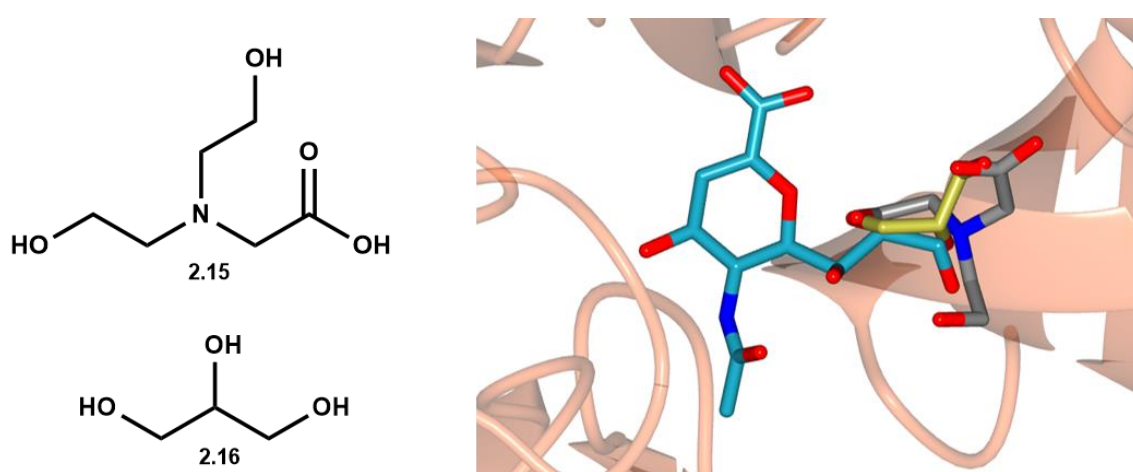
Co-crystallisation conditions involved addition of Pse5Ac7Ac (**1.13**) to the purified PA2794 enzyme to a concentration of 1 mM. Sitting drop trays containing well solutions of imidazole (0.1 mM, pH 7.5, 8.0 or 8.5) and PEG 8K (4 % - 18 %) were constructed. Drops consisting of 0.3  $\mu\text{L}$  or 0.5  $\mu\text{L}$  PA2794 (3.5 mg mL<sup>-1</sup>), Pse5Ac7Ac (1 mM) and 0.5  $\mu\text{L}$  well solution were added alongside the corresponding well solution. No crystalline substances were observed under any of the conditions, suggesting that either addition of Pse5Ac7Ac (**1.13**) had altered the drop conditions, preventing crystallisation, or that binding of Pse5Ac7Ac (**1.13**) had invoked a conformational change that would require alternate conditions for crystallisation.

Following refinement, the crystal structure developed in the alternate crystallisation conditions was superimposed with the *apo* structure from crystals developed under the original conditions (**Figure 2.15a**). There was some deviation in the position of electron density assigned to solute molecules but there was precise agreement (R.M.S 0.20 Å for 432 C $\alpha$  atoms) between the conformation of the PA2794 residues (**Figure 2.15a,b**). Electron density was observed in the PA2794 active site that could not be attributed to the enzymatic structure, however it was much smaller than expected for the desired Pse5Ac7Ac (**1.13**) ligand (**Figure 2.15c**). Fitting of molecules to the electron density elucidated that it could be assigned as a glycerol molecule in a favourable rotamer (**Figure 2.15b**).



**Figure 2.15** Overlay of the PA2794 *apo* structure (ice blue, grey) and the structure soaked with a potential ligand; Pse5Ac7Ac (**1.13**) (coral, gold), demonstrating the similarity of **a**) their overall structure, **b**) active site side chain conformations and the position of binding of an active site solute molecule, **c**) electron density observed in the active site, fitted to a glycerol molecule.

The presence of the bicine (**2.15**) and glycerol (**2.16**) molecules in the active site are somewhat expected when considering that they mimic portions of the Pse5Ac7Ac (**1.13**) structure (**Figure 2.16**) and are present in high concentrations. Both the glycerol and bicine molecules are bound in the active site proximal to where the C6 sugar propane chain is predicted to bind (**Figure 2.16**). Hence it was postulated that attempts to soak Pse5Ac7Ac (**1.13**) to PA2794 crystals may have been unsuccessful due to the glycerol molecule blocking access to the binding site. Even though *in silico* studies have previously demonstrated that Pse5Ac7Ac (**1.13**) can be docked into the PA2794 active site in an energetically favourable conformation,<sup>3</sup> it was also hypothesised that structural modifications may be required for optimal Pse5Ac7Ac (**1.13**) binding. Therefore providing an explanation for why crystals failed to develop in co-crystallisation trials and why crystallised PA2794 does not appear to have a high affinity for Pse5Ac7Ac (**1.13**).



**Figure 2.16** Comparison of the structures of bicine (grey, **2.15**) and glycerol (gold, **2.16**) and the positions they were observed compared to the predicted sugar binding site.

To promote crystallisation of PA2794 in complex with Pse5Ac7Ac (**1.13**), it was attempted to eliminate the use of bicine or glycerol (or other structurally related molecules) to prevent competition for binding with the active site. Although crystallisation conditions without the use of bicine have been established previously,<sup>3</sup> conditions have not been identified without inclusion of glycerol.

Glycerol was removed from the PA2794 storage buffer (Tris-HCl (50 mM, pH 7.4), 10 % glycerol) using 30 kDa molecular weight cut off falcon tubes. Three repeats of addition of 9 mL Tris-HCl (50 mM, pH 7.4) followed by concentration of the enzyme to 1 mL was utilised to buffer exchange the glycerol out of solution. However during this process precipitation of the PA2794 enzyme occurred suggestive that the glycerol is required in aqueous solution to prevent exponential protein aggregation at this concentration. Additionally it was found that the precipitated protein did not re-solubilise upon addition of glycerol up to 20 %.



## 2.4 Conclusions and Future work

A F129A mutation was successfully introduced to the *PA2794* gene and recombinant plasmids transformed into *E. coli* cells for both the native and mutant genes. Expression conditions were optimised for both enzymes and procedures established to obtain enzymes with high purity. Crystallisation of PA2794 was achieved and the whole structure refined to a resolution of 1.2 Å, including the *N*-terminal sequence that had previously not been assigned. The introduction of the putative Pse5Ac7Ac (**1.13**) ligand was attempted pre- and post- crystallisation of PA2794. However electron density within the active site could only be assigned to solute molecules.

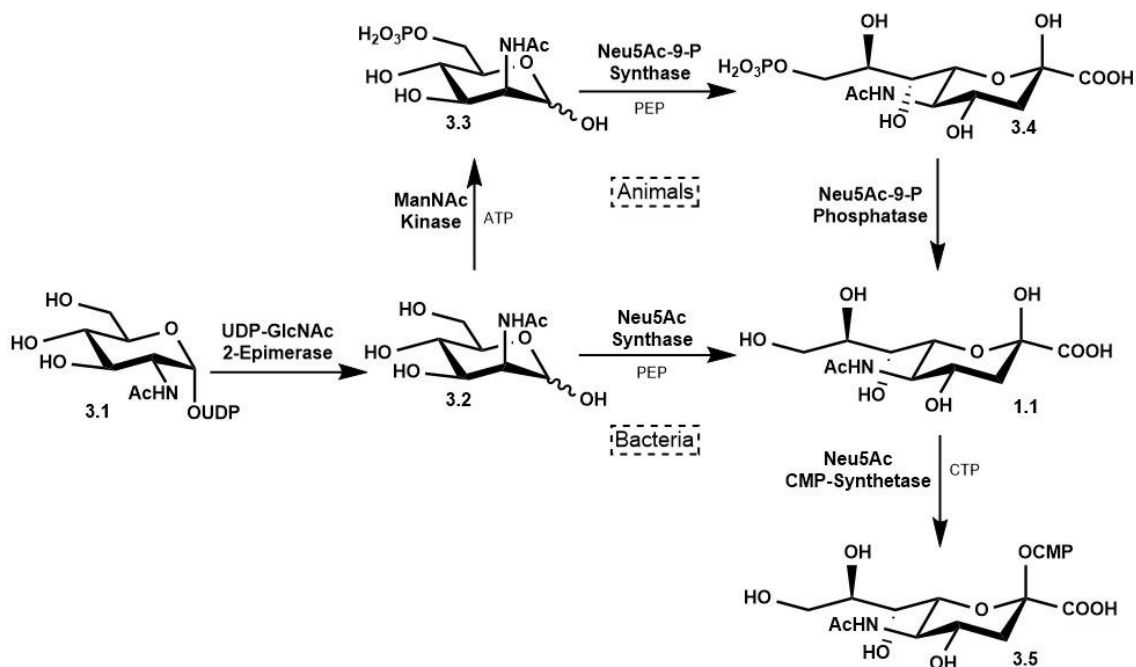
As the supply of pure Pse5Ac7Ac (**1.13**) was limited, attempts at screening new PA2794 purification and crystallisation conditions for enzyme complexes with Pse5Ac7Ac (**1.13**) were discontinued. It was acknowledged that in order to fully characterise this putative pseudaminidase (and other pseudaminic acid processing enzymes), access to a supply of pseudaminic acid chemical probes is required. Therefore this research turned to focussing on the design of a strategy for the large scale (> 100 mg) production of Pse5Ac7Ac (**1.13**) that could be further modified into the desired analogues for probing pseudaminic acid processing enzymes.

## **Chapter 3 Pseudaminic acid biosynthesis**

### 3.1 Introduction

#### 3.1.1 Sialic acid biosynthesis

Commonly, sialic acids are biosynthesised from the activated sugar UDP-GlcNAc (**3.1**),<sup>39</sup> and are then converted to their glycosyl donor activated analogues by enzyme catalysed transfer of a CMP group. In bacteria Neu5Ac (**1.1**) synthesis from UDP-GlcNAc (**3.1**) requires only two enzymes; a UDP-GlcNAc 2-epimerase to hydrolyse the UDP group and alter the stereochemistry at C2,<sup>221</sup> and a Neu5Ac synthase that catalyses the condensation reaction of ManNAc (**3.2**) and phospho-enol-pyruvate (PEP) to produce the nonulosonic acid, Neu5Ac (**1.1**).<sup>222</sup> In non-mammalian higher organisms the pathway follows the same initial step to ManNAc (**3.2**) however prior to the condensation step, a ManNAc kinase phosphorylates the C6 position (**3.3**).<sup>223</sup> Mammals follow the same synthetic pathway but only require one enzyme; the bifunctional UDP-GlcNAc 2-epimerase/ManNAc kinase to synthesise ManNAc-6P (**3.3**).<sup>224</sup> This group is consequently removed after formation of the nonulosonic acid (**3.4**) to produce Neu5Ac (**1.1**), in all organisms Neu5Ac (**1.1**) can then be activated with CTP form CMP-Neu5Ac (**3.5**) to enable transfer onto other structures (**Scheme 3.1**). It was hypothesised that the Pse5Ac7Ac (**1.13**) biosynthetic pathway would resemble this biosynthetic route; with enzymes catalysing the conversion of the UDP-GlcNAc starting material into an intermediate that could undergo a Pse5Ac7Ac synthetase catalysed reaction to produce Pse5Ac7Ac (**1.13**).

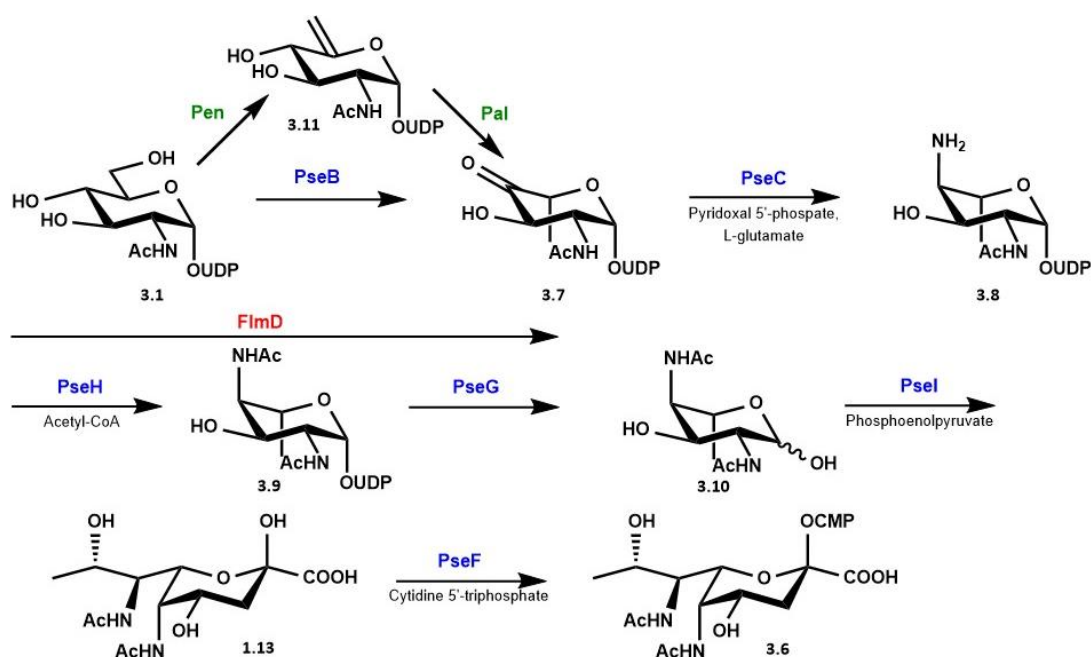


**Scheme 3.1** The biosynthetic pathway from UDP-GlcNAc (**3.1**) to CMP-Neu5Ac (**3.5**) in bacteria and mammals.

### 3.1.2 Identification of the pseudaminic acid biosynthetic enzymes

*H. pylori* and *C. jejuni* flagella are predominantly glycosylated with Pse5Ac7Ac (**1.13**), and derivatives, therefore the flagellin glycosylation gene clusters in these bacteria were inspected for potential pseudaminic acids biosynthetic genes. Schirm *et al.* initially identified four genes (*HP0326A*, *HP0326B*, *HP0840* and *HP0178*) in *H. pylori* that were found to be essential for production of CMP-Pse5Ac7Ac (**3.6**).<sup>180</sup> *HP0840* was identified as the first enzyme in the Pse5Ac7Ac (**1.13**) biosynthetic pathway and upon inactivation of this gene there was an accumulation of UDP-GlcNAc (**3.1**) providing evidence for this as the Pse5Ac7Ac (**1.13**) biosynthetic starting material.<sup>225</sup> Sequence homology assigned the *HP0178* enzyme as a Neu5Ac synthetase and identified the *HP0326A* protein to have the highest homology with CMP-sialic acid synthetases.<sup>180</sup> Thus it was proposed that the final stages of the biosynthetic pathway to CMP-Pse5Ac7Ac (**3.6**) did indeed mimic that of CMP-Neu5Ac (**3.5**) and that other biosynthetic enzymes would be required to process UDP-GlcNAc (**3.1**) into the Pse5Ac7Ac synthetase precursor (**3.10**). *HP0326B* displayed homology with a GT suggesting it may be required for the processing of CMP-Pse5Ac7Ac (**3.6**) rather than involved in Pse5Ac7Ac (**1.13**) synthesis. However upon mutation of this gene there was an accumulation of Pse5Ac7Ac biosynthetic intermediates (**3.7**, **3.8**, and **3.9**) rather than the expected CMP-Pse5Ac7Ac (**3.6**) that would accumulate if the gene encoded for a GT.<sup>180</sup>

UDP-GlcNAc (**3.1**) dehydratase/aminotransferase pairs were isolated and their reaction products analysed by NMR. *H. pylori* (*HP0840/HP0366*) and *C. jejuni* (*Cj1293/Cj1294*) products (**3.7** and **3.8**) were assigned as the first two intermediates in the Pse5Ac7Ac (**1.13**) biosynthetic pathway based on their observed stereochemistry (**Scheme 3.2**).<sup>226</sup> *HP0326B* and *HP0327* were identified as potential Pse5Ac7Ac (**1.13**) biosynthetic genes by gene comparison to bacterial species known to produce CMP-Pse5Ac7Ac (**3.6**). Sequence homology of *HP0327* suggested the enzyme was the acetyl-transferase required to install the second acetamido group (**3.9**).<sup>227</sup> Although *HP0326B* had originally been assigned as a GT and not part of the biosynthetic pathway, it was suggested that if the sugar was “transferred” onto a water molecule, this would result in the apparent UDP hydrolysis producing the desired sugar (**3.10**).<sup>228</sup> Thorough structural and biochemical analyses of each enzyme has led to confirmation of all of the biosynthetic intermediates and have been shown as analogous in *H. pylori* and *C. jejuni*. These five biosynthetic enzymes required for the production of Pse5Ac7Ac (**1.13**) are now generally termed as PseB, PseC, PseH, PseG, PseI with the Pse5Ac7Ac CMP-synthetase labelled as PseF (**Scheme 3.2**).<sup>203</sup>



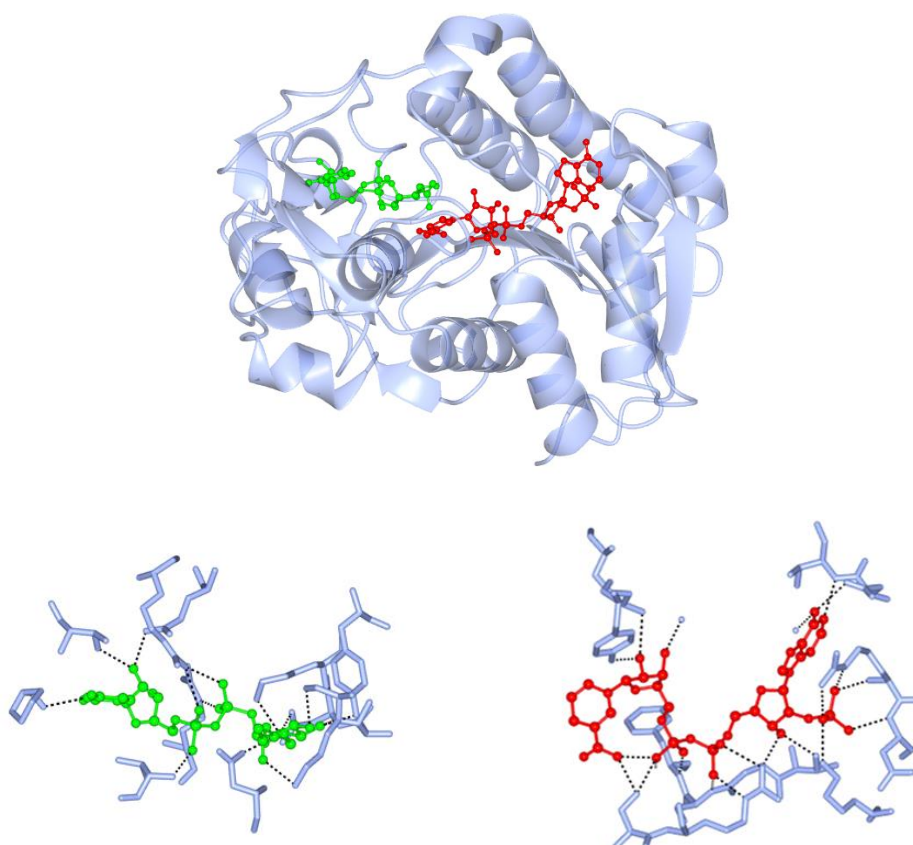
**Scheme 3.2** The CMP-Pse5Ac7Ac (**3.16**) biosynthesis detailing the *C. jejuni* and *H. pylori* pathway (blue) with any enzymatic deviations observed in *A. caviae* (red) and *B. thuringiensis* (green) highlighted.

A slightly different Pse5Ac7Ac (**1.13**) biosynthetic pathway has been assigned in *A. caviae*; another pathogen that displays Pse5Ac7Ac (**1.13**) on its flagellin as well as in the LPS. Mutations to the *flm* gene locus caused loss of motility, flagella and the LPS suggesting this locus has a role in flagellar assembly and LPS biosynthesis.<sup>229</sup> A cluster of genes in this locus displayed homology to the biosynthetic genes found in *C. jejuni* and *H. pylori* and comparison displayed conserved domains for FlmA with PseB, FlmB with PseC and NeuB with PseI. However in *A. caviae* the FlmD protein was much larger than expected and displayed conserved domains with both PseH and PseG suggesting this protein catalyses both the desired amine acetylation and UDP-hydrolysis.<sup>229</sup> The *A. caviae* enzymes have not been as rigorously investigated as the *C. jejuni* and *H. pylori* enzymes and full biochemical analysis of the route in this organism is yet to be carried out.

Further deviations from the original Pse5Ac7Ac (**1.13**) biosynthetic pathway were discovered when the pathway from the gram positive bacteria *B. thuringiensis* was investigated. An operon was identified containing seven genes that were proposed to encode enzymes in the CMP-Pse5Ac7Ac (**1.13**) biosynthetic pathway.<sup>182</sup> Sequence alignment allocated homologues for the PseC, PseH, PseG, PseI and PseF enzymes and two *B. thuringiensis* enzymes that were predicted as dehydratases. LC/MS and NMR characterisation of the product of each enzymatic reaction showed that two enzymes were required to carry out the PseB function. The first enzyme (Pen) converts UDP-GlcNAc (**3.1**) into UDP-6-deoxy-D-GlcNAc-5,6-ene (**3.11**) and the second (Pal) acts as a C4 oxidase and C5,6 reductase resulting in epimerisation at C5 compared to the original starting material (**3.1**).<sup>182</sup>

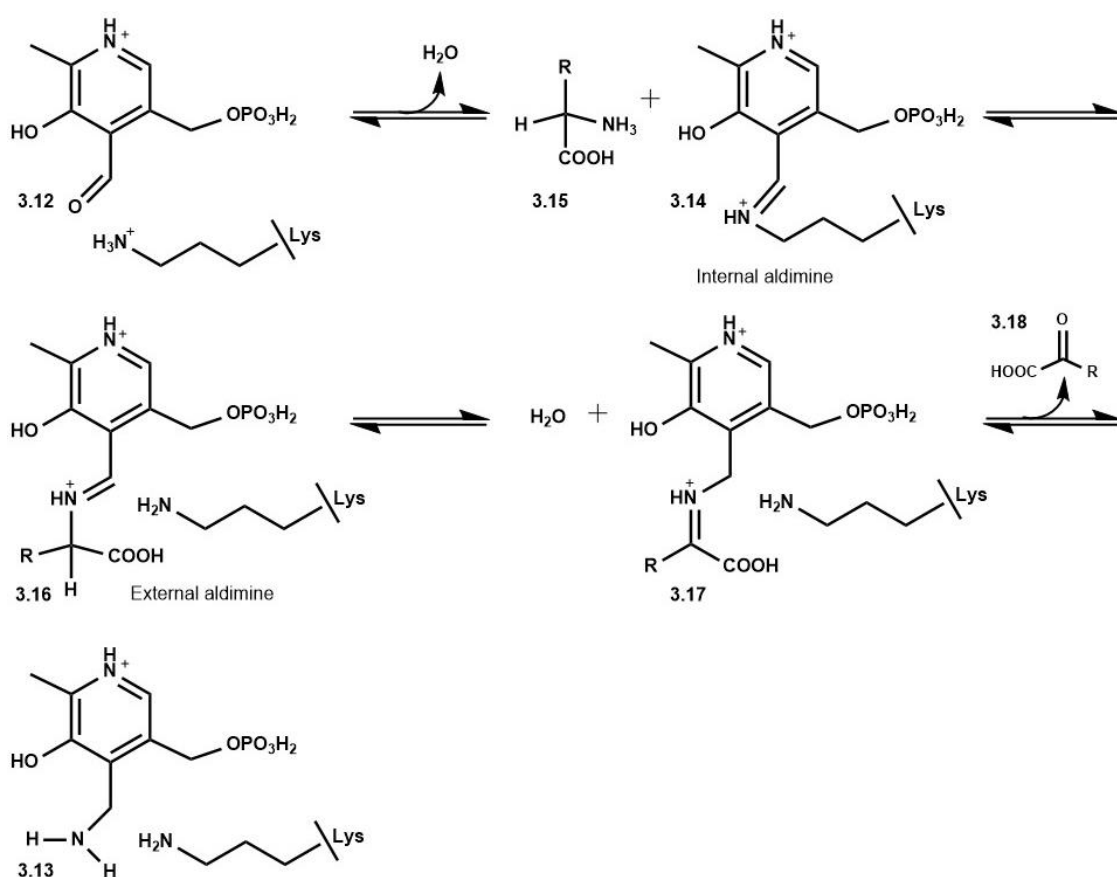
### 3.1.3 Pseudaminic acid biosynthetic enzyme structure and function

PseB is the first enzyme in the pathway and converts UDP-GlcNAc (**3.1**) to UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**). The proposed reaction mechanism follows three sequential steps; oxidation of the C4 hydroxyl, dehydration at C6 to form the alkene, followed by reduction to the methyl.<sup>230</sup> Initially controversy over the *H. pylori* PseB co-factor occurred with spectrophotometric studies showing that NADP<sup>+</sup> in the reaction mixture was not used by the enzyme at all and that additional NAD<sup>+</sup> could be used but with a very poor efficiency.<sup>225</sup> Furthermore conversion to the PseB product (**3.7**) was observed without addition of potential co-factor molecules. Therefore it was suggested that PseB may behave equivalently to some UDP-GlcNAc C4 epimerases; with the NAD(P)<sup>+</sup> co-factor remaining tightly bound to the enzyme during purification and throughout the reaction with internal regeneration. This was confirmed when crystal structures of PseB revealed electron density concurrent with a bound NADPH molecule even without addition of exogenous molecules to the enzyme (**Figure 3.1**).<sup>230</sup> The deeply buried co-factor is surrounded by residues orientated to have favourable interactions with the molecule including multiple H-bonding residues (**Figure 3.1**).



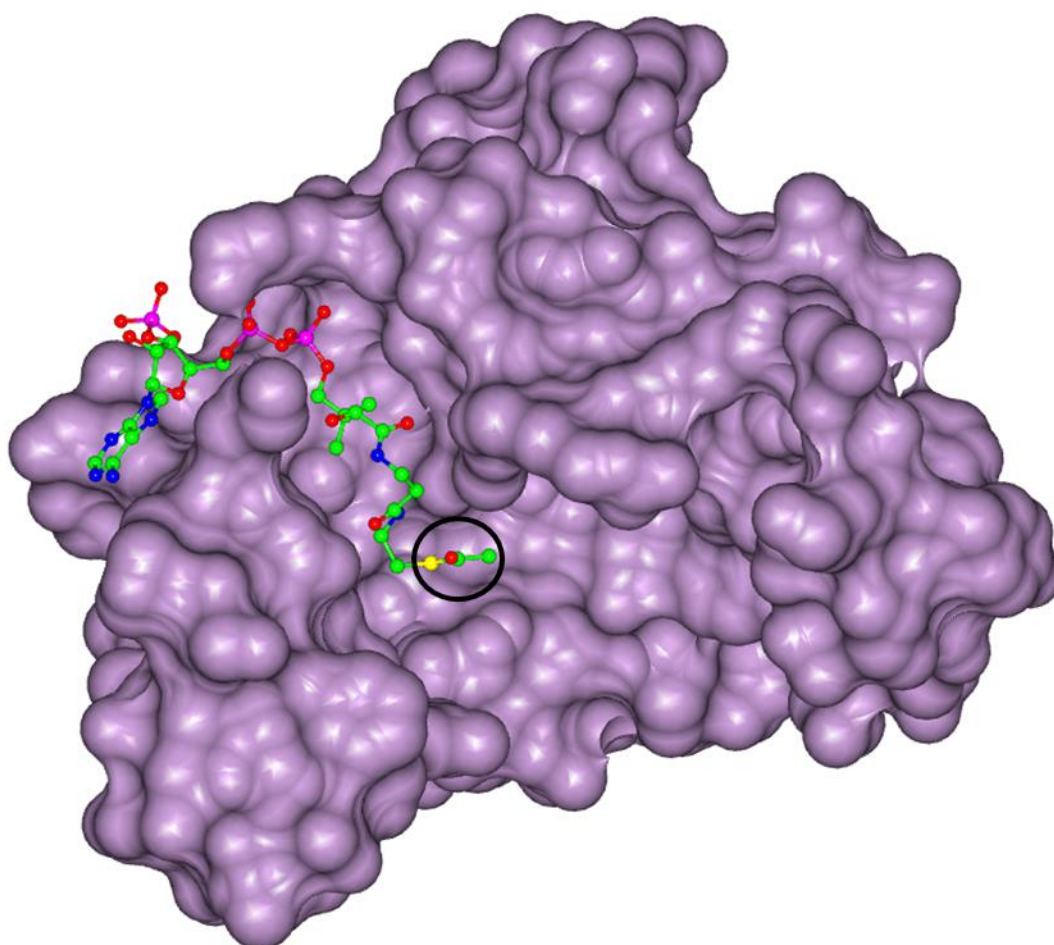
**Figure 3.1** *H. pylori* PseB (PDB 2GN4)<sup>230</sup> crystal structure in complex with the UDP-GlcNAc substrate (**3.1**) (green) and the tightly bound NADPH co-factor (red) and close-ups of the substrate and co-factor binding sites displaying the numerous hydrogen bonds between molecules and enzyme.<sup>230</sup>

PseC was identified as the second enzyme in the Pse5Ac7Ac (**1.13**) biosynthetic pathway, acting as an aminotransferase to catalyse synthesis of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) from the 4-keto derivative (**3.7**). *H. pylori* PseC was found to share sequence similarity with PLP-dependent aminotransferases, and assigned as a Type 1 aminotransferase.<sup>231</sup> PLP-dependent aminotransferases undergo two subsequent half-transamination reactions; an acetyl transfer to the PLP (**3.12**) co-factor by an acetyl donor followed by aminotransfer from this to the sugar (**3.7**). Lys183 was predicted as the catalytic active site residue in *H. pylori* PseC. During the first half transamination, initially an enzyme-PLP Schiff base forms (**3.14**), then a free amino donor (**3.15**) releases the enzyme from the internal aldimine (**3.14**) forming an external aldimine (**3.16**) via a transaldimination reaction. Finally hydrolysis occurs resulting in release of a glyoxylic acid (**3.18**) to produce free PMP (**3.13**) in the active site (**Scheme 3.3**).<sup>231</sup>



**Scheme 3.3** The first half-transamination reaction of PLP-dependent aminotransferases, catalysing conversion of PLP (**3.11**) to PMP (**3.12**) utilising a free amino donor (**3.14**).

PseH is the third enzyme in the *H. pylori* and *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic pathways and catalyses acetyl transfer from acetyl-CoA to the C4 amine formed in the previous step to yield UDP-4-amino-4,6-dideoxy-AltdiNAc (**3.9**).<sup>203</sup> Surprisingly it was found that *C. jejuni* PseH displays very little structural and binding similarity to other acetyltransferases associated with the biosynthesis of sugars for protein glycosylation. However crystallisation of PseH in complex with the co-factor shows the acetyl group protruding into the predicted substrate binding groove in a favourable position for direct acetyl transfer to occur to the desired C4 amino sugar substituent (**Figure 3.2**).<sup>232</sup>

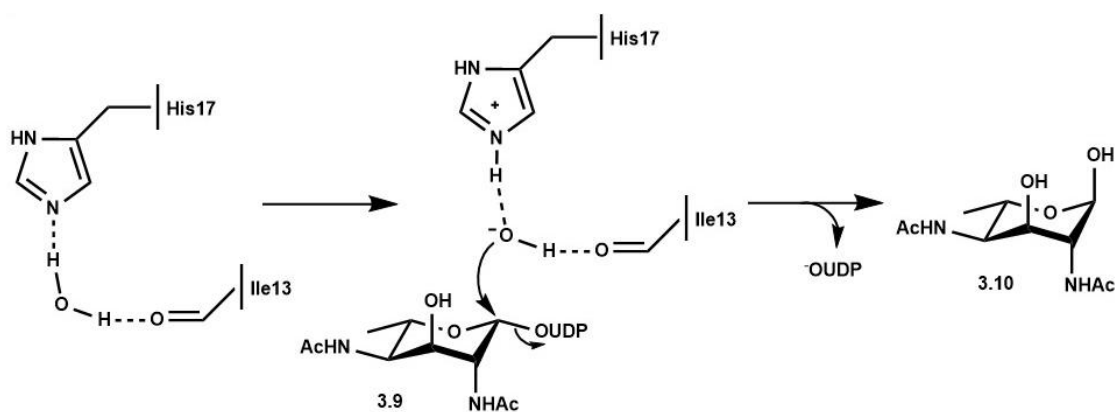


**Figure 3.3** *C.jejuni* PseH (PDB 4XPL)<sup>232</sup> crystal surface structure (purple, globular) in complex with the acetyl-CoA co-factor (ball and stick model) with the thio-acetyl group (circled) protruding into the predicted substrate binding groove.



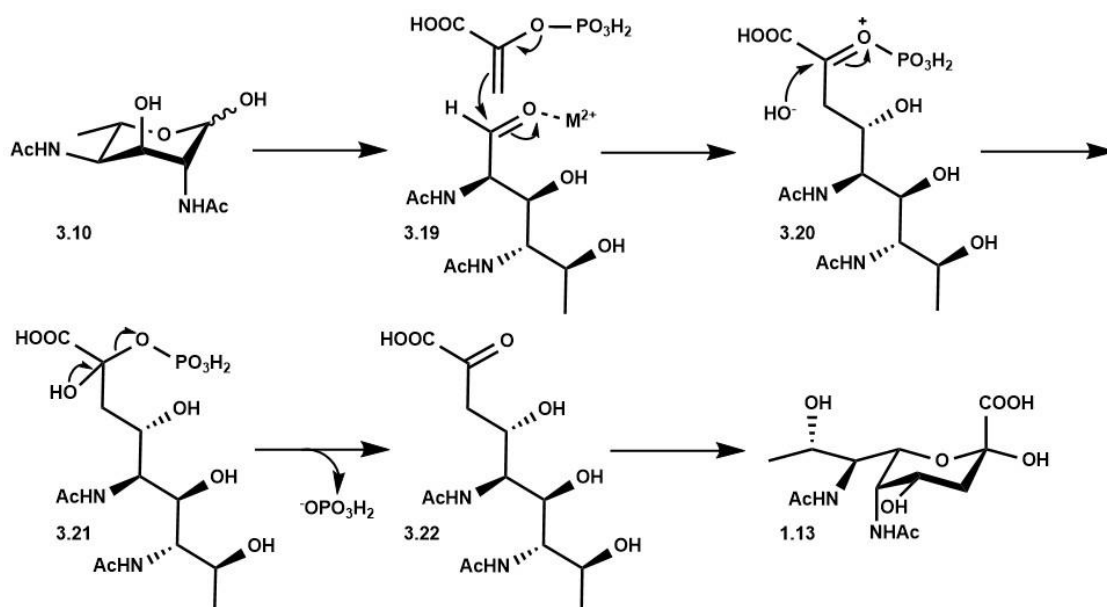
The *PseG* gene locus (*Cj1312*) *H. pylori* homologue was originally assigned to code for a putative glycosyltransferase as it displayed significant alignment with conserved sequences of some UDP-sugar transferases.<sup>180</sup> PSI-BLAST homology sequencing of *Cj1312* classified it as belonging to the metal-independent GT-B superfamily,<sup>233</sup> however the amino acid sequence displays only modest overall identity with this class of protein. However due to a lack of other candidates, and its genome positioning, this gene was also highlighted as a potential glycosyl hydrolase for catalysing the desired cleavage of the UDP group during Pse5Ac7Ac (**1.13**) biosynthesis.<sup>228</sup>

Biochemical analysis confirmed the role of PseG as a UDP-hydrolase, with the enzyme thought to employ a water molecule as the acceptor to yield the Pse5Ac7Ac precursor; 6-deoxy-AltdiNAC (**3.10**).<sup>228</sup> Crystal structures highlighted an active site histidine residue that was found to be essential for hydrolysis and was proposed to abstract a water proton to activate it as a nucleophile.<sup>234</sup> Investigations into the enzyme mechanism in the presence of D<sub>2</sub>O displayed that UDP hydrolysis of UDP-4-amino-4,6-dideoxy-AltdiNAC (**3.9**) results in an inversion of the stereochemistry at C1 (**Scheme 3.4**).<sup>228</sup>



**Scheme 3.4** The proposed mechanism for PseG catalysed UDP hydrolysis, utilising a histidine-activated and isoleucine-stabilised water molecule as a nucleophile to attack at the anomeric centre.<sup>234</sup>

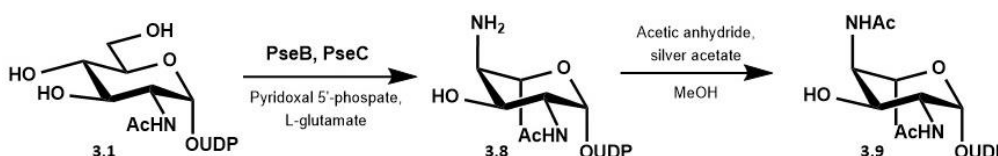
The final enzyme, PseI was the first to be identified as being involved in the Pse5Ac7Ac (**1.13**) biosynthetic pathway as it displays homology to Neu5Ac synthetase sequences and mimics the final biosynthetic step for the production of sialic acids.<sup>180</sup> This final step requires a sialic acid synthase (SAS) to catalyse the condensation of phosphoenolpyruvate with the sugar resulting in the desired nonulosonic acid.<sup>235</sup> Genes displaying conserved features with other SAS were identified in *C. jejuni* and biochemical experiments confirmed one enzyme (PseI) that converted 6-deoxy-AltdiNAC (**3.10**) (a potential Pse5Ac7Ac precursor) to Pse5Ac7Ac (**1.13**).<sup>236</sup> The mechanism employed by the pseudaminic acid synthetase was found to be a metal dependent, C-O bond cleavage mechanism that proceeds in the ring open form (**3.19**) *via* generation of an oxocarbenium ion (**3.20**). Attack by a free hydroxyl produces a tetrahedral intermediate (**3.21**) which is followed by release of the phosphate group resulting in the formation of a nonulosonic acid (**3.22**) which can then ring close to Pse5Ac7Ac (**1.13**) (**Scheme 3.5**).<sup>236</sup>



**Scheme 3.5** *C. jejuni* PseI pseudaminic acid synthetase PEP condensation mechanism.

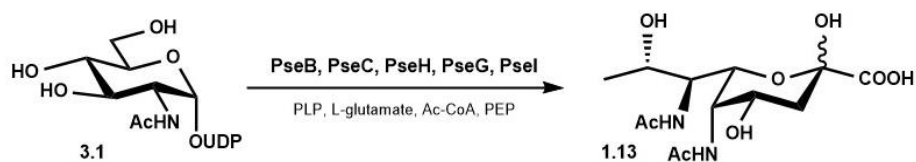
### 3.1.4 Exploitation of the pseudaminic acid biosynthetic pathway for synthesis of pseudaminic acid analogues

During the biochemical analyses of the enzymes in the Pse5Ac7Ac (**1.13**) biosynthetic pathway, access to the enzyme substrates was required but some are not commercially available or easy to chemically derive. In particular, it was deemed impractical to chemically synthesise the proposed PseG substrate (**3.9**) with previous synthetic routes being costly, lengthy and extremely low yielding.<sup>196</sup> Therefore PseB and PseC were utilised as synthetic tools to convert UDP-GlcNAc into UDP-4-amino-2,4,6-trideoxy- $\beta$ -L-AltNAc (**3.8**) which could then be selectively chemically acetylated to yield the desired product (**3.9**) (**Scheme 3.6**). Although the synthesis of UDP-4-amino-4,6-dideoxy-AltNAc (**3.8**) displayed high conversion, significant amounts of material were lost during intermediate purification steps reducing the overall yield to just 34 % of desired product (**3.9**).<sup>228</sup>



**Scheme 3.6** A chemoenzymatic route to the PseG substrate, UDP-6-deoxy-AltNAc (**2.9**); incubation of PseB, PseC, PLP and L-Glu in 50 mM Tris-HCl, pH 7.4, 37 °C, followed by acetylation with acetic anhydride and silver acetate in methanol.<sup>228</sup>

Knowledge and potential exploitation of the full Pse5Ac7Ac (**1.13**) biosynthetic pathway provides an alternative route to the synthesis of Pse5Ac7Ac (**1.13**) and displays significant advantages over chemical synthesis. For example, stereochemistry is controlled, no toxic reagents are required and the synthesis does not require lengthy, complicated steps. Schoenhofen *et al* showed that, once purified, all of the recombinant biosynthetic enzymes could be combined along with their co-factors and UDP-GlcNAc (**3.1**), to produce Pse5Ac7Ac (**1.13**) in a “one-pot” reaction with a quantitative yield (**Scheme 3.7**). However less than 5 mgs of Pse5Ac7Ac (**1.13**) was produced using this method with the major limitation being the requirement for very expensive co-factors. Another disadvantage to this scheme was that, although quantitative conversion to Pse5Ac7Ac (**1.13**) was reported, the methods for protein production and purification had been generalised and as such may not be the conditions for optimal production of each enzyme.<sup>203</sup>



**Scheme 3.7** One-pot enzymatic production of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**); incubation of PseB, PseC, PLP, L-Glu, acetyl-CoA and PEP in 25 mM sodium phosphate, 50 mM NaCl pH 7.3, 37 °C.<sup>203</sup>

This project began with the aim to standardise a route to Pse5Ac7Ac (**1.13**) that did not suffer from the shortcomings of the existing chemical and enzymatic syntheses. In particular it was desired to find a route whereby Pse5Ac7Ac (**1.13**) could feasibly be produced on a large scale > 50 mg. It was decided that utilisation of enzymes from the Pse5Ac7Ac (**1.13**) biosynthetic pathway had the most scope for this; to gain maximum yields at minimal cost. Therefore initial focus fell on optimising production of the biosynthetic enzymes and probing their activity for production of Pse5Ac7Ac (**1.13**).

## 3.2 Pseudaminic acid biosynthetic enzyme production

### 3.2.1 Optimising enzyme expression conditions

PseB and PseC recombinant plasmids were a kind gift from the Tanner lab with the PseB pET30a recombinant plasmid designed to express C-terminal HexaHis tagged PseB and a PseC pFO4 recombinant plasmid to express N-terminal HexaHis tagged PseC. The *PseH*, *PseG* and *PseI* genes were codon optimised for *E. coli* and recombinant plasmids designed and purchased from GenScript (**Appendix 6**). The pET15b vectors were utilised to yield N-terminal HexaHis tagged sequences for each gene with expression controlled by the T7 promoter. The PseB and PseC recombinant plasmids were transformed *via* electroporation into *E. coli* BL21 DE3 cells and PseH, PseG and PseI recombinant plasmids transformed *via* heat shock. Successful transformation confirmed by growth on antibiotic-containing LB agar plates (artificial selection) (**Table 3.1**).

**Table 3.1** Table of the recombinant plasmids containing pseudaminic acid biosynthetic enzymes and their associated antibiotic resistance.

Enzyme	Bacterial source	Plasmid	Restriction Enzyme sites	Antibiotic resistance
PseB	<i>C. jejuni</i>	pET30a	NdeI-XhoI	Kanamycin
PseC	<i>C. jejuni</i>	pFO4	BamHI-EcoRI	Ampicillin
PseH	<i>C. jejuni</i>	pET15b	BamHI-EcoRI	Ampicillin
PseG	<i>C. jejuni</i>	pET15b	NdeI-BamHI	Ampicillin
PseI	<i>C. jejuni</i>	pET15b	NdeI-BamHI	Ampicillin

Initially small scale grow ups were employed to confirm transformation, expression and purification techniques for the desired biosynthetic protein. Incorporation of the PseB, PseC, PseH, PseG and PseI expressing plasmids did not appear to be detrimental to bacterial growth which was monitored prior to induction of the protein of interest. OD<sub>600</sub> measurements displayed that the growth of cells occurred at different rates; with PseH reaching the desired OD<sub>600</sub> significantly faster than the others, and PseI significantly slower. It was hypothesised that the smaller length of *PseH* could aid in the smaller lag time however this only infers a relatively small change when comparing the sequence length of the entire plasmid (**Appendix 6**). Therefore it is expected that the plasmid copy number would be similar for all recombinant plasmids and that the observed differences was due to discrepancies in the health of the cells post transformation rather than any negative effects of basal protein expression. Although using the induction conditions detailed by Schoenhofen (0.1 mM IPTG, 37 °C, 4 hrs) produced enough protein to synthesise Pse5Ac7Ac (**1.13**) on a small scale, expression trials were carried out to optimise production and make this route more feasible for large scale synthesis. Three different

variables were explored in order to find optimum conditions for protein production; IPTG induction concentration, temperature post induction and time of incubation post induction (**Table 3.2**).

**Table 3.2** Expression trial conditions to optimise production of the biosynthetic enzymes.

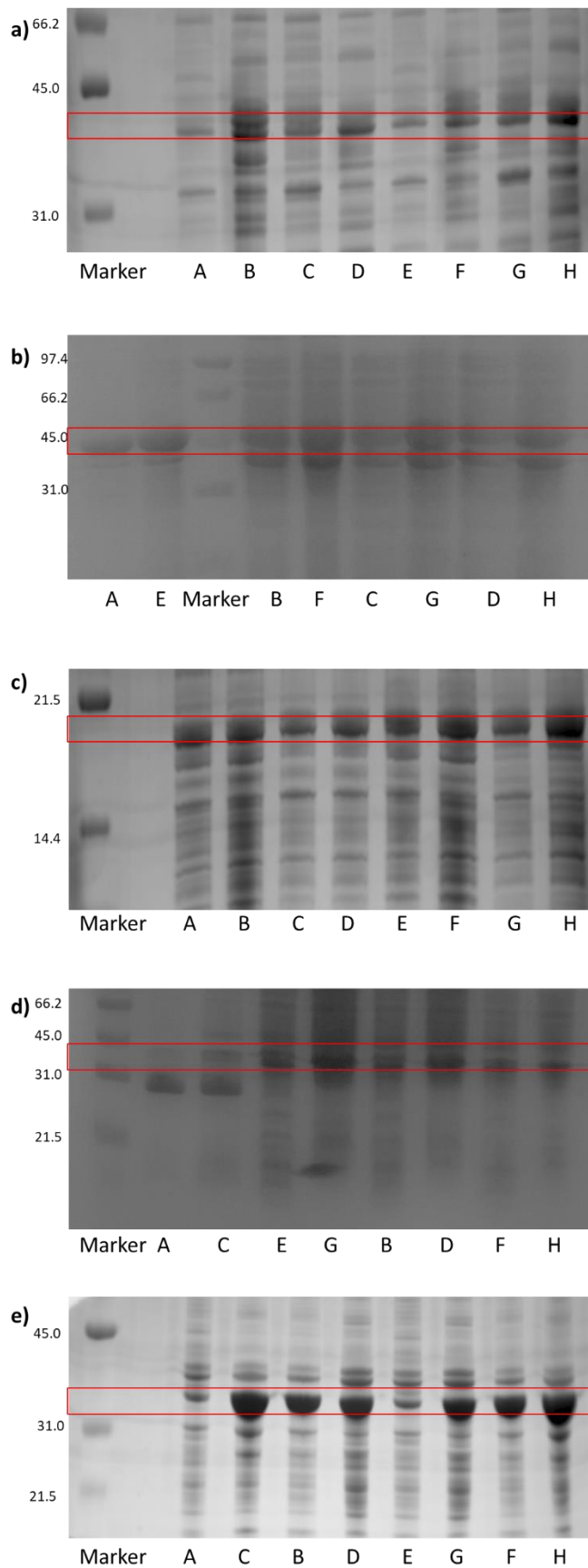
Condition	A	B	C	D	E	F	G	H
Post induction incubation temp	16 °C				37 °C			
Post induction concentration	0.1 mM		0.5 mM		0.1 mM		0.5 mM	
Post induction incubation time	4 hrs	20 hrs	4 hrs	20 hrs	4 hrs	20 hrs	4 hrs	20 hrs

Protein production was monitored after lysis by imaging on SDS-PAGE and comparison of which conditions produced the most prominent protein bands (**Figure 3.4**). Only marginal increases in protein production of PseB and PseC were displayed; with little difference apparent between the two concentrations of IPTG and the time left incubating after IPTG induction. However by incubating at 37 °C instead of 16 °C, a more concentrated protein was observed indicative of higher protein production (**Figure 3.4a, b**). PseH protein production was equally deemed unaffected by IPTG concentration but production was increased when incubating for a longer period of time and marginally by the increase in temperature post induction (**Figure 3.4c**).

Analysis of the band of interest on the SDS-PAGE gel demonstrated that using the higher temperature post induction was the variable that most improved production of PseG and PseI production (**Figure 3.4d, e**). Additionally adding more IPTG had a small positive effect, as did incubating at 37 °C compared to 16 °C. Therefore these modified conditions (**Table 3.3**) were employed during the subsequent large scale enzymatic grow ups in order to maximise enzyme production and hence reduce resources required for production of Pse5Ac7Ac (**1.13**) on a large scale.

**Table 3.3** Optimised conditions for production of pseudaminic acid biosynthetic enzymes.

	IPTG concentration / mM	Post induction incubation temp / °C	Post induction incubation time / hrs
PseB	0.1	37	4
PseC	0.1	37	4
PseH	0.1	16	20
PseG	0.5	37	4
PseI	0.5	37	20



**Conditions:**

**A** - 16 °C, 0.1 mM IPTG, 4 hrs.

**B** - 16 °C, 0.1 mM IPTG, 20 hrs.

**C** - 16 °C, 0.5 mM IPTG, 4 hrs.

**D** - 16 °C, 0.5 mM IPTG, 20 hrs.

**E** - 37 °C, 0.1 mM IPTG, 4 hrs.

**F** - 37 °C, 0.1 mM IPTG, 20 hrs.

**G** - 37 °C, 0.5 mM IPTG, 4 hrs.

**H** - 37 °C, 0.5 mM IPTG, 20 hrs.

**Figure 3.3** SDS PAGE analysis of crude samples of the biosynthetic enzymes under different induction conditions **a)** PseB, **b)** PseC, **c)** PseH, **d)** PseG, **e)** PseI.

### 3.2.2 Purification of the biosynthetic enzymes

Following large scale production of the desired enzymes, the protein of interest was purified from other buffer soluble proteins by loading onto a Nickel affinity column and applying an increasing imidazole concentration gradient. Presence of a UV Abs<sub>280</sub> peak during washing with higher concentrations of imidazole (after loading and washing the column with low concentrations of imidazole) confirmed over-expression of the His<sub>6</sub>-tagged protein. This peak occurred at similar concentrations of imidazole across purification of each biosynthetic enzyme (100-150 mM imidazole) (**Appendix 7**). The eluted proteins corresponding to these UV Abs<sub>280</sub> peaks were confirmed as the desired biosynthetic enzymes by SDS-PAGE analysis which displayed over-expressed, pure protein of the expected molecular weight (**Figure 3.5**). Favourably these larger scale expressions appeared to further increase the amount of purified enzyme per volume of media which was attributed to economies of scale with predicted lower percentage of cell loss during the collection and lysis of cells, and enzymatic loss during purification. Production of purified enzyme under these conditions (**Table 3.4**) was either equal to or greater than previously reported yields (7-11 mgL<sup>-1</sup> LB).<sup>203</sup>

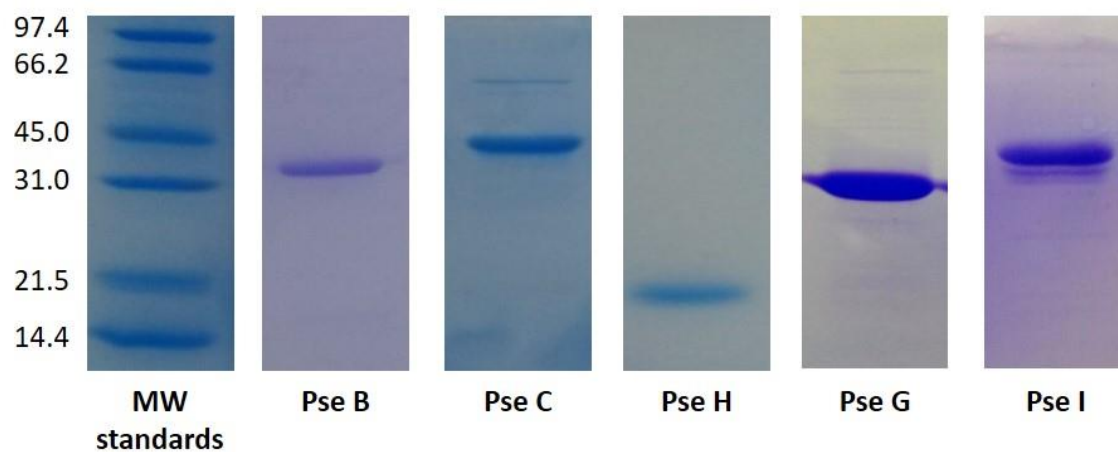
Unfortunately, this increase in enzyme concentration during purification proved detrimental to the yield of soluble Pse C, with enzyme precipitation occurring during purification. As precipitation had not appeared during the smaller scale preparations it was hypothesised that aggregation could be a result of the increased concentration of PseC in imidazole containing buffer (a molecule that has previously been shown to affect protein stability in high concentrations). Dialysis buffer was therefore added to each elution fraction prior to purification of PseC to dilute its concentration (and that of the imidazole) as soon as possible after elution. However, precipitation still occurred using this method, even when adding detergents, such as glycerol and PEG, to the dialysis buffer. Subsequently, lysed cells containing PseC were purified in portions to mimic the smaller scale original purification process, unfortunately, not only was this method both time and cost ineffective, but it was to no avail and precipitation still occurred.

It was hypothesised that modified PseC induction conditions, rather than purification conditions, may account for the apparent aggregation of PseC after purification. Therefore a large scale grow up of PseC was carried out using the lower induction temperature of 16 °C and still only left to incubate for 4 hours. Upon purification this batch of protein did not appear to precipitate and was subjected to desalting ahead of storage without apparent aggregation. Although these conditions were not optimum for PseC production, it was deemed a necessary reduction in order to maintain pure PseC solubility in the desired buffer system and hence these were used for all further protein preparations.



**Table 3.4** The quantity of each biosynthetic enzyme purified per L of LB culture.

	Molecular weight / kDa	Quantity of enzyme per L LB / mg
<b>PseB</b>	37.4	14
<b>PseC</b>	42.3	9
<b>PseH</b>	18.7	17
<b>PseG</b>	31.3	11
<b>PseI</b>	38.6	18



**Figure 3.4** Depiction of the purified *C. jejuni* Pse5Ac7Ac (1.13) biosynthetic enzymes on a SDS PAGE gel.

### 3.3 Enzymatic synthesis of pseudaminic acid

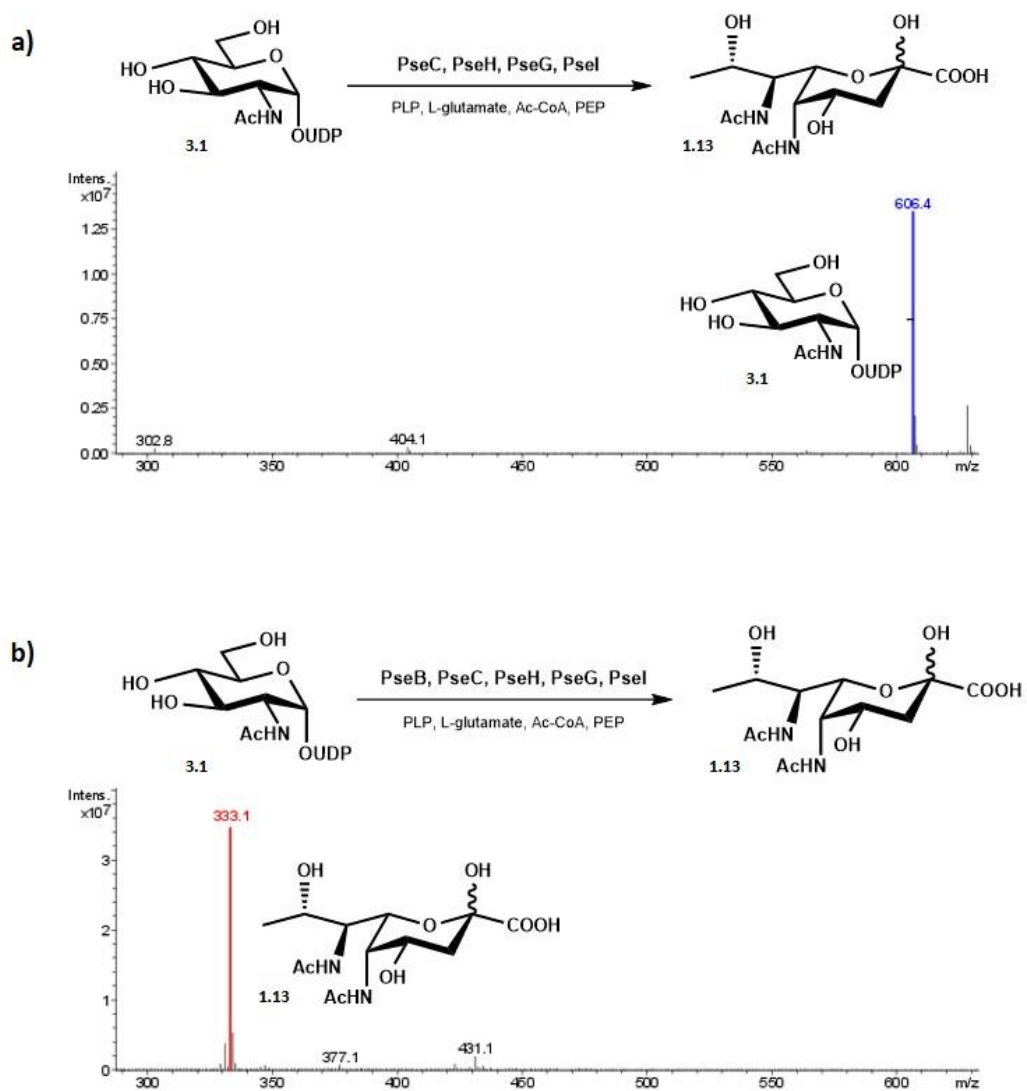
#### 3.3.1 Enzymatic synthesis of pseudaminic acid utilising enzymes from the *Campylobacter jejuni* biosynthetic pathway

Subsequent to the large scale purifications of the biosynthetic enzymes, standardisation of the “one-pot” enzymatic synthesis of Pse5Ac7Ac (**1.13**) was attempted *in house*. Following the procedure outlined previously,<sup>203</sup> a small-scale reaction mixture (0.1 mg) was constructed utilising the pure enzymes. UDP-GlcNAc (**3.1**) and the five biosynthetic enzymes were incubated (shaken 120 rpm, 37 °C) in a sodium phosphate buffer (50 mM, pH 7.4) with an excess of co-factors (**Table 2.5**). A negative control reaction was also set up utilising identical conditions as above but without the inclusion of the first enzyme in the biosynthetic pathway; PseB and the reactions were monitored at time intervals by negative ion ESI LC-MS.

Analysis of the negative control was run in parallel to the reaction mixture and importantly there was no detection of a peak at  $[M-H]^-$  333 Da, or appearance of any additional peaks compared to when the reaction mixture was first constructed. Thus suggesting that the UDP-GlcNAc (**3.1**) starting material was stable in the “one-pot” reaction mixture and that the biosynthetic enzymes could not catalyse any side-reactions without PseB first converting UDP-GlcNAc (**3.1**) to UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**) (**Figure 3.6a**). Appearance of a peak corresponding to production of Pse5Ac7Ac (**1.13**) ( $[M-H]^-$  333 Da) was observed in the reaction mixture containing all of the biosynthetic enzymes (**Figure 3.6b**). Under these conditions, no peaks corresponding to the biosynthetic intermediate products were observed and full conversion to Pse5Ac7Ac (**1.13**) was estimated after two and a half hours, when the LCMS peak area relating to the starting material UDP-GlcNAc  $[M-H]^-$  606 Da could no longer be detected (**Figure 3.6b**).

**Table 3.5** General reaction components and concentrations for the “one-pot” small scale enzymatic synthesis of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**2.1**)

Reagent	Concentration / mM
UDP-GlcNAc	1.0
Pyridoxal 5' phosphate (PLP)	1.5
L-glutamic acid	10.0
Acetyl-Coenzyme A	1.5
Phosphoenolpyruvate (PEP)	2.0
PseB, PseC, PseH, PseG, PseI	0.38 mg mL <sup>-1</sup>



**Figure 3.5** Negative ESI LC-MS of the small scale Pse5Ac7Ac (**1.13**) enzymatic reaction, **a)** negative control (reaction mixture without the first biosynthetic enzyme; PseB) after 2.5 hrs, demonstrating substrate UDP-GlcNAc (**3.1**) ( $[M-H]^-$  606 Da) stability and no detection of any of the biosynthetic products and **b)** analysis of the enzymatic synthesis at 2.5 hrs, demonstrating full conversion to the desired Pse5Ac7Ac (**1.13**) product  $[M-H]^-$  333 Da.

### 3.3.2 Enzymatic synthesis of pseudaminic acid utilising enzymes from the *Aeromonas caviae* biosynthetic pathway

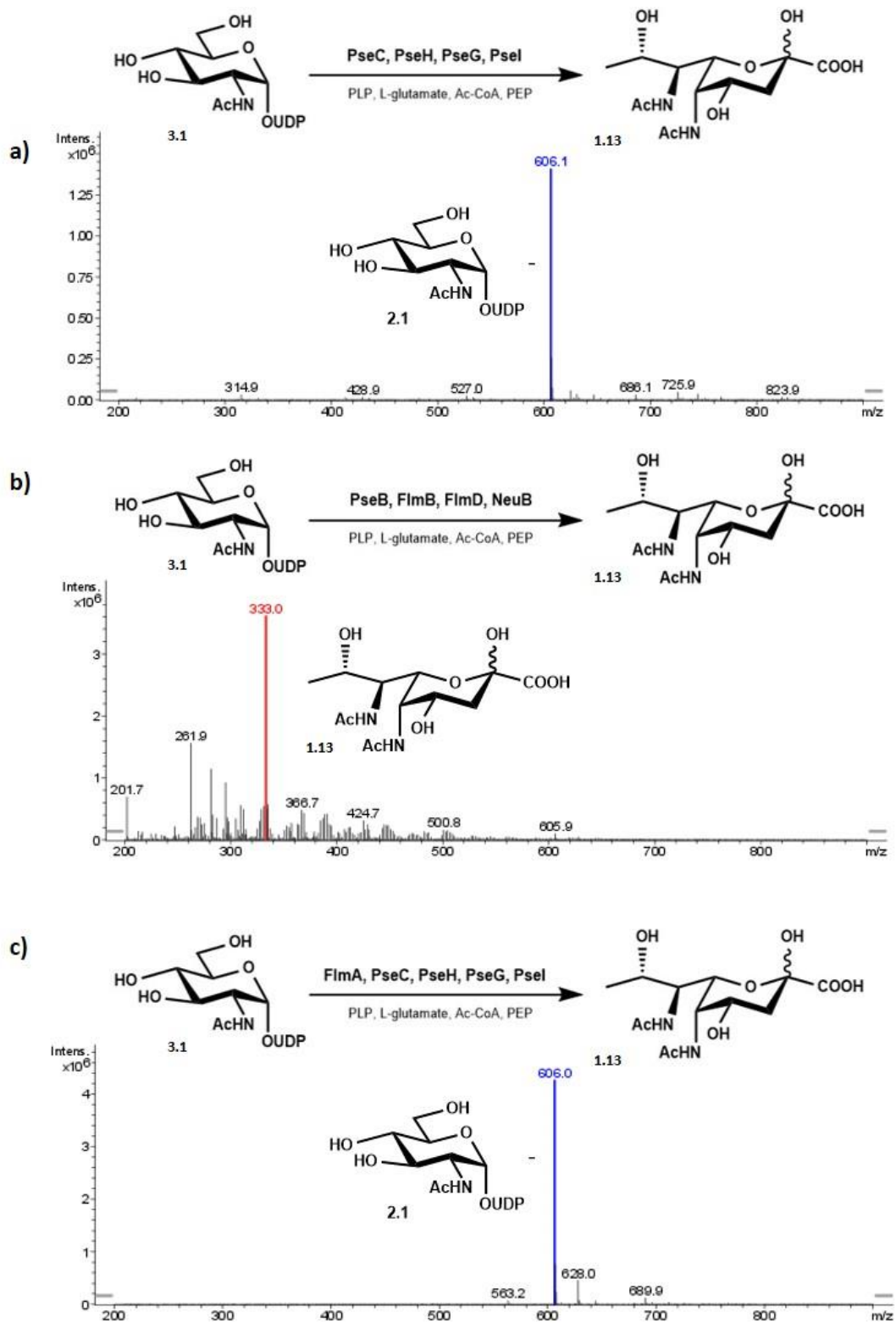
Although synthesis of Pse5Ac7Ac (**1.13**) has been previously demonstrated using the recombinant biosynthetic enzymes encoded by *C. jejuni* and *H. pylori*,<sup>203</sup> the pathway has not been fully elucidated *in vitro* using the recombinant enzymes from *A. caviae* (FlmA, FlmB, FlmD and NeuB) (**Scheme 3.2**). *A. caviae* flagellin and LPS are glycosylated with Pse5Ac7Ac (**1.13**) and when transposon mutants of FlmA, FlmB, FlmD or NeuB were inserted, a loss of motility and LPS O-antigen bands were recorded.<sup>229</sup> Previous studies discovered these four genes in a genetic locus similar to that responsible for Pse5Ac7Ac (**1.13**) biosynthesis in *C. jejuni* and *H. pylori* and showed FlmD and NeuB to display homology with Neu5Ac (**1.1**) biosynthesis enzymes; UDP-GlcNAc 2-epimerase and Neu5Ac synthetase respectively.<sup>229</sup>

In order to ascertain if these enzymes (FlmA, FlmB, Flm D and NeuB) did make up the *A. caviae* Pse5Ac7Ac (**1.13**) biosynthetic pathway, collaborators (Jon Shaw group, University of Sheffield) subjected purified enzymes to similar reaction conditions as the *C. jejuni* enzymes (**Table 3.5**). However under these conditions they were unable to detect turnover of UDP-GlcNAc (**3.1**) to Pse5Ac7Ac (**1.13**), or indeed identify production of any of the biosynthetic intermediates. Therefore it was decided to test these *A. caviae* enzymes in the standardised procedure established for the *C. jejuni* enzymes in order to assess if these enzymes did indeed turnover to produce Pse5Ac7Ac (**1.13**) equivalently.

Reactions were set up using a mixture of biosynthetic enzymes from *A. caviae* and *C. jejuni* so that activity of each individual *A. caviae* enzyme could be established (**Table 3.6**). Reactions containing *C. jejuni* PseB with the rest of the enzymes from either *C. jejuni* or *A. caviae* displayed turnover of UDP-GlcNAc (**3.1**) to Pse5Ac7Ac (**1.13**) with equivalent conversions. However it was found that the reaction mixtures containing FlmA from *A. caviae* displayed no turnover of UDP-GlcNAc (**3.1**) in our standardised system (**Figure 3.7**).

**Table 3.6** The combination of *C. jejuni* and *A. caviae* proposed Pse5Ac7Ac (**1.13**) biosynthetic enzymes trialled during activity experiments to elucidate the *A. caviae* biosynthetic pathway.

Experiment	Enzyme composition
A	<i>C. jejuni</i> PseB, PseC, PseH, PseG, PseI
B	<i>C. jejuni</i> PseB, PseC, PseH, PseG, <i>A. caviae</i> NeuB
C	<i>C. jejuni</i> PseB, PseC, PseH, PseI, <i>A. caviae</i> FlmD
D	<i>C. jejuni</i> PseB, PseH, PseG, PseI, <i>A. caviae</i> FlmB
E	<i>C. jejuni</i> PseC, PseH, PseG, PseI, <i>A. caviae</i> FlmA
F	<i>A. caviae</i> FlmA, FlmB, FlmD, NeuB



**Figure 3.6** Negative ESI LC-MS of the Pse5Ac7Ac (**1.13**) enzymatic reaction **a)** the negative control with no PseB or FlmA showing detection of only UDP-GlcNAc (**3.1**) starting material after 4 hrs, **b)** conversion to Pse5Ac7Ac (**1.13**) after 4 hrs with *C. jejuni* PseB and *A. cavaie* Flm B, Flm D and NeuB and **c)** demonstrating *A. cavaie* FlmA inability to turnover UDP-GlcNAc (**3.1**).

It has previously been shown that it is not necessary to add in the PseB co-factor (NADPH) when utilising the purified enzyme as it remains tightly bound to the enzyme during this process and is re-oxidised during the enzymatic reaction. It was also considered that the PseB co-factor, NADPH may not be bound as tightly in FlmA as it is in PseB and hence addition of it to the reaction mixture may be required for FlmA activity. Comparison of the *C. jejuni* PseB sequence with *H. pylori* PseB and *A. caviae* FlmA demonstrated that PseB from *C. jejuni* and *H. pylori* had a slightly higher level of identity than when comparing *C. jejuni* and *A. caviae* sequences, 61.1 % compared to 51.7 %. However general features were the same and in particular residues predicted to hydrogen bond the substrate and co-factor were conserved across all three sequences (**Figure 3.8**). However it was acknowledged that subtle differences between the enzymatic sequences can infer large alterations in the overall structure. Therefore it cannot be assumed that the NADPH co-factor is equivalently bound in FlmA as it is in PseB and hence addition of exogenous NADPH co-factor may be required for enzymatic activity.

```

H. pylori  MPNHQNMLDNQTILITGGTGSFGKCFVRKVLDTTNA-KKIIVYSRDELKQSEMAMEFN---DPRMR
          |   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
C. jejuni  -----MFNGKNILITGGTGSFGKTYTKVLLENYKP-NKIIYSRDELKQFEMSSIFN---SNCMR
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
A. caviae  -----MLNKNKTVLITGGTGSFGKQFIKTLERYPQVKRIVIFSRDELKQSELRLNYPQKDYPLR

H. pylori  FFIGDVRDLERLNYALEGVDICIHAAALKHVPIAEYNPLECIKTNIMGASNVINACLKNAISQVIA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
C. jejuni  YFIGDVRDKERLSVAMRDVDFVIHAAAMKHVPVAEYNPMECIKTNIHGAQNVIDACFENGVKKCIA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
A. caviae  FFIGDVRDRNRMVQACEGIDVIHAAAIKHQVDTAEYNPTECI RTNVDGAENVIHAALQCGVKEVVA

H. pylori  LSTDKAANPINLYGATKLCSDKL FVSANNFKGSSQTQFSVVRYG NVVGSRGSVVPFFKCLVQNKAS
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
C. jejuni  LSTDKACNPVNL YGATKLASDKL FVAANNIAGNKQTRFSVTRYG NVVGSRGSVVPFFKCLIAQGSK
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
A. caviae  LSTDKACAPINLYGATKLTSDKLFTAANNIKGSRNIRFSVVRYG NVMSGSGSVIPFFLKKRAEG--

H. pylori  EIPI TDIRMTRFWITLDEGVSFVLKSLKRMHGGEIFVPKIPSMKMTDLAKALAPNTPTKIIGIRPG
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
C. jejuni  ELPI TDIRMTRFWISLEDGVKLVLSNFERMHGGEIFIPKIPSMKITNLAHALAPNLSHKIIGIRAG
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
A. caviae  VLPI THEEMTRFNISLQDGVNMVYALEHHLGGEIFVPKIPSYRILDIATAISPECKTKVVGIRPG

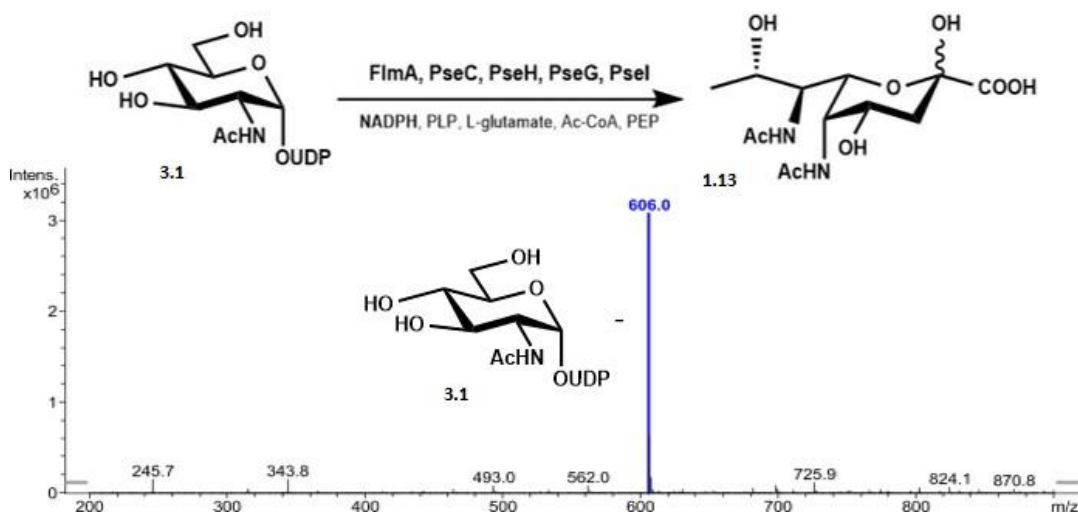
H. pylori  EKLHEVMIPKDESHLALAEFEDFFIIQPTISFQ-TPKDYTLTKLHEKGQKVAPDFEYSSHN NNQWLE
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
C. jejuni  EKLHEIMISSDDSHLT YEFENY AISPSIKLVDQESDFSINALGEKGQKV KDGFSYSSDN NPQWAS
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
A. caviae  EKLHEEMITD TDSLNTIDLGRYAILPSVSNHTEEDYIKH---HSAVKVPFGFKYNSGTNT EWET

H. pylori  PDDLKLL-----
          | |
C. jejuni  EKELLDIINHTEGF----
          | |
A. caviae  VESLRELKTHVDANFEV

```

**Figure 3.7** Sequence comparison of the *H. pylori*, *C. jejuni* and *A. caviae* PseB enzymes with residues highlighted that are involved in substrate (green) and co-factor (red) hydrogen bonding.

An additional reaction was set up incorporating NADPH (1.5 mM), the reaction mixture (**Table 3.5**), *A. caviae* FlmA, and *C. jejuni* PseC, PseH, PseG and PseI to test the requirement for exogenous co-factor. Production of Pse5Ac7Ac (**1.13**) was monitored using negative ESI LC-MS, however even after 24 hrs the only detectable sugar was the UDP-GlcNAc (**3.1**) starting material (**Figure 3.9**). As this result was identical to reactions without the addition of NADPH it was suggested that a lack of co-factor was not responsible for the lack of activity in this case. However there is still uncertainty over whether FlmA NADPH binding is analogous to *C. jejuni* and *H. pylori* PseB.



**Figure 3.8** Negative ESI LC-MS of the small scale enzymatic reaction employing *A. caviae* FlmA and *C. jejuni* PseC, PseH, PseG and PseI in the standard reaction mixture with the inclusion of NADPH (1.5 mM).

These results imply that *A. caviae* FlmA may not carry out the role it has been assigned; a UDP-GlcNAc 5-inverting, 4,6- dehydratase, and that the Pse5Ac7Ac (**1.13**) biosynthetic pathway in this bacteria may diverge further from that found in *C. jejuni* and *H. pylori*. For example, as with *B. thuringiensis*,<sup>182</sup> a partnership of enzymes may be required to convert UDP-GlcNAc (**3.1**) into the PseB product (**3.7**). Alternatively this FlmA enzyme could be less stable than the other *A. caviae* enzymes and the integrity of the protein may have been affected during transportation hence resulting in the lack of enzymatic activity. Further experiments into the potential activity of *A. caviae* FlmA were discontinued due to a lack of available protein. The comparison between *C.jejuni* PseB and *A. caviae* FlmA activity does however highlight how minor changes in the synthesis of Pse5Ac7Ac (**1.13**) can have catastrophic downstream effects.

### 3.4 Conclusions and future work

Promising developments along the pathway to a viable large scale synthesis of Pse5Ac7Ac (**1.13**) have been made. Induction conditions for the *C. jejuni* biosynthetic enzymes were explored and pure enzymes were combined in one pot to produce Pse5Ac7Ac (**1.13**). The activity of the proposed *A. cavaie* Pse5Ac7Ac (**1.13**) biosynthetic enzymes were investigated and FlmB, FlmD and NeuB were shown to behave analogously to their *C. jejuni* counterparts.

Optimal conditions for the production of Pse5Ac7Ac (**1.13**) biosynthetic enzymes were established; different conditions were required for the maximum induction of each enzyme. Under the new conditions one of the enzymes; PseC, appeared to aggregate and precipitate during purification resulting in inactive protein. Therefore alternate purification conditions were investigated, such as the addition of detergents, to promote enzymatic solubility however under all conditions precipitation could not be prevented. Induction conditions were re-evaluated and it was found that by lowering the temperature after induction, precipitation of PseC could be avoided. Unfortunately this was detrimental to the yield of PseC and it could be useful in the future to explore further methods for promoting solubility whilst retaining yield. For example, co-expression of PseC with one of the other biosynthetic enzymes could help to promote PseC solubility. Alternatively re-design of the plasmid could be beneficial with incorporation of either a fusion solubility tag such as MBP or fusion to another of the biosynthetic genes. However it is unknown if this would result in detrimental effects on the activity of the enzymes and hence further investigations would be required.

The purified Pse5Ac7Ac (**1.13**) displayed activity in a one-pot reaction (with their required co-factors) to allow for the standardised fully enzymatic conversion of UDP-GlcNAc (**3.1**) to Pse5Ac7Ac (**1.13**). Mimicry of these conditions with the proposed Pse5Ac7Ac (**1.13**) biosynthetic enzymes from *A. cavaie* confirmed three of the enzymes to be active and form part of the Pse5Ac7Ac (**1.13**) biosynthetic pathway. However the proposed *A. cavaie* UDP-5-inverting-4,6-dehydratase (FlmA) did not display turnover of UDP-GlcNAc (**3.1**), even upon addition of the NADPH co-factor to the reaction. Previous evidence presents FlmA as the first enzyme of the *A. cavaie* Pse5Ac7Ac (**1.13**) biosynthetic pathway and hence further investigations into its activity are required. Initial experiments should focus on gene expression and purification with regards to minimising potential de-stabilising effects that could have occurred during transit from Sheffield. Additionally crystallographic studies of this enzyme would aid in elucidating any potential structural differences compared to *H. pylori* PseB, which would help explain their difference in activity and confirm whether addition of exogenous co-factor is required.



## **Chapter 4 Optimising the PseB and PseC transformation of UDP-GlcNAc**

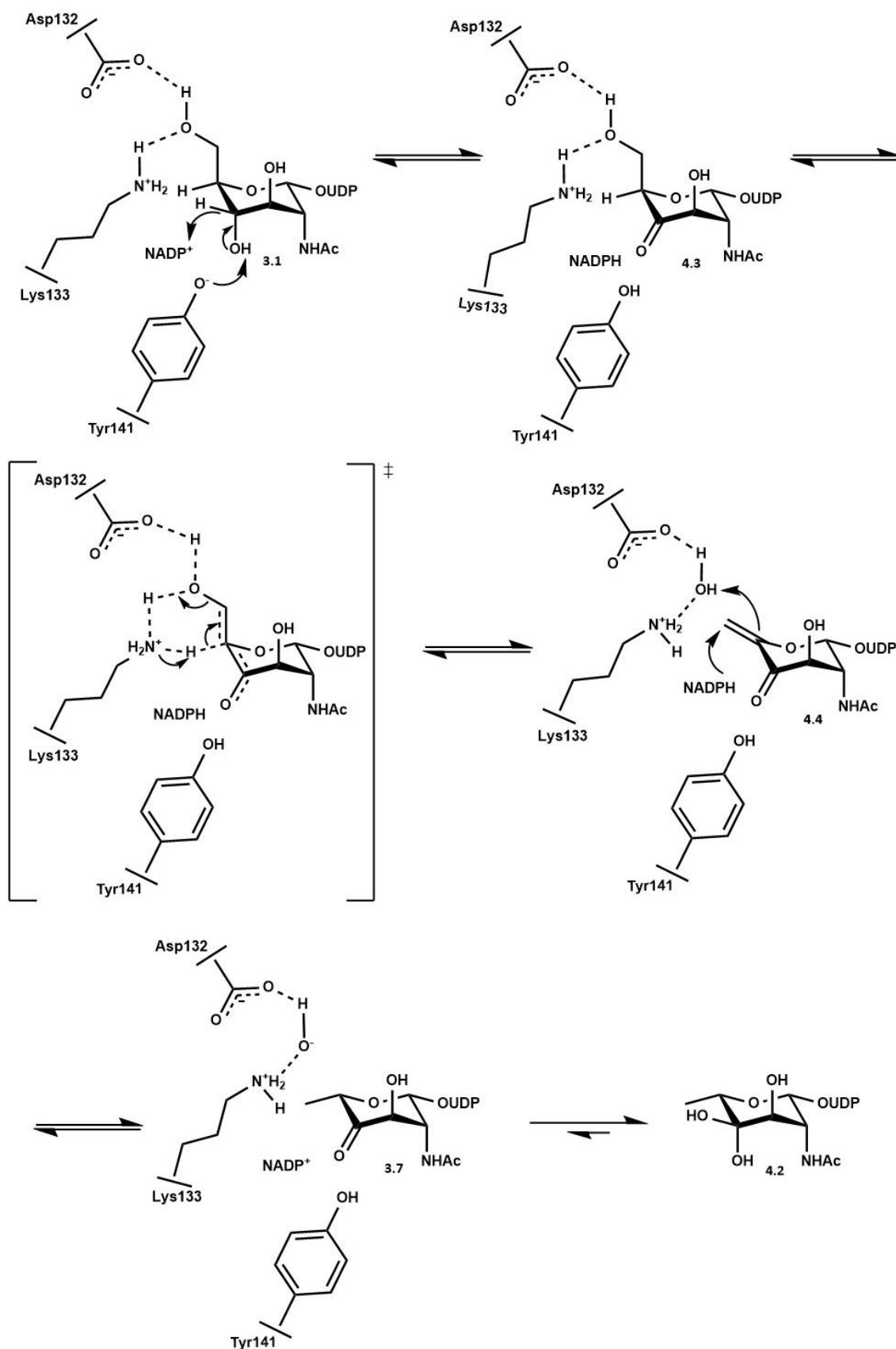
## 4.1 PseB and PseC activity in the Pse5Ac7Ac biosynthetic pathway

### 4.1.1 *Campylobacter jejuni* PseB; a UDP-GlcNAc 5-inverting-4,6-dehydratase

Deoxysugars are an important class of carbohydrates that are often used to glycosylate cell surface structures, for example *C. burnetii* virenose is an important component of the LPS which has been shown to be critical for evading host defences.<sup>237</sup> During the biosynthesis of deoxy sugars C6 deoxygenation is usually catalysed during the first step and the deoxy mechanisms have been heavily investigated.<sup>238</sup> A proposed C6 deoxygenation enzyme was originally annotated as FlaA1 in *H. pylori* and sequence similarities suggested a nucleotide activated sugar as the likely substrate and a NAD(P)H co-factor binding site.<sup>225</sup> UDP-GlcNAc (**3.1**) was initially proposed to be the substrate in activity assays and later confirmed by crystal structure complexes.<sup>230</sup>

However, initial characterisation of this enzyme was ambiguous, with confusion over its catalytic activity, co-factor requirement and products released. Although PseB (*FlaA1*, *Cj1293*) displays conserved domains with 4,6-dehydratases it also displays remarkable sequence similarities with UDP-GlcNAc (**3.1**) C4 epimerases<sup>225</sup> such as WbpP from *P. aeruginosa*.<sup>239</sup> The 4,6-dehydratase activity was observed in all studies and PseB was assigned as a member of the short chain dehydrogenase/reductase (SDR) family<sup>240</sup> and part of a sub-group that exhibits 4,6-dehydratase activity on nucleotide activated sugars to form deoxy-hexoses.<sup>241</sup> However there was some uncertainty regarding the functionality and configuration of the observed 4-keto product with reports of production of both of the C4 epimers; UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**)<sup>230</sup> and UDP-4-keto-6-deoxy-GlcNAc (**4.1**).<sup>225</sup> NMR data collected *in situ* confirmed the *C. jejuni* PseB initial product as UDP-4-keto-6-deoxy-L-IdoNAc (**3.13**),<sup>230</sup> observed as the hydrated form (**4.2**) in aqueous solution (**Scheme 4.1**).<sup>242</sup>

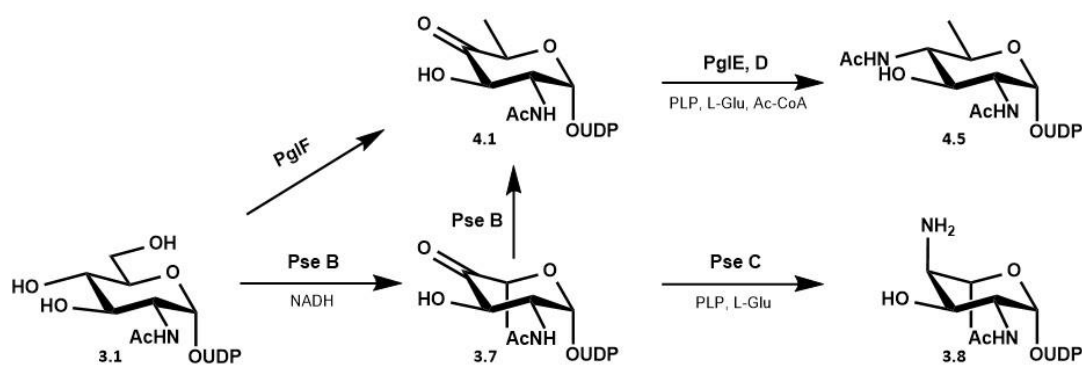
The crystal structure revealed the characteristic SDR (S/T)YK catalytic triad in close proximity to the GlcNAc moiety as well as an aspartate and lysine. It also confirmed the presence of the tightly bound NADP(H) co-factor that requires regeneration during the reaction. An accumulation of the biochemical research and crystal structure of the UDP-GlcNAc 5-inverting-4,6-dehydratase allowed for the mechanism to be proposed (**Scheme 4.1**).<sup>230</sup> This mechanism first employs the (S/T)YK triad and NADP<sup>+</sup> to oxidise the C4 hydroxyl producing the ketone intermediate (**4.3**) and reduced co-factor. This step is followed by Lys133 and Asp132 catalysed dehydration across the C5 (deprotonation) and C6 bonds (dehydration) forming the 4-keto-5,6-ene derivative (**4.4**). Finally the reduced NADPH co-factor delivers a hydride to C6 and simultaneously a stabilised water molecule donates a proton to C5 from the opposite face to form the inverted methyl group of the first PseB product (**3.7**).



**Scheme 4.1** PseB catalysed oxidation, dehydration and reduction of the substrate UDP-GlcNAc (**3.1**) to form the initial product UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**) which is in equilibrium with the hydrated form (**4.2**) in aqueous conditions.

#### 4.1.2 Additional PseB catalysed epimerisation

Further investigation of *C. jejuni* PseB biochemical mechanism revealed that in addition to acting as an inverting 4,6-dehydratase, PseB can also catalyse a further C5 epimerisation of the initial product, albeit at much lower rate (**Scheme 4.2**).<sup>242</sup> Incubation of the initial PseB product, UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**), with PseB resulted in conversion to the C5 epimer UDP-4-keto-6-deoxy-GlcNAc (**4.1**), with no epimerisation occurring in the control containing no PseB. UDP-4-keto-6-deoxy-GlcNAc (**4.1**) is the first intermediate (synthesised by PglF) in the biosynthesis of UDP-diNAcBac (**4.5**),<sup>241</sup> a sugar essential in *C. jejuni* N-glycosylation.<sup>243</sup> It may be that *C. jejuni* has evolved this secondary PseB catalysed epimerisation to ensure UDP-diNAcBac (**4.5**) production if PglF activity is compromised.

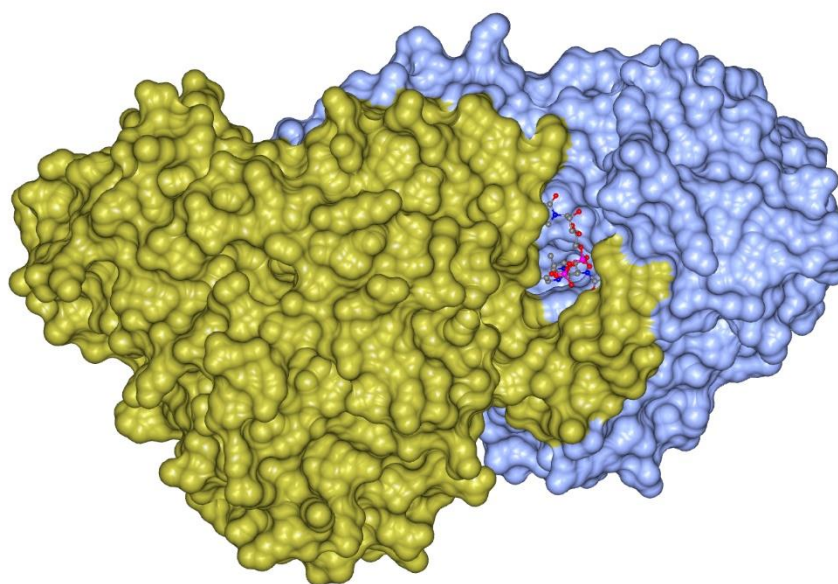


**Scheme 4.2** *C. jejuni* interlinking schemes from UDP-GlcNAc to two epimeric deoxysugar intermediates (**3.7**, **4.1**) in the biosynthetic pathways of two sugars important for protein glycosylation.

Full understanding of the complex mechanics of PseB turnover can only aid *in vitro* conversion of UDP-GlcNAc (**3.1**) to UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**), the first intermediate in the Pse5Ac7Ac (**1.13**) biosynthetic pathway. Prevention of the secondary PseB catalysed epimerisation is desired in order to promote activity of the Pse5Ac7Ac (**1.13**) pathway, however information is lacking on the catalytic turnover to this product (**4.1**).<sup>242</sup> It would be useful to carry out additional mutation studies to elucidate any modifications which decrease the activity of this secondary reaction without perturbing the initial PseB reaction. However for the purpose of use in the enzymatic synthesis of Pse5Ac7Ac (**1.13**) it was deemed more practical to investigate and optimise the PseB and PseC reaction design to increase the efficiency of turnover of UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**) to the desired product UDP-4-amino-4,6-dideoxy-β-L-AltNAc (**3.8**).

#### 4.1.3 PseC; a UDP-4-keto-6-deoxy-L-IdoNAc aminotransferase

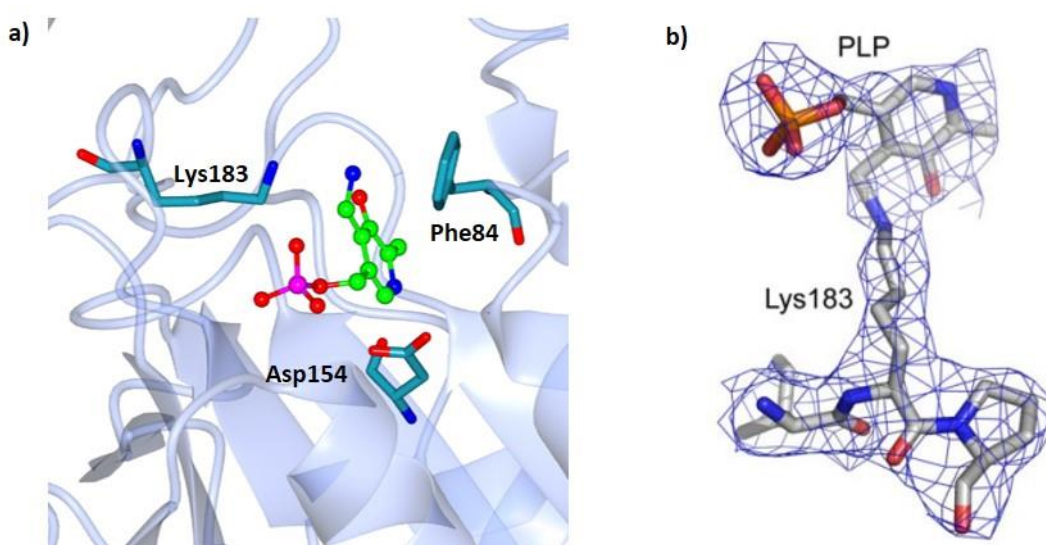
A number of biomacromolecules employ amino sugars within their structures and biosynthesis often requires an aminotransferase catalysed transamination. PLP-dependent aminotransferases are classified into four subgroups *via* comparison of amino acid sequences aligned based on the predicted secondary structure.<sup>244</sup> PseC enzymes are characterised as PLP-dependent Type 1 aminotransferases and have been shown to catalyse transfer of an amino group to the C4 of UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**).<sup>4</sup>



**Figure 4.1** The *H. pylori* PseC (PDB 2FNU)<sup>4</sup> homodimer (chain A ice blue, chain B gold), with the bound co-factor and substrate indicating position of the active site at the dimer interface.

*H. pylori* PseC exists as a homodimer in solution and in the crystal structure with both subunits contributing to each active site which is located near the dimer interface (**Figure 4.1**). Crystal structures of *H. pylori* PseC revealed the characteristic Type 1 aminotransferase PLP-binding site adjacent to the active site.<sup>245</sup> In particular the highly conserved aspartic acid and phenylalanine residue were identified as Asp154 and Phe84 respectively in *H. pylori* PseC. Asp154 is orientated to interact with the pyridinium nitrogen, enhancing the electron sink nature of the co-factor, and the Phe84 ring is orientated to  $\pi$ -stack with the co-factor pyridine ring, stabilising binding of the co-factor (**Figure 4.2a**).

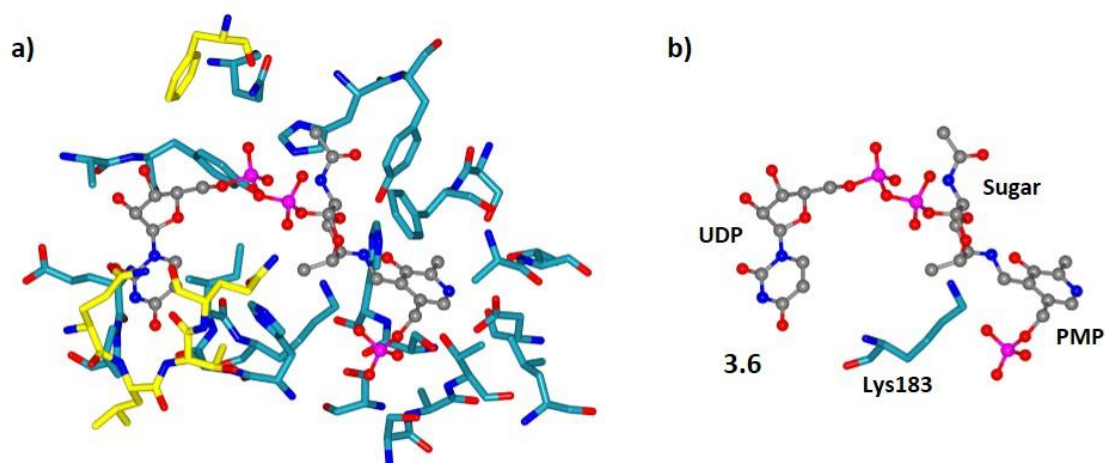
Crystal structures of *H. pylori* PseC in complex with PLP (**3.11**) showed it to form an internal aldimine with the Lys183 residue (**Figure 4.2b**) providing evidence for one of the intermediates in the proposed mechanism of the first half transamination reaction (**Scheme 3.3, Page 79**).<sup>4</sup> The identification of the natural amino donor of PLP-dependent aminotransferases has previously been ambiguous. However spectroscopic analysis of such enzymes consistently show turnover is achieved with L-glutamate (L-Glu) as the free amino donor and activity with this donor is more efficient when compared to other amino acids e.g. L-glutamine or L-alanine.<sup>246</sup>



**Figure 4.2** *H. pylori* PseC crystal studies (PDB 2FN6)<sup>4</sup> **a)** the co-factor binding site in complex with the PLP co-factor (**3.11**), highlighting Type 1 aminotransferase conserved residues and **b)** electron density of the PLP-enzyme internal aldimine intermediate (figure adapted from the original paper).<sup>4</sup>

Co-crystallisation of PseC, PLP (**3.11**) and the proposed product UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) also provided insight into the second half transamination reaction as the enzyme was seen to act in reverse. Electron density in the active site could be attributed to a PMP-sugar aldimine (**4.6**) (**Figure 4.3a**) suggesting that a direct aminotransfer from the PMP (**3.12**) to the keto-sugar occurs and confirming the proximity of the co-factor binding site and active site.

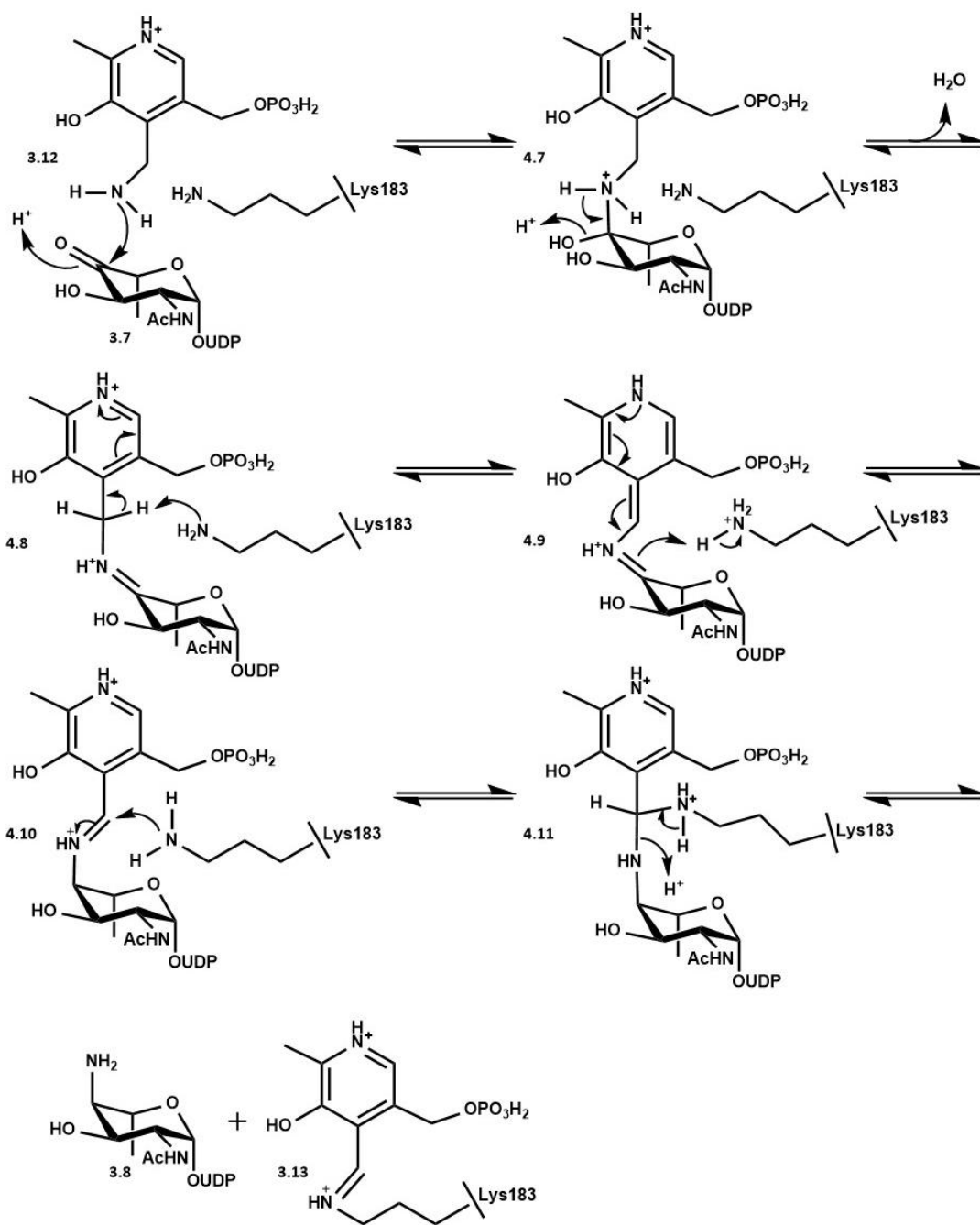
Lys183 was also identified as the catalytic residue for the second half reaction due to its orientation (**Figure 4.3**).<sup>4</sup> A drastically reduced aminotransferase activity was identified upon Lys183Arg mutation hence confirming the importance of Lys183.<sup>4</sup> These active site structural features eluded to the proposal of a mechanism mimicking that of ArnB (another UDP-4-keto aminotransferase)<sup>246</sup> whereby the PMP amine attacks the substrate C4 and Lys183 is utilised in a transaldimination reaction trigger release.



**Figure 4.3** The PMP-substrate external aldimine (ball and stick model) **a)** in complex with the surrounding *H. pylori* PseC residues (cyan chain A, yellow chain B) and **b)** with the catalytic Lys183 residue.

During the first half transamination (**Scheme 3.3**), initially an enzyme-PLP Schiff base forms (**3.14**), then a free amino donor (**3.15**) releases the enzyme from the internal aldimine (**3.14**) forming an external aldimine (**3.16**) *via* a transaldimination reaction. Finally hydrolysis occurs resulting in release of a glyoxylic acid (**3.18**) to produce free PMP (**3.13**) in the active site.<sup>231</sup> In *C. jejuni* PseC a screen of all twenty amino acids as the amino donor revealed that maximum conversion was achieved using 10 mM glutamate (20 molar equivalents of the substrate).<sup>247</sup>

The second amino transfer (**Scheme 4.3**) occurs *via* attack of the PMP amine to the sugar C4 keto producing water and a ketimine intermediate (**4.8**). Lys183 then acts as a base to abstract a labile proton (made so by the electron sink nature of the PMP pyridine ring) from the ketimine structure resulting in formation of a quinoid intermediate (**4.9**). Re-protonation at the sugar C4 position allows for formation of the sugar-co-factor aldimine (**4.10**) which is then released in a transaldimination reaction to give the bound co-factor enzyme complex (**3.13**) and the free aminated sugar (**3.8**).



**Scheme 4.3** The proposed PseC mechanism transferring an amino group from the PMP co-factor, generated *in situ*, to the keto-sugar (**3.8**) via formation of an external aldimine (**4.10**) that crystal structures have been shown as present in the active site during the reverse reaction.



#### 4.1.4 Campylobacter jejuni PseC

Although a structure for *C. jejuni* PseC has not been solved, detailed biochemical characterisation has been performed on this enzyme and similarities with the *H. pylori* sequence infers a similar fold. Sequence alignment reveals a 43.0 % identity between these enzymes and importantly, residues proximal to the co-factor and active site are almost always identical when comparing the *C. jejuni* PseC sequence with the *H. pylori* PseC sequence (**Figure 4.4**). For example the *H. pylori* catalytic residue Lys183 is aligned with the predicted *C. jejuni* catalytic residues; Lys181 surrounded by homologous residues.

```

H. pylori  MKEF AYS EPCLDKEDKKAVLEVLNSK QLTQ GKRSLLFEEALCEFLGVKHALVFN SATSAL LLTLY
          ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
C. jejuni  --MLT YSHQNIDQSDIDTLTKALKDEI LTGGKKVNEFEEALCEYMGVKHACVLN SATSAL HLAY

H. pylori  RNFSEFSADRNEI IITPTI SFVATANMLLESYGTPVFAGIKNDGNIDELALEKLINE ---RTKAI
          ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
C. jejuni  TA---LGVQEKIVLTTPLT FAATANAALMAGAKVEFIDIKNDGNIDEKKLEARLKLESENIGAI

H. pylori  VSVDYAGKSVEVESVQKLCKKHSLSFLS DS SHALGSEYQNKKVGGFALASVFSFHAIK PITTAE
          ||  ||  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
C. jejuni  SVVDFAGNSVEMDEI SNLTKKYNIPLID DASHALGALYKSEKVGKKADLSIFSFHPVK PITTFE

H. pylori  GGAVVTNDSELHEKMMLFRSHGMLKKE FFEGEVKSIGH NFRNLNEIQSALGLSQLKKAPFLMQKR
          ||  ||  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
C. jejuni  GGAVVSDNEELIDKIKLLRSHGIVKKRLWDS DMVELGY NYRLSDVACALGINQLKKLDHNLKRR

H. pylori  EEAALTYDRI FKDNPYFTPLHPLLKDKSSNHLYPILMHQKFFTCCKLILESLHKGILAQVHYK
          ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |  ||  |
C. jejuni  EEIANFYDKEFEKNPYFSTIKIKDYKSSRHLYPILLFPEFYCQKEELFESLLHAGIGVQVHYK

H. pylori  PIYQYQLYQQLFNTAPLKS AEDFYHAE ISLPCHANLNLESVQNI AHSVLKTFESFKIE ---
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
C. jejuni  PTYEF SFYKLLGEIKLQ NADNFYKAELS IPCHQEMNLKDAKFVKDTLFS ILEKVKKGYCG

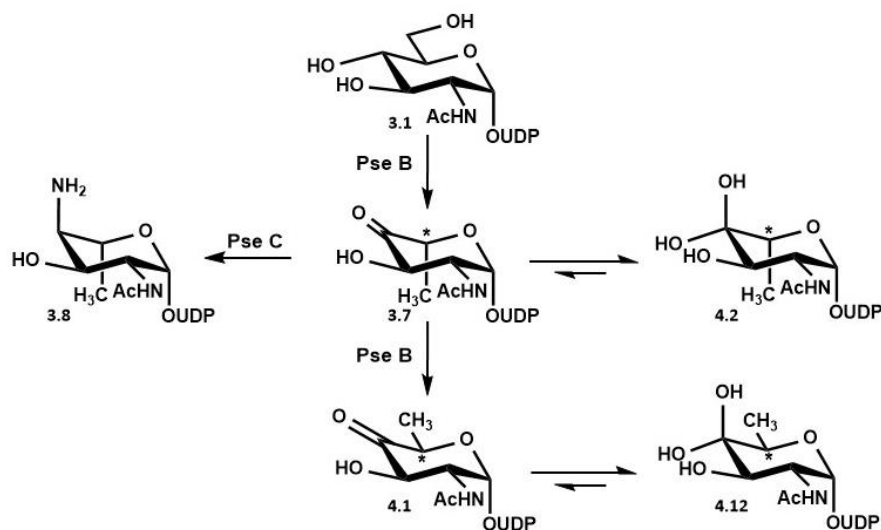
```

**Figure 4.4** Alignment comparison of the *H. pylori* and *C. jejuni* PseC sequences with *H. pylori* residues shown to be proximal to the co-factor and active site highlighted (green). The conserved *C. jejuni* residues in these sites are also highlighted (cyan), along with the conserved catalytic lysine (magenta).

NMR of the purified product from coupled *C. jejuni* PseB and PseC reactions confirmed the identity of the product as UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) and suggested its role as the second enzyme in the Pse5Ac7Ac (**1.13**) biosynthetic pathway.<sup>226</sup> Coupled reactions also revealed that aminotransferase activity occurred without the addition of exogenous PLP (**3.11**), suggesting that similarly to PseB, there is tight binding of the co-factor throughout purification.<sup>247</sup> However it was found that introduction of PLP (**3.11**) into the reaction mixture could increase activity suggesting that the intracellular level of PLP (**3.11**) was not sufficient for saturation of PseC with its co-factor.

## 4.2 Investigations into the relationship between PseB and PseC

### 4.2.1 Complications whilst monitoring the coupled PseB and PseC reaction



**Scheme 4.4** *C. jejuni* PseB and PseC full reaction scheme highlighting conversion to the PseB inverted by-product (4.1) and the non-enzymatic production of the hydrated PseB products (4.2, 4.12).

As previously discussed, when left in the presence of UDP-4-keto-6-deoxy-L-IdoNAc (3.7) *C. jejuni* PseB catalyses the epimerisation of this product at C5 (highlighted with an \*) (Scheme 4.4).<sup>242</sup> This is problematic during the enzymatic synthesis of Pse5Ac7Ac (1.13) as this epimer is not a substrate for PseC and does not form part of the Pse5Ac7Ac (1.13) biosynthetic pathway. Additionally activity analysis is complicated as the PseB ketone products (3.7, 4.1) exist in equilibrium with their hydrated forms (4.2, 4.12) which have the same molecular weight as the starting material UDP-GlcNAc (3.1) (Scheme 4.4). Therefore monitoring the reaction with negative ESI LC-MS becomes more complex as the  $[M-H]^-$  peak at 589 could be attributed to either the desired or un-desired PseB catalysed products (3.7, 4.1) and the  $[M-H]^-$  peak at 606 could be attributed to three different molecules (3.1, 4.2, 4.12), of which only one (4.2) is desired (Table 4.1).

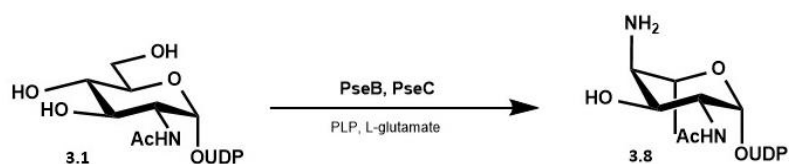
**Table 4.1;** LC-MS ESI  $[M-H]^-$  values for the components of the *C. jejuni* PseB, PseC enzymatic synthesis

Compound number	3.1	3.7	3.8	4.1	4.2	4.12
LC-MS ESI $[M-H]^-$ / Da	606	589	590	589	606	606

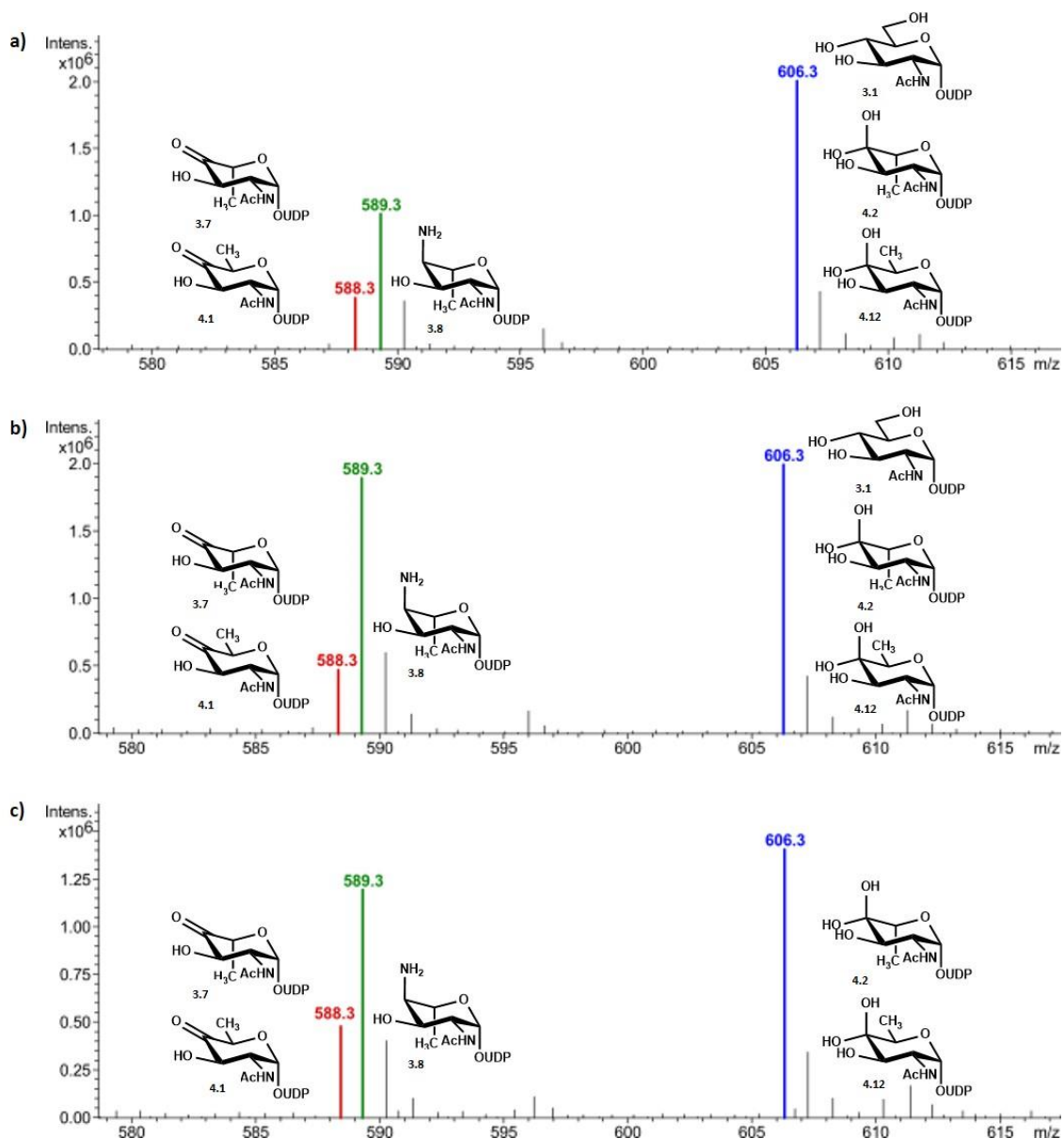
#### 4.2.2 Initial investigations into the coupled PseB and PseC reaction

Although pleasing that all of the *C. jejuni* enzymes were active and that enzymatic turnover to Pse5Ac7Ac (**1.13**) had been achieved on a small scale (< 1mg), it was proposed that further investigations into each step were required in order to optimise turnover and prevent potential issues upon increasing the scale of the reaction. It was decided to initially focus on the PseB and PseC enzymatic reactions with the intention to obtain conditions that would achieve optimal conversion to the PseC product (**3.8**). At the previously used concentrations of enzymes no biosynthetic intermediates were observed during the enzymatic synthesis of Pse5Ac7Ac (**1.13**) therefore lower concentrations of PseB and PseC were utilised in the following experiments to allow for tracking conversion (**Scheme 4.5**). A 50 mg UDP-GlcNAc (**3.1**) reaction mixture was set up by addition of PseB (25  $\mu$ M), PseC (23  $\mu$ M), PLP (1.5 mM) and L-Glu (10 mM) to UDP-GlcNAc (1 mM) in sodium phosphate buffer (50 mM, pH 7.4). The reaction was incubated (shaking 120 rpm, 37 °C) and conversion to the PseC product (**3.8**) was monitored with negative ESI LC-MS (**Figure 4.5**). Although the use of LC-MS is a useful comparative tool and allows for relative conversions to be assigned, within this report it is not used as a quantitative measure. Different molecules have different ionisation potentials and hence conversions have only been assigned in order to relatively compare the effect of modifying the conditions.

Upon addition of PseB and PseC the intensity of the  $[M-H]^-$  606 Da peak decreased which was attributed to PseB catalysed turnover of UDP-GlcNAc (**3.1**), and there was an increase in intensity of the peak corresponding to the PseB product (**3.7**) ( $[M-H]^-$  588 Da) and PseC product (**3.8**) ( $[M-H]^-$  589 Da) from 10 mins to 45 mins (**Figure 4.5a, b**). For example, the relative intensity of the PseC product (**3.7**) peak was 29 % after 10 minutes and increased to 42 % after 45 minutes (relative intensity calculated as a percentage of the total intensity of all peaks corresponding to sugar intermediates). It has previously been shown that hydration of the PseB product occurs in just a few minutes (producing a structure with the same molecular weight as UDP-GlcNAc **3.1**) and hence upon the time scale of LC-MS analysis it can be assumed that conversion to the PseB hydrated product had begun prior to analysis. Therefore turnover to the PseB product could not be quantified during the reaction as the PseB hydrated product and the UDP-GlcNAc starting material (**3.1**) will both be contributing to the intensity of the  $[M-H]^-$  606 Da peak.



**Scheme 4.5** The coupled *C. jejuni* PseB and PseC reaction converting UDP-GlcNAc (**3.1**) to the second Pse5Ac7Ac (**1.13**) biosynthetic intermediate UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**).



**Figure 4.5** *C. jejuni* PseB and PseC reaction progression **a)** after incubating for 10 mins displaying conversion to the PseC product (**3.8**), **b)** after incubating for 45 mins displaying increased conversion to the PseC product (**3.8**), and **c)** after incubating for 6 hrs displaying no further increase in conversion to the PseC product (**3.8**).

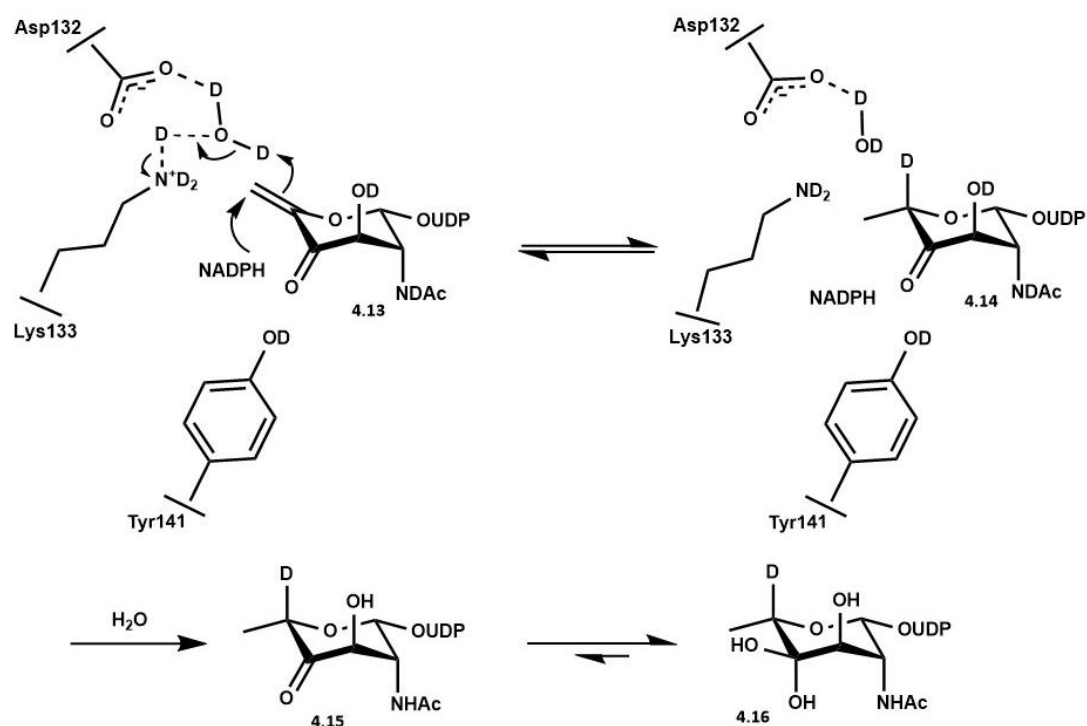
It was hoped that the inclusion of the PseC enzyme converting the ketone-PseB product (**3.7**) to the aminated product (**3.8**) would promote both the PseB catalysed depletion of UDP-GlcNAc (**3.1**) and the un-catalysed dehydration of the PseB hydrated product (**4.2**) and hence a decrease in intensity of the  $[M-H]^-$  606 Da peak would be observed as the reaction progressed. However the reaction did not progress as expected, with similar relative intensities of all peaks observed after 1 hr and after 6 hrs of incubation, with a significant  $[M-H]^-$  606 peak remaining (45 %) (**Figure 4.5c**). This was somewhat unprecedented considering the previously achieved conversion during the enzymatic synthesis of Pse5Ac7Ac (**1.13**). However as the reaction was on a larger scale and progressing more slowly it was proposed that the slower PseB catalysed secondary epimerisation had begun to compete with the PseC reaction and converted some of the UDP-4-keto-6-deoxy-L-IdoNAc (**3.8**) into the PseB epimeric by-product (**4.1**) which in the hydrated form (**4.12**) also has a  $[M-H]^-$  of 606.

Even though it was observed that turnover to the desired PseC product sugar (**3.8**) had halted under these conditions it was impossible to distinguish which sugar structures were still present in the solution. For example, the intensity of the  $[M-H]^-$  606 peak could be attributed to UDP-GlcNAc (**3.1**) starting material or hydrated PseB by-product (**4.2**, **4.12**) and the  $[M-H]^-$  588 peak could be the desired PseB product (**3.7**) or its C5 epimeric form (**4.1**). However the lack of further conversion was indicative that an equilibrium had been reached and the peaks at  $[M-H]^-$  588 Da and  $[M-H]^-$  606 Da could be attributed to the presence of the PseB by-products (**4.1** and **4.12** respectively).

These initial results suggest that production of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) on a large scale could be more challenging than originally assumed. Although a modest conversion was achieved  $\sim$  40 %, when comparing the relative intensity of the PseC product peak at  $[M-H]^-$  589 Da, to the combined relative intensities of the other sugar substrates ( $[M-H]^-$  589 Da and  $[M-H]^-$  606 Da) after 6 hrs, it was proposed that optimisation of the enzyme concentrations could promote higher levels of conversion. In particular it was apparent that a significant proportion of the sugar compounds had undergone conversion to the PseB by-product (**4.1**), affecting the overall conversion to the desired PseC product (**3.8**) hence it was desired to attain conditions to prevent this.

### 4.2.3 The PseB and PseC coupled reaction in deuterium oxide

To prevent formation of the proposed by-product it was necessary to remove UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**) from the solution as soon as it was formed. Therefore a series of experiments were designed using varying ratios of PseB:PseC in order to observe if an increase in PseC concentration could promote conversion to the PseC product (**3.8**) and out-compete the PseB secondary reaction (**Scheme 4.4**). These experiments were carried out in deuterated buffer to allow for the PseB hydrated forms to be distinguishable from UDP-GlcNAc (**3.1**) by negative ESI LC-MS as a solvent molecule is employed in the PseB mechanism.

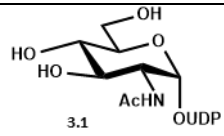
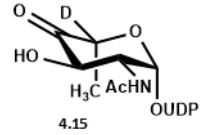
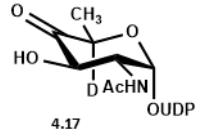
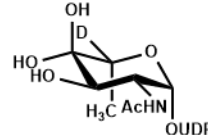
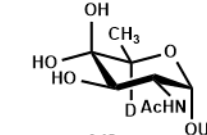
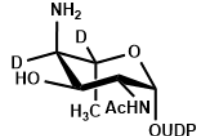


**Scheme 4.6** Partial PseB mechanism highlighting the D<sub>2</sub>O solvent molecule utilised during the mechanism and the corresponding product detected by negative ESI LC-MS whereby all of the labile deuterium have undergone solvent exchange with the H<sub>2</sub>O/MeCN mobile phase.

Previous experiments, in D<sub>2</sub>O, tracked by NMR identified formation of a C5-D bond during the PseB catalysed reaction, which is formed during the reduction of the alkene, hence formation of a stable C-D bond is observed (**Scheme 4.6**). Solvent exchange will also occur between labile bonds (such as hydroxyls) in these sugar molecules in D<sub>2</sub>O, however this potential complication is minimised during LC-MS analysis as they are subjected to an excess of H<sub>2</sub>O in the mobile phase and all labile bonds undergo deuterium-proton exchange. Therefore carrying out the reaction in

D<sub>2</sub>O allows for the production of PseB products (**4.15**, **4.16**, **4.17**, **4.18**) with a [M-H]<sup>-</sup> of +1 compared to their non-deuterated analogues resulting in the PseB hydrated forms (**4.14**, **4.15**) being distinguishable from UDP-GlcNAc (**3.1**) (Table 4.2).

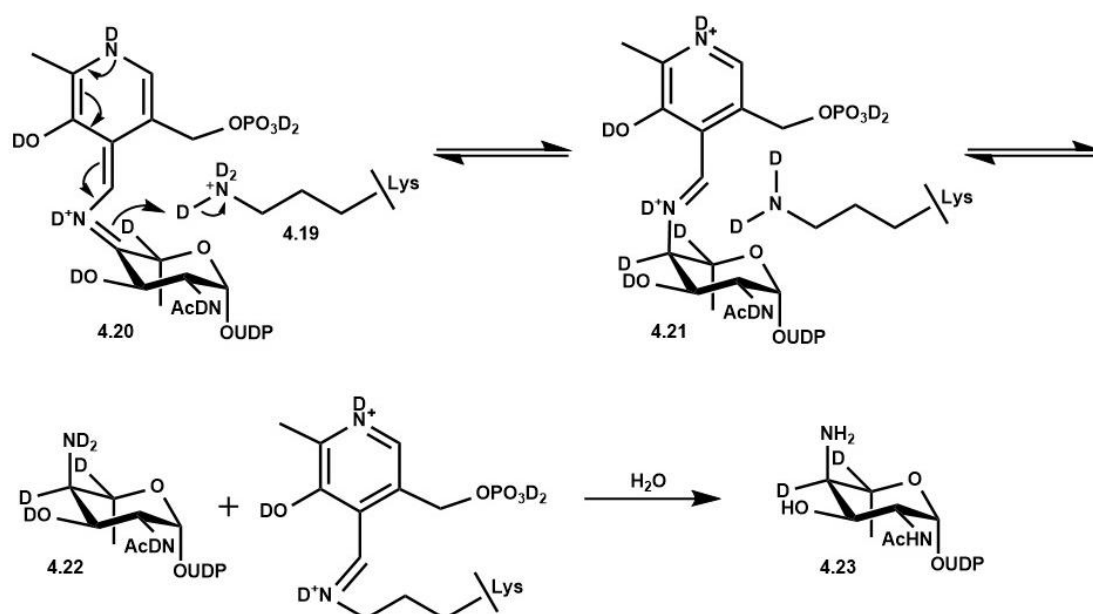
**Table 4.2** Negative ESI [M-H]<sup>-</sup> values for sugar molecules present in the deuterated PseB, PseC reaction

Sugar molecule	LC-MS ESI [M-H] <sup>-</sup> in D <sub>2</sub> O
 3.1	606
 4.15	589
 4.17	
 4.16	607
 4.18	
 4.23	591

In a

deuterated solvent all of the labile N-H and O-H bonds will undergo solvent exchange and be deuterated so in negative ESI mass spectrometry a peak at a [M-H]<sup>-</sup> 600 would be expected for the deuterated PseC product (**4.18**). This peak is not observed using LC-MS analysis however, as the solvent system has an excess of H<sub>2</sub>O and therefore during analysis solvent exchange occurs with all labile deuterated bonds. Based on the proposed PseC mechanism, it was expected that a second stable deuterium would be incorporated into the sugar structure. Predictions of the position of the deuterium incorporation can be made based on the proposed PseC mechanism as during the synthesis of UDP-4-amino-4,6-dideoxy-β-L-AltNAc (**3.8**) there is only one non-labile bond made; (Scheme 3.3 and Scheme 4.3).

In deuterated solvent, it can be assumed that all labile protons will be solvent exchanged, for example the catalytic Lys181 amine will be deuterated (**4.19**) (**Scheme 4.7**). During the rearrangement from the quionoid intermediate (**4.20**) to the PMP-substrate aldimine (**4.21**), the active site lysine donates a deuterium to the sugar C4 position. Therefore a deuterium from the labile N-D bond is incorporated into the sugar ring (**4.22**) forming a stable C-D bond that is retained in the sugar ring during LC-MS analysis when all of the labile deuterium undergo solvent exchange (**4.23**) (**Scheme 4.7**). Negative ESI LC-MS analysis confirmed the incorporation of a non-labile deuterium during the PseC reaction as well as the one incorporated during the PseB reaction as the product peak was observed at an  $[M-H]^-$  of +2 compared to when the PseB, PseC coupled reaction is carried out in  $H_2O$  ( $[M-H]^-$  589 in  $H_2O$ ,  $[M-H]^-$  591 in  $D_2O$ ) (**Table 4.2**).



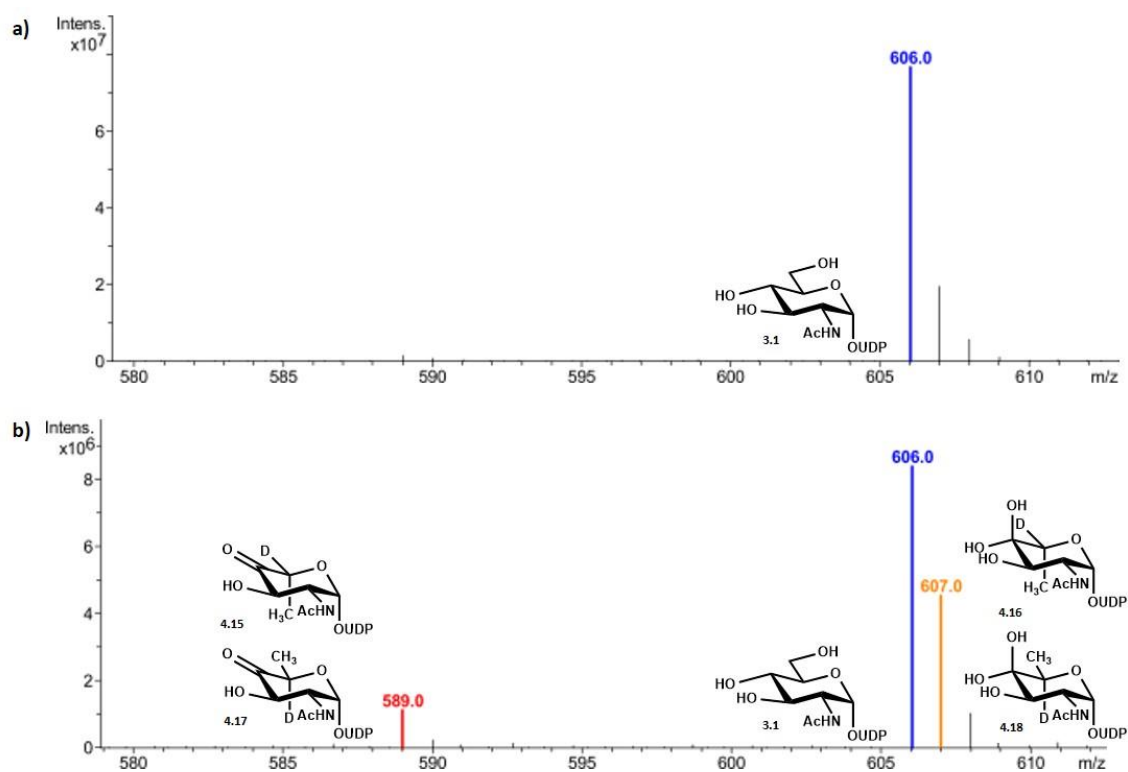
**Scheme 4.7** Partial PseC mechanism in deuterium proposing formation of a stable C4 carbon-deuterium bond and the resulting sugar identified by negative ESI LC-MS.

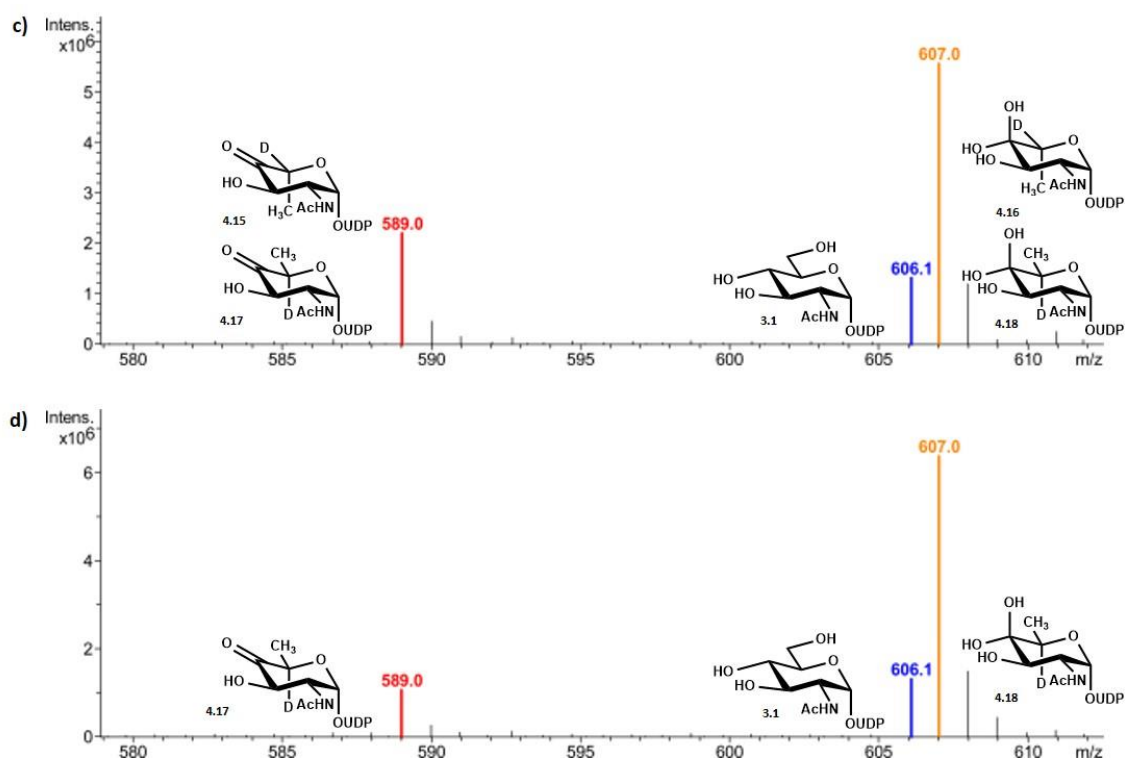


#### 4.2.4 Preventing formation of the PseB by-product

Upon incubation (shaking 120 rpm, 37 °C) of Pse B (25 μM) and UDP-GlcNAc (1 mM) in a deuterated sodium phosphate buffer (50 mM, pH 7.4), the LC-MS peak corresponding to UDP-GlcNAc (**3.1**) drastically diminished over 2 hours (**Figure 4.6b,c**). In its place arose two new peaks that were assigned to the singly deuterated versions of the keto tautomer (**4.15, 4.17**) and hydrated (**4.16, 4.18**) Pse B products. No such peaks were observed in the control (the reaction mixture without PseB) with only the  $[M-H]^-$  606 peak observed, indicative of no retaining solvent exchange occurring with the starting material, even after 12 hrs (**Figure 4.6a**).

Negative ESI LC-MS analysis showed that after 30 minutes a considerable amount of UDP-GlcNAc (**3.1**) remained in the reaction mixture containing PseB (64 % of the combined relative intensities of all of the sugar moieties). After two hours this had decreased to only a minor peak (13 %) showing efficient activity of the enzyme under these conditions (**Figure 4.6b,c**). During analysis it was observed that there was a higher intensity for the peak corresponding to the hydrated PseB product ( $[M-H]^-$  607) compared to the ketone PseB product peak ( $[M-H]^-$  589) suggesting that the position of equilibrium lay to the hydrated form and that conversion was fast. For example after 30 minutes, the ratio of deuterated ketone-PseB product (**4.15, 4.17**) to hydrated-PseB product (**4.17, 4.18**) was 1 : 4 (**Figure 4.6b**).





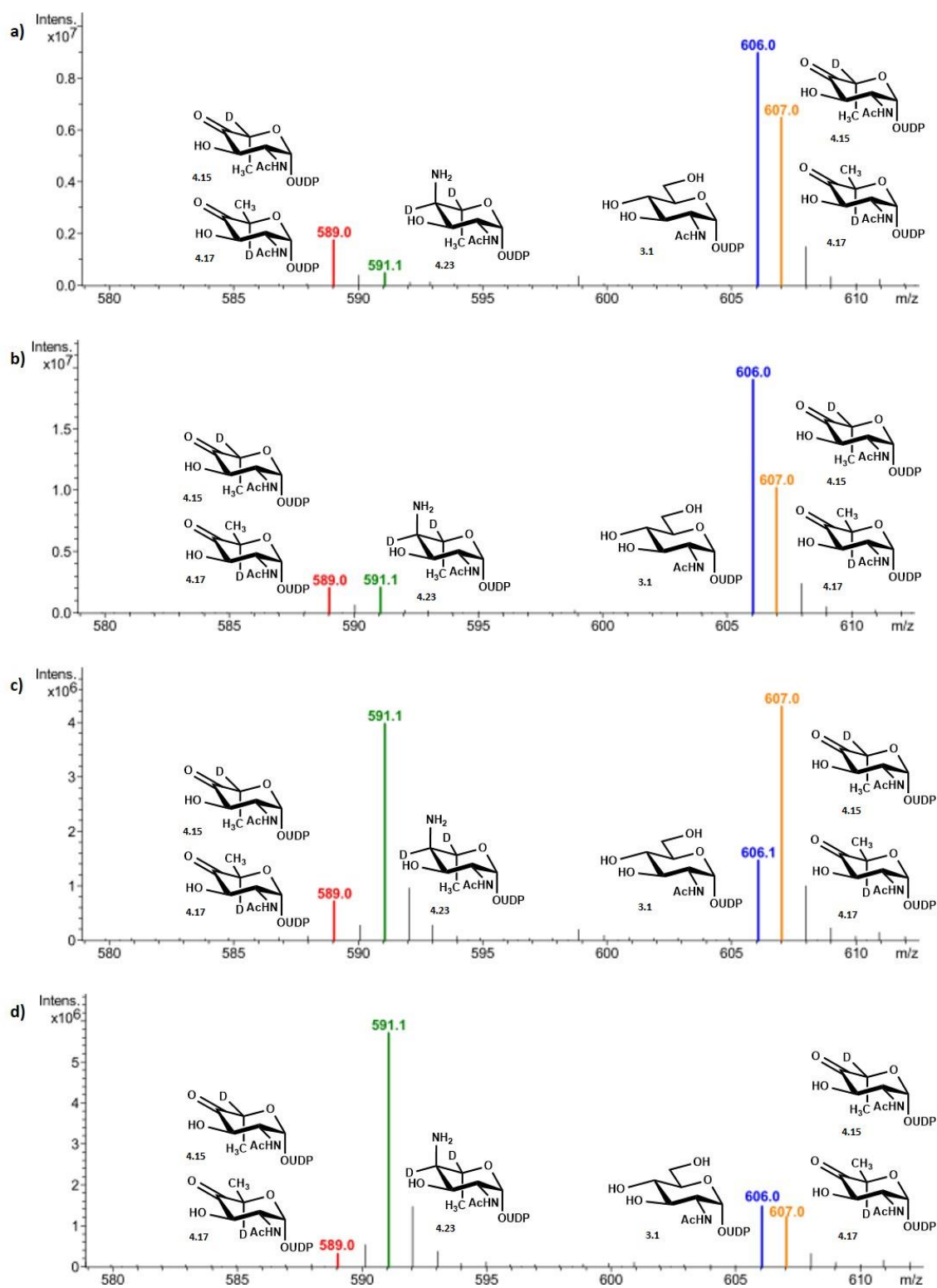
**Figure 4.6** Negative ESI analysis of formation of the *C. jejuni* PseB products in deuterated solvent; **a**) control reaction of UDP-GlcNAc (**3.1**) after 12 hrs incubation displaying no conversion **b**) 30 mins after incubating UDP-GlcNAc (**3.1**) with PseB, **c**) 120 mins after incubating UDP-GlcNAc (**3.1**) with PseB and **d**) 120 mins after addition of PseC and co-factors to the reaction mixture **c**).

To investigate the C5 stereochemistry of the reaction mixture after two hours i.e whether it was still an intermediate of the Pse5Ac7Ac (**1.13**) biosynthetic pathway, PseC (and its co-factors) in D<sub>2</sub>O were added to the mixture. Negative ESI LC-MS analysis after two hours displayed no conversion to the deuterated PseC product (**4.23**) indicative that PseB had already catalysed all of its original product (**4.15**) into the epimeric by-product (**4.17**) (**Figure 4.6d**). Comparison of the ketone-PseB product (**4.15**, **4.17**) with the hydrated product (**4.16**, **4.18**) after this four hours of incubation showed an increased ratio of 1 : 5.2. Thus the previously observed remaining [M-H]<sup>-</sup> 606 peak in non-deuterated PseB reactions was attributed to the formation of these hydrated products (**4.16**, **4.18**) and the apparent low conversion rates to the PseB product (**3.7**) accounted for.

Unlike the sequential addition of PseC to the PseB reaction mixture, when PseB and PseC were added to the reaction mixture simultaneously, negative ESI LC-MS analysis confirmed the presence of deuterated PseB products (**4.15-4.18**) in addition to a peak corresponding to the doubly deuterated PseC product ( $[M-H]^-$  591) (**4.23**) (**Figure 4.7**). Repeats of the  $D_2O$  experiments allowed for optimisation of enzymatic concentrations in order to maximise conversion to the Pse C product (**4.23**). Reaction mixtures in deuterated sodium phosphate buffer (50 mM, pH 7.4) containing UDP-GlcNAc (1 mM), PLP (1.5 mM), L-glu (10 mM) PseB (25  $\mu$ M) and PseC (25  $\mu$ M or 125  $\mu$ M) were incubated (shaking 120 rpm, 37 °C). Negative ESI LC-MS after ten minutes using [PseB : PseC] ratios of either 1:1 or 1:5 revealed little difference in the consumption of UDP-GlcNAc (**3.1**); with the  $[M-H]^-$  606 peak 56.7 % compared to 56.9 % respectively (of the total peak intensities of the combine sugar products) (**Figure 4.7a,b**). This was also the case after two hours, whereby in both reactions the peak corresponding to UDP-GlcNAc (**3.1**) had decreased similarly to 14 % and 16 %, suggesting that the concentration of PseC did not affect the overall rate of consumption of UDP-GlcNAc (**3.1**) in these conditions (**Figure 4.7c,d**).

However, the concentration of PseC was found to have an effect on the production of PseC product (**4.23**), with the reaction mixture containing an excess of PseC, resulting in a higher conversion to the PseC product (**4.23**) and lower peak intensities for the PseB products (**4.15-4.18**). Even after ten minutes of incubation the reaction mixtures showed some deviation with the reaction mixture with the higher PseC concentration having a 6 % conversion to the PseC product whereby the reaction with 25  $\mu$ M PseC only had a 3 % conversion (**Figure 4.7a,b**).

After incubating the reactions for two hours, a much larger difference in relative peak intensity for the conversion of PseB products to PseC product were observed between the two reactions containing different concentrations of PseC (**Figure 4.7c, d**). At a [PseB : PseC] ratio of 1:1, there was only a 38 % conversion to the PseC product (**4.23**) after two hours, with 48 % of the sugar species remaining as PseB products (**4.15-3.18**) (**Figure 4.7c**). The reaction mixture containing five fold higher concentration of PseC showed an increased conversion of 66 % to the PseC product (**4.23**) with only 20 % of the sugar species remaining as PseB products (**Figure 4.7d**). This ratio did not change after incubating the reaction mixture for 24 hours suggesting that the PseB, PseC equilibrium had been reached. This improvement upon conversion to the PseC product (**3.8**) was deemed sufficient for preventing significant formation of the PseB epimeric by-product during enzymatic synthesis of Pse5Ac7Ac (**1.13**) whereby the formation of the PseC product would be even further promoted by its subsequent conversion to the other biosynthetic intermediates.



**Figure 4.7** Negative ESI analysis of formation of the *C. jejuni* PseB (4.13-4.16) and PseC products (4.17) in deuterated solvent whilst investigating the effect of increasing the PseC concentration; **a)** concentration ratio of [PseB : PseC] of 1 : 1 after 10 mins incubation, **b)** concentration ratio of [PseB : PseC] of 1 : 5 after 10 mins incubation, **c)** concentration ratio of [PseB : PseC] of 1 : 1 after 2 hrs incubation, **d)** concentration ratio of [PseB : PseC] of 1 : 5 after 2 hrs incubation.

### 4.3 Conclusions and future work

Although the previously optimised production of the biosynthetic enzymes demonstrated *in vitro* turnover from UDP-GlcNAc (**3.12**) to Pse5Ac7Ac (**1.13**) in a one pot reaction on a small scale, restrictions to using the biosynthetic pathway were highlighted. Before large scale synthesis of Pse5Ac7Ac (**1.13**) could be attempted, it was imperative that these were addressed and resolved.

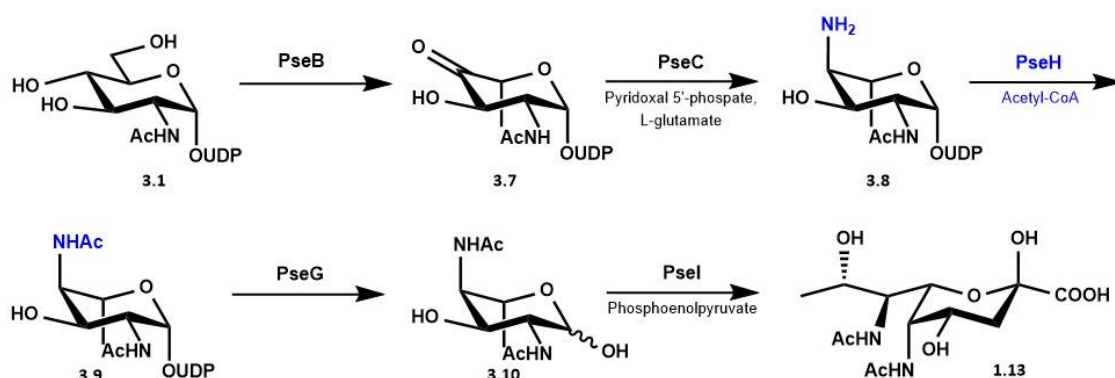
These experiments focussed on exploring the relationship between the PseB and PseC enzymatic reactions, and the potential deviations that can occur from the Pse5Ac7Ac (**1.13**) biosynthetic pathway. Notably the PseB secondary epimerisation posed an issue for the enzymatic synthesis of Pse5Ac7Ac (**1.13**) as the resulting product (**4.1**) does not make up part of the Pse5Ac7Ac (**1.13**) biosynthetic pathway, but the UDP-diNAcBac (**4.5**) pathway instead. Although the mechanism to the desired PseB product has been defined, little is known about the secondary reaction and hence PseB could not be engineered to prevent formation of this product. Therefore it was proposed that the reaction would have to be optimised to remove UDP-4-keto-6-deoxy-L-IdoNAc (**3.7**) as soon as it was formed to prevent formation of the epimeric by-product. This was achieved by monitoring the PseB and PseC reactions in D<sub>2</sub>O and modifying the ratio of the concentrations of the enzymes. It was found that by increasing the concentration of PseC to five times that of PseB, conversion to the PseC product could be considerably increased after two hours compared to using equal concentrations of PseB and PseC.

Further investigations into preventing the formation of the PseB by-product could focus on determining the residues implicated in the PseB mechanism. Mutation studies have highlighted active site residues essential for the catalysis of the initial inversion and dehydration but were found not to perturb the secondary epimerisation step.<sup>242</sup> Therefore this suggests that different residues could be involved in the different mechanisms and hence further studies to identify the secondary reaction catalytic residues would aid in the design of rational mutations that could reduce the rate of the second epimerisation reaction without perturbing the initial reaction. Furthermore, investigations of the inter-linking of sugar pathways in other Pse5Ac7Ac (**1.13**) biosynthesising species could allow for identification of an UDP-GlcNAc (**3.1**) 5-inverting-4,6-dehydratase that does not catalyse the secondary epimerisation to form a Pse5Ac7Ac (**1.13**) biosynthetic by-product.

## **Chapter 5 Identification of alternative acetyl-transfer strategies**

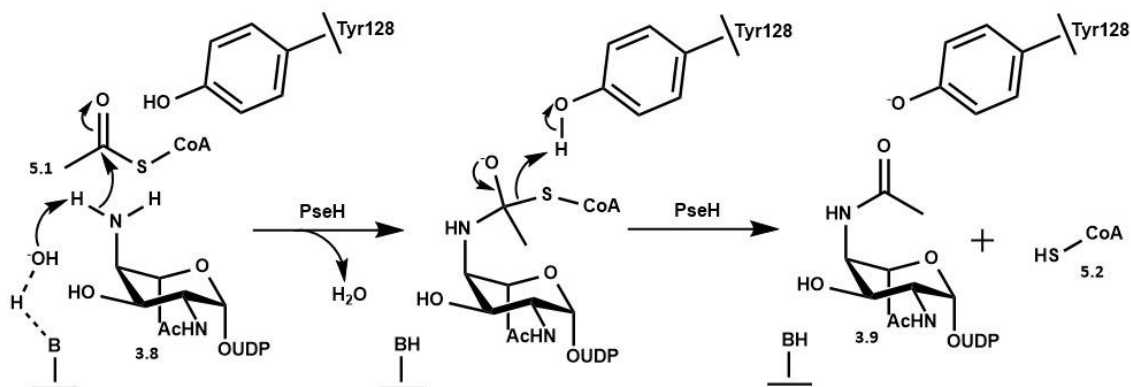
## 5.1 Utilisation of acetyltransferases as synthetic tools

### 5.1.1 The PseH catalysed acetyl transfer step



**Scheme 5.1** *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic pathway highlighting the acetyl transfer step.

The third step in the Pse5Ac7Ac (**1.13**) biosynthetic pathway is the transfer of an acetyl group to C4 of UDP-4-amino-4,6-dideoxy- $\beta$ -l-AltNAc (**3.8**) (**Scheme 5.1**).<sup>248</sup> Enzymatically this step is catalysed by PseH, an aminoglycoside *N*-acetyltransferase from the GNAT superfamily, and as such, utilises acetyl-CoA (**5.1**) as a co-factor.<sup>249</sup> It has been proposed that this family of acetyltransferase enzymes drives nucleophilic attack by orientating a basic residue so that it can deprotonate the amino group and employing an acidic residue to stabilise the de-acetylated thiolate anion.<sup>250</sup> Crystal structures of *C. jejuni* PseH identified Tyr128 as the general acid to stabilise the proposed thiolate ion formed following acetyl transfer to UDP-4-amino-4,6-dideoxy- $\beta$ -l-AltNAc (**3.8**) however an acidic residue close enough to deprotonate the C4 amino group was not observed. Docking of the substrate into the *H. pylori* PseH crystal structures showed a well-ordered water molecule hydrogen bonded to a serine positioned proximal to the C4 amino group that could mediate deprotonation of the substrate (**Scheme 5.2**).<sup>251</sup>



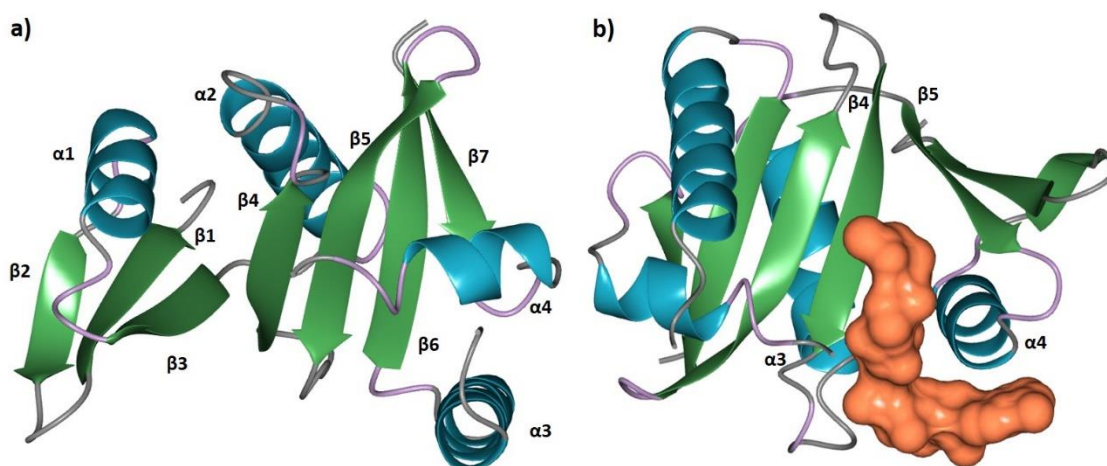
**Scheme 5.2** Acetyltransfer mechanism for synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -l-AltNAc (**3.9**) showing the *C. jejuni* PseH catalytic acid; Tyr128, and the proposed water mediated amine deprotonation.

### 5.1.2 Acetyl-CoA binding in GNAT enzymes

Acylated amines are common in many biological, industrially relevant molecules and pharmaceuticals,<sup>252</sup> and are generally introduced *via* acetyl transfer onto an amine.<sup>253</sup> Chemical attachment of this functionality onto complex structures can be selectively challenging, however. Enzyme catalysed amine acetylation is commonly observed in nature with the majority employing acetyl-CoA (**5.1**) as an acetyl donor with a vast range of acceptor molecules.<sup>254</sup>

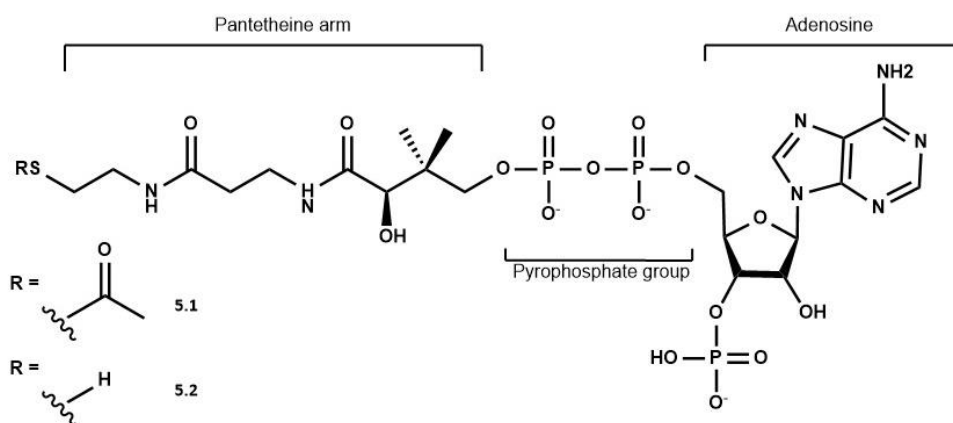
The GCN5-related *N*-acetyltransferase (GNAT) superfamily consists of more than 100,000 members found in all kingdoms of life.<sup>255</sup> It encompasses a wide range of acceptor molecules including bacterial aminoglycoside transferases,<sup>256</sup> histone (and other proteins) *N*-acetyltransferases<sup>257</sup> and arylamine *N*-acetyltransferases.<sup>258</sup> This diversity of function is reflected in the vast array of cellular processes that these enzymes are involved in; for example drug resistance,<sup>259</sup> biosynthesis,<sup>232</sup> metabolism,<sup>260</sup> and DNA replication.<sup>261</sup>

Although the GNAT family has low sequence homology there is conservation of the core fold in all structures to date.<sup>250</sup> Structural characterisation of over twenty GNAT enzymes has revealed conserved secondary elements consisting largely of antiparallel  $\beta$  sheets (six or seven strands) connected by loops or one of the four conserved  $\alpha$  helices (**Figure 5.1a**).<sup>249</sup> Another notable structural feature is the “ $\beta$ -4 bulge” forcing the  $\beta$ -4 and  $\beta$ -5 strands apart at one end to form a V-shape cleft (**Figure 5.1b**).<sup>250</sup> Crystal structures in complex with the (acetyl)-CoA co-factor (**5.1/5.2**) show this cleft to be the co-factor binding site in all structures. Acetyl-CoA (**5.1**) is bound shaped like the letter ‘L’ and interactions with the enzyme occur with residues in the proximal  $\alpha$ 3,  $\alpha$ 4,  $\beta$ 4 and  $\beta$ 5 secondary structures (**Figure 5.1b**).<sup>249</sup>



**Figure 5.1** Crystal structure of the *C. jejuni* PseH acetyltransferase (PDB 4XPL)<sup>232</sup> highlighting **a**) the canonical secondary structure of GNAT enzymes and **b**) the co-factor V-shape cleft binding site between strands  $\beta$ 4 and  $\beta$ 5 with the bound acetyl-CoA (**5.1**) co-factor (orange globular structure).

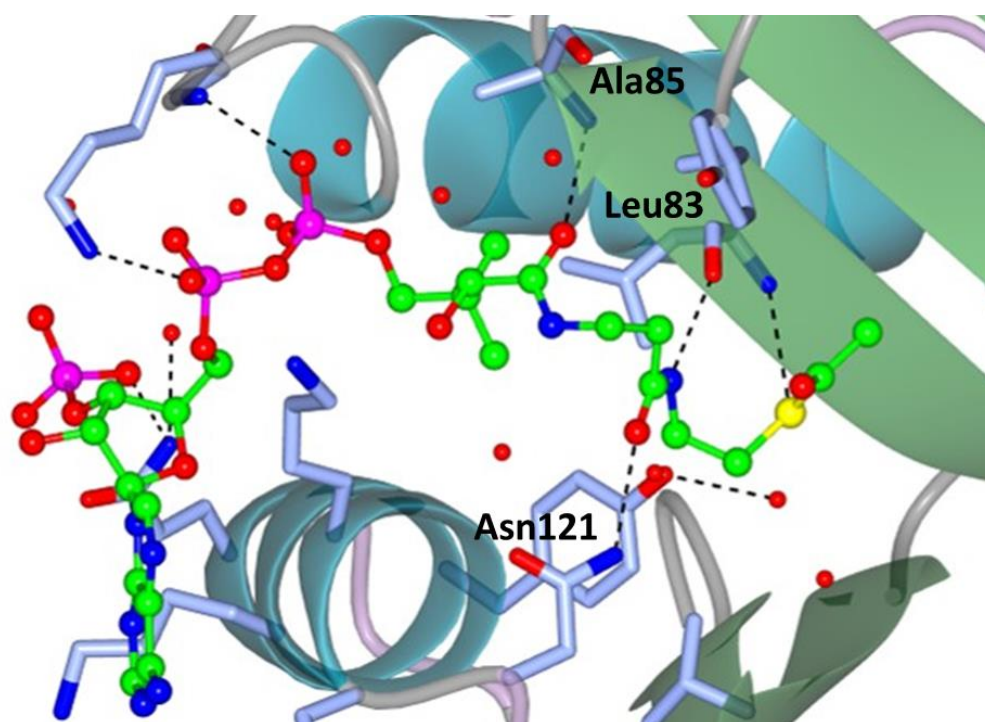




**Figure 5.2** Breakdown of the structure of acetyl-CoA (5.1) and CoA (5.2) into substructures; the pantetheine arm, pyrophosphate group and adenosine moiety.

Whilst considering the binding interaction of acetyl-CoA (5.1) with GNAT enzymes it is useful to divide the molecule into sub-structures; the thio-acetate, the pantetheine arm, the pyrophosphate group and the adenosine group (Figure 5.2). Across all members, binding of the acetyl-CoA co-factor (5.1) appears to be largely reliant on the strong interactions with the acetyl-CoA pantetheine arm with seemingly fewer contributory interactions at more distal distances from the acetyl group.<sup>262</sup> Certain residues in the  $\beta$ -4 and  $\beta$ -5 strands are homologous across species and structural studies complexed with (acetyl)-CoA (5.1/5.2) depict strong hydrogen bonding interactions between the acetyl-CoA pantetheine arm and these residues.<sup>249</sup> For example in *C. jejuni* PseH three hydrogen bonds exist between the pantetheine arm and the surrounding residues; the Ala85 backbone amide can hydrogen bond with one of the pantetheine carbonyls whilst the other carbonyl hydrogen bonds with the Asn121 side chain and the amide proximal to the co-factor acetyl group has a hydrogen bond to the Leu83 backbone carbonyl (Figure 5.3).<sup>232</sup> It is believed that these hydrogen bonds, along with a host of other pantetheine-enzyme interactions help to hold the thio-ester in place for catalytic transfer.

Generally in GNAT enzymes, the pyrophosphate group oxygens are found to interact with the  $\alpha$ 3 helix backbone amide bonds although the number of interactions available and their strength varies greatly between enzymes.<sup>249</sup> The *C. jejuni* PseH crystal structure in complex with acetyl-CoA (5.1) shows the pyrophosphate group arranged between two lysine residues with the Lys90 sidechain suggested to form a hydrogen bond to one of the pyrophosphate oxygens. Other hydrogen bonds occur between pyrophosphate oxygens and amides in the protein backbone such as with Gly91 (Figure 5.3).<sup>232</sup> Unsurprisingly the solvent exposed adenosine part of acetyl-CoA (5.1) has been shown to bind in a variety of conformations in GNAT crystal structures. The *C. jejuni* PseH crystal structure shows adenosine electron density in a number of positions with only a few weak interactions with enzyme residues.<sup>232</sup>



**Figure 5.3** Crystal structure of *C. jejuni* PseH (PDB 4XPL)<sup>232</sup> in complex with its co-factor acetyl-CoA (5.1) (ball and stick model, carbon-bright green, oxygen-red, nitrogen-blue, sulfur-yellow, phosphorus-magenta) showing hydrogen bonds (black dotted lines) between residues (stick model, carbon-ice blue, oxygen-red, nitrogen-blue) and the pantetheine arm and pyrophosphate group.

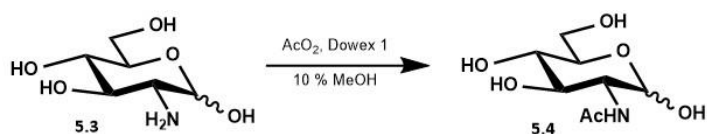
Members of the GNAT superfamily appear to mainly rely on the strong interactions with the acetyl-CoA pantetheine arm for binding of the co-factor with seemingly fewer contributory interactions at more distal distances from the acetyl group. The tight binding of the pantetheine arm is somewhat unsurprising as the acetyl group needs to be held in the correct conformation to be transferred to the substrate. The exact mechanism of transfer is unknown for PseH and, in particular, there is uncertainty surrounding the catalytic base residue. Crystal structures in complex with the product or substrate and mutagenesis studies could provide further insight into the potential the catalytic residues.

Although acetyl Co-A (5.1) is critical in multiple cellular processes,<sup>263</sup> its complex structure renders it an expensive co-factor for use *in vitro*. This expense of acetyl-CoA (5.1) is the major limitation for the utilisation of GNAT enzymes for synthesis and therefore investigations have focussed on eliminating the requirement for excess acetyl-CoA (5.1). Elucidation of the binding interactions has aided in rational design of other acetyl donors that could act as the co-factor for GNAT enzymes. It has previously been shown that for some enzymes, binding affinity may be somewhat retained with substrates that mimic only the acetyl-CoA pantetheine arm and hence these simple molecules may be used in place of acetyl-CoA (5.1) for the enzyme catalysed acetyl transfer.<sup>264</sup>

### 5.1.3 Acetyl-transfer methods during enzymatic syntheses

It has been estimated that acetyl-CoA (**5.1**) is used in ~4 % of all known enzyme reactions,<sup>264</sup> many involving highly specific and selective acetyl transfers that could be synthetically useful tools for desired acetyl transfer reactions. Despite this molecule being abundant in cells and used ubiquitously as an acetyl donating group, its complex structure means that it is also costly and complicated to chemically synthesise.<sup>255</sup> Therefore methods to reduce the amount of acetyl-CoA (**5.1**) required have been investigated to allow for use of acetyltransferases in synthesis to be more economically viable. Chemical acetylation can be an attractive alternative instead of use of an acetyltransferase as it completely removes the need for the acetyl-CoA (**5.1**) co-factor, and is especially effective in the synthesis of simple molecules.<sup>265</sup> However in some cases it can be more advantageous to perform enzymatic acetyl-transfer; during an enzymatic synthesis<sup>266</sup> or when reacting more complex structures with multiple sites of potential acetylation,<sup>267</sup> for example. Both chemical and enzymatic schemes for acylation of amines have been explored and are discussed below with emphasis on methods that are selective, low cost and applicable to use within enzymatic pathways to carbohydrates.

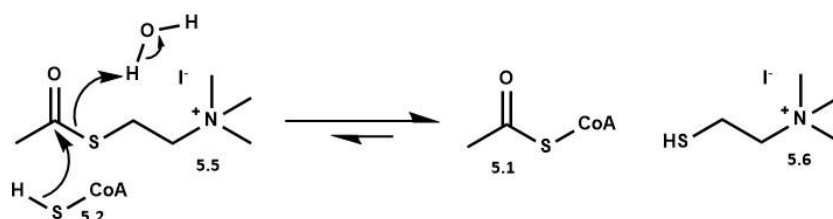
A multitude of chemical acetylation protocols exist, suggesting many different potential reagents and conditions for amine acetylation,<sup>268-271</sup> however not all are suitable for acetyl transfer onto aminoglycosides. Most importantly for this reaction, it is desired for the method to be highly selective for amine acetylation so that it is not necessary to protect other functional groups such as the free hydroxyls often present in sugar molecules, however this is difficult to achieve fully. For example, the simple sugar GlcN (**5.3**) can be *N*-acetylated with acetic anhydride in 10 % methanol and Dowex 1 (carbonate form). A 43 % yield is reported after filtration and crystallisation and although the GlcNAc (**5.4**) product is the major form, minor (< 1 %) *O*-acetylated products can be detected.<sup>272</sup> Additionally chemical acetylation methods should preferentially only use non-toxic, readily available reagents and solvents under mild conditions. Chemical methods are often flawed for use within *in vitro* enzymatic workflows, as the components of the chemical reactions are invariably incompatible with subsequent enzyme reactions and can reduce enzyme efficiency or require additional purification and solvent exchange steps.



**Scheme 5.3** A synthetic route of acetyltransfer to GlcN (**5.3**) to produce GlcNAc (**5.4**).

Nature has developed acetyl-CoA (**5.1**) regeneration systems itself to prevent the need for continual synthesis of such a complex and important molecule. For example in aerobic organisms acetyl-CoA (**5.1**) is an initiator of the Krebs cycle and its regeneration is catalysed by three enzymes involved in fixing pyruvic acid, formation of a disulphide bond and donation of the acetyl group.<sup>273</sup> Less complex natural systems have also been identified that only require one enzyme for the catalysis of acetyl transfer onto CoA (**5.2**) and of these, acetyl-CoA synthetase (ACS), phosphotransacetylase (PTA) or carnitine acetyltransferase (CAT) are generally employed *in vitro*.<sup>274</sup> Each enzyme displays efficient transfer of an acetyl group to CoA (**5.1**) and can be used *in situ* to regenerate acetyl-CoA (**5.2**).<sup>275</sup> However industrial utilisation is still far from simple as issues arise with the substrates and co-factors needed for these enzymatic reactions. ACS requires another expensive co-factor ATP to catalyse the coupling of carboxylic acids to CoA (**5.2**),<sup>276</sup> and PTA and CAT require acyl substrates that can be difficult to prepare and are unstable limiting their practicality for wide use as acetyl-CoA (**5.1**) regenerators.<sup>277-278</sup>

Chemical methods for regeneration of acetyl-CoA (**5.1**) have also been investigated, with syntheses detailing the use of small molecules such as acid anhydrides<sup>279</sup> or N-hydroxysuccinimide esters of fatty acids.<sup>280</sup> However these methods cannot be utilised efficiently in all reaction schemes. They are generally not specific for CoA reacylation and require organic solvents so cannot be carried out *in situ* within enzymatic reactions, necessitating extra purification steps and extended synthetic routes. One chemical regeneration factor, acetyl-thiocholine iodide (**5.5**), has been identified that can overcome these issues.<sup>281</sup> This regeneration factor is water soluble and importantly is primed for attack by the CoA thiol (**5.2**) as it has an electron withdrawing group which increases the electrophilicity of the carbonyl. The electron withdrawing group also aids in prevention of the reverse reaction by decreasing the nucleophilicity of the resulting thiol (**5.6**) (**Scheme 5.4**). Acetyl-thiocholine iodide (**5.5**) was found to act as an efficient regeneration factor of acetyl-CoA (**5.1**) during enzymatic reactions with substoichiometric levels of acetyl-CoA (**5.1**) for example during the citrate synthase catalysed synthesis of citric acid.<sup>281</sup>

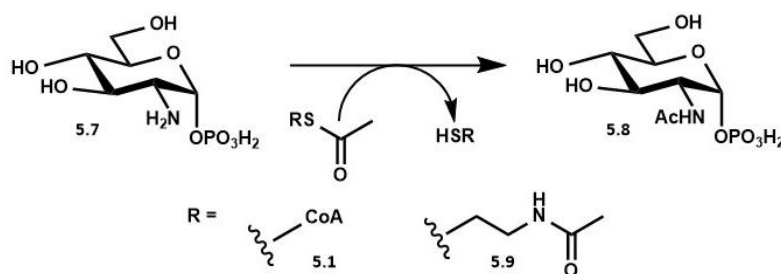


**Scheme 5.4** Mechanism for the acetyl transfer from acetyl-thiocholine iodide (**5.5**) to CoA (**5.1**).

#### 5.1.4 Introduction to the promiscuity of aminoglycoside *N*-acetyltransferases co-factor binding sites

A number of acetyl-transferase enzymes utilising acetyl-CoA (**5.1**) as the co-factor have displayed co-factor promiscuity in the structure the CoA (**5.2**) molecule and in the structure of acylated group they will transfer to the acceptor molecule. This suggests that acetyltransferases could be used *in vitro* to enzymatically transfer a variety of small acylated derivatives from CoA analogues onto natural acceptor molecules.

It has been proposed that the reactivity of the acetyl transfer utilising acetyl-CoA (**5.1**) stems from the thio-ester and is largely unaffected by the distal parts of the acetyl-CoA (**5.1**) structure.<sup>282</sup> Additionally acetyltransferase crystal structures in complex with the co-factor have shown that the majority of enzyme-co-factor binding interactions occur between the parts of the acetyl-CoA (**5.1**) structure proximal to the thioacetate group.<sup>262</sup> Therefore CoA analogues have been synthesised that mimic the thioacetate and pantetheine part of the acetyl-CoA (**5.1**) structure, such as *S*-acetylcysteamine, with the aim to provide less expensive acetyltransferase co-factor alternatives to acetyl-CoA (**5.1**).<sup>264</sup> For example, the rate of acetyl-transfer to GlcN-1-P (**5.7**), to form GlcNAc-1-P (**5.8**), catalysed by the *Pyrococcus furiosus* bifunctional GlcN-1-P (**5.7**) acyltransferase/GlcNAc-1-P uridylyltransferase was reduced only 2-3 fold when using *N*-acetylcysteamine thioester (**5.9**) compared to acetyl-CoA (**5.1**) (**Scheme 5.5**).<sup>283</sup> This demonstrates that the need for acetyl-CoA (**5.1**) can be removed by exploring the activity of acetyl-transferase enzymes with much less expensive unnatural co-factor substitutes.



**Scheme 5.5** Acetyltransferase catalysed acetylation of GlcN-1-P utilising the natural co-factor acetyl-CoA (**5.1**) or *N*-acetylcysteamine thioester (**5.9**).

Additionally PatB, a peptidoglycan *O*-acetyltransferase, that catalyses transfer of an acetyl group from a trans-membrane protein to the C6 hydroxyl of MurNAc<sup>284</sup> can accept acetyl-CoA (**5.1**) as an acetyl donor as well as a number of structurally different molecules. For example, PatB was shown to catalyse acetyltransfer from *p*-nitro-phenol-acetate to chitotriose<sup>285</sup> and 4-methylumbelliferyl acetate to chitopentaose<sup>286</sup>. This suggests that some acetyltransferases may

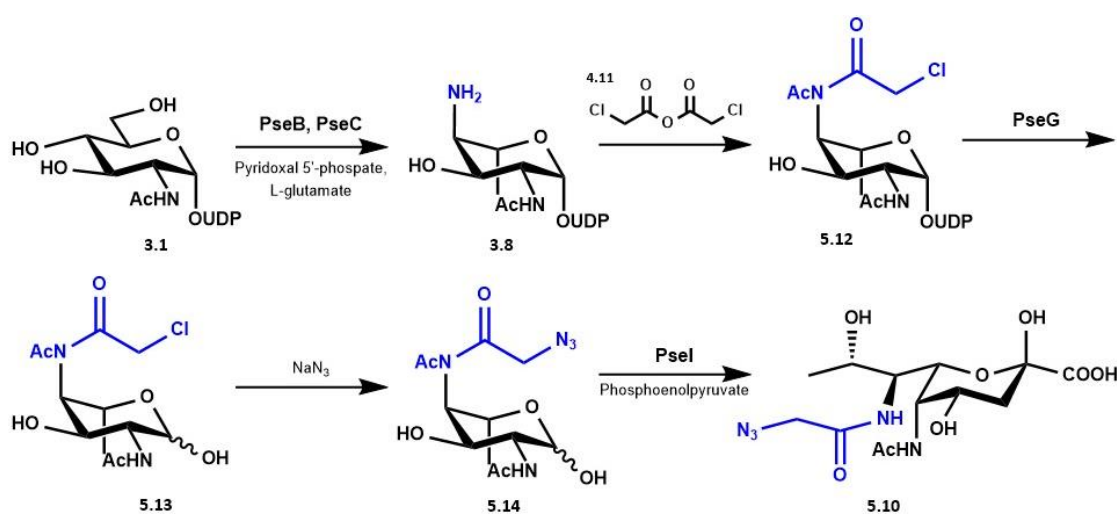
have even greater donor promiscuity and that co-factor substitutes may not need to be limited to acetyl-CoA (**5.1**) analogues.

Some acetyltransferase enzymes have also been shown to catalyse transfer a number of different acyl-CoA groups onto substrates, highlighting a potential strategy for the introduction of additional functionality onto amine groups. Naturally, a number of acyl-CoA structures exist and are utilised as co-factors for specific transfer reactions. For example *N*-myristoyl transferases exhibit specificity for their native co-factor; myristoyl-CoA, catalysing transfer of the myristoyl group onto amino-terminal glycine residues.<sup>211</sup> As acetyl-CoA (**5.1**) is more abundant than other acyl derivatives enzymes utilising it often display less specificity and have been shown to accept a number of acyl-CoA donors. However activity can quickly decrease as the acyl group differs more greatly from an acetyl group, for example during analysis of *Streptomyces rubellomurinus* FrbF catalysed acyl transfer to CMP-3-aminopropylphosphonate, with propionyl-CoA, malonyl-CoA and acetoacetyl-CoA were found to yield quantifiable amounts of product however the expected product was not detected using hydroxybutyryl-CoA and glutaryl-CoA.<sup>287</sup>

Visualisation of glycans in their native cellular environment is heavily desired to aid insight into their roles and to better understand the significance of these cellular structures. The Bertozzi group has been at the front-line for enabling such studies *via* metabolic labelling of glycans with bioorthogonal handles.<sup>288-290</sup> It has been shown that a number of simple azido-labelled sugars can be transported into cells and successfully displayed in place of the native sugars on cell structures within cell surface glycoconjugates.<sup>289</sup> In this strategy the azido-tagged sugars are processed by biosynthetic enzymes to afford more complex sugars to be displayed on glycans e.g. chemically synthesised ManNAz can be fed to cells and processed into Neu5Az<sup>291</sup> which can then be bioorthogonally tagged on cell surfaces using a variety of “click” chemistries.<sup>292</sup>

Although production of ManNAz is not synthetically challenging, labelling of more complex sugars with azido groups can be chemically challenging. Therefore other techniques for installation of azido groups into carbohydrates have been investigated. Due to their promiscuity surrounding the acyl group, some acetyltransferase enzymes have been proposed as suitable for catalysing azido-acyl transfer to carbohydrate substrates.<sup>293</sup> However due to the cost and complexity of the CoA structure (**5.1**), it is desired to be able to use azido-derivatives of small structural mimics of CoA (**5.1**) as azido-acyl donors with acetyltransferases. This has been demonstrated by incorporation of bioorthogonal handles during the PatB catalysed acylation of MurNAc with *N*-acylcysteamine thioesters.<sup>294</sup>

Synthesis of Pse5Ac7Az (**5.10**) has previously been achieved *in vitro* and *in vivo* and visualised through fluorescent click labelling of *C. jejuni* flagellin.<sup>295</sup> This strategy used PseB and PseC to synthesise a C4 amino group (**3.8**) which was then chemically acetylated with chloroacetic anhydride (**5.11**) to give UDP-2-acetamido-4-chloroacetamido-4,6-dideoxy- $\beta$ -l-AltNAc (**5.12**). Following enzymatic hydrolysis of the UDP group with PseG, treatment with sodium azide afforded the Pse5Ac7Az biosynthetic precursor, UDP-4-azidoacetamido-4,6-dideoxy- $\beta$ -l-AltNAc (**5.14**). It was shown that this molecule could be turned over by PseI *in vitro* (**Scheme 5.6**) and *in vivo* when fed to *C. jejuni* cells. However this strategy to Pse5Ac7Az (**5.10**) requires alternating between chemical and enzymatic steps resulting in a number of time consuming purifications and solvent exchanges *en route* to the desired final compound.<sup>295</sup>



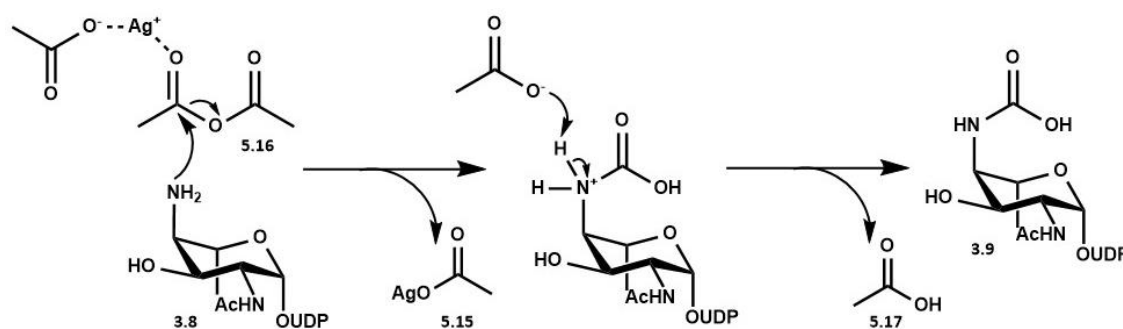
**Scheme 5.6** A chemo-enzymatic route to the Pse5Ac7Az (**5.10**) utilising a combination of enzymatic and chemical steps.

One of the major limitations of scaling up the *in vitro* enzymatic synthesis of Pse5Ac7Ac (**1.13**) lies with the expense of acetyl-CoA (**5.1**) which is required in excess during the PseH catalysed acetyl-transfer. Three alternate methods for acetyl-transfer have been explored; chemical acetylation, enzyme catalysed acetylation coupled with acetyl-CoA (**5.1**) co-factor regeneration and use of other acetyl-donating co-factor substitutes. The methods were evaluated based on availability, cost, practicality and viability for use within the enzymatic system for production of Pse5Ac7Ac (**1.13**).

## 5.2 Acetylation during the chemo-enzymatic synthesis of Pse5Ac7Ac

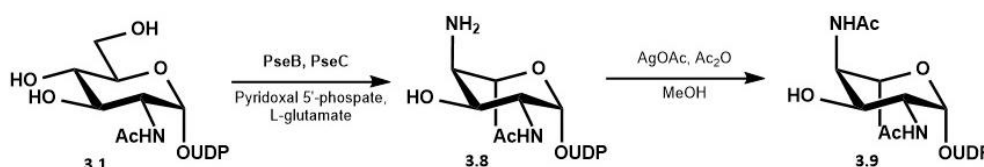
### 5.2.1 Chemical acetylation of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc

After considering a number of chemical methods it was decided to use a heterogeneous silver acetate catalyst with acetic anhydride to selectively acetylate the C5 amine.<sup>296</sup> This method had already been detailed for the synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) and observed to selectively form the desired acetamido product with quantitative conversion.<sup>228</sup> The silver acetate (**5.15**) is utilised as a catalyst, activating the acetic anhydride (**5.16**) by coordinating to a carbonyl oxygen. Hence, electron density is further drawn away from the carbonyl carbon, increasing its electrophilicity and making it more susceptible to nucleophilic attack. This promotion of carbonyl electrophilicity removes the requirement for addition of a base to deprotonate the amine prior to the acetylation step and this occurs after acetyl transfer, producing acetic acid (**5.17**) as a by-product (**Scheme 5.7**).



**Scheme 5.7** Chemical acetylation of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) to afford UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) using a silver acetate catalyst (**5.15**) to activate the acetic anhydride making it more susceptible to nucleophilic attack.

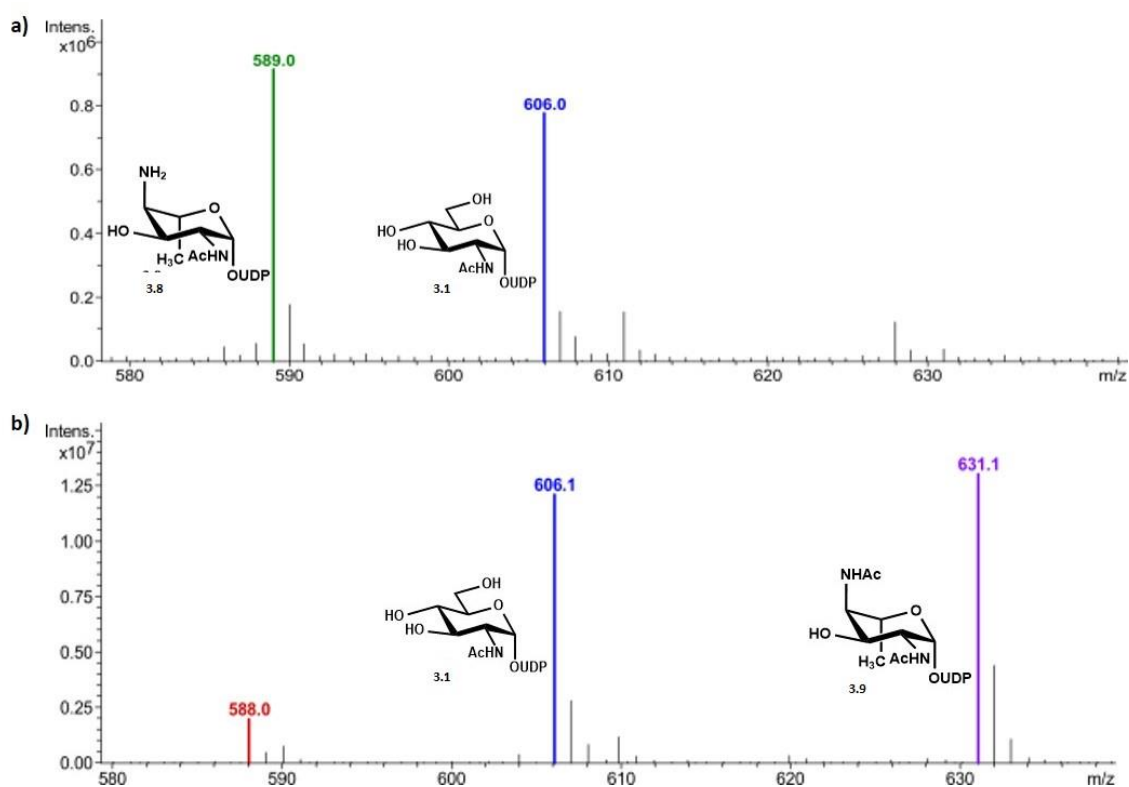
A chemo-enzymatic synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) was proposed (**Scheme 5.8**). The PseH reagent was prepared using a coupled PseB and PseC reaction; UDP-GlcNAc (1 mM) was suspended in 50 mM Tris-HCl (pH 7.4) with PseB (25  $\mu$ M) and PseC (25  $\mu$ M) along with the PseC co-factors PLP (2.5 mM) and L-Glu (10 mM). The reaction was shaken (120 rpm) at 37  $^{\circ}$ C and monitored by negative ESI LC-MS for 5 hrs (**Figure 5.4a**).



**Scheme 5.8** Enzymatic synthesis of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) followed by chemical acetylation to UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) using a silver acetate catalyst (4 equiv) and acetic anhydride (11.5 equiv) in MeOH, at room temperature for 3 hrs.



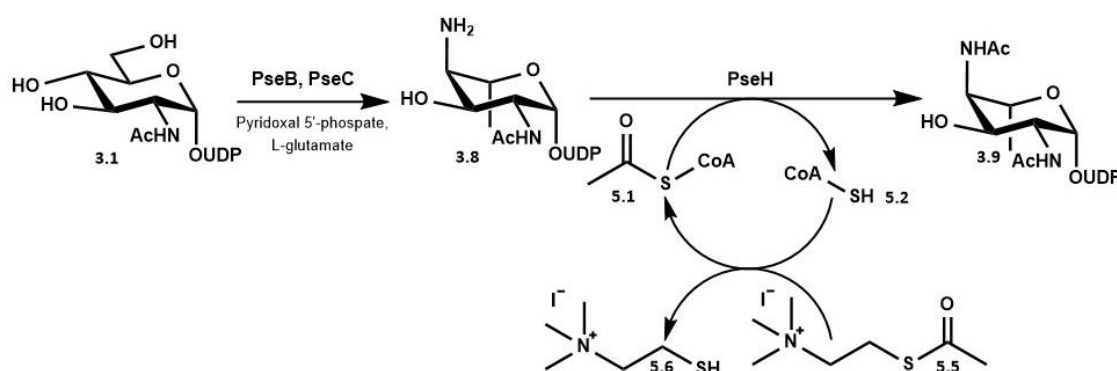
Acetic anhydride (11.5 equivs) and silver acetate (1 equiv) were added to the crude PseC product (**3.8**) in methanol and the reaction monitored by negative ESI LC-MS. Equivalentts were calculated based on quantitative conversion (1.65  $\mu\text{mol}$ ) to the PseC product (**3.8**) as yield could not be calculated from the mass of the crude reaction. A decrease in relative intensity of the  $[\text{M}-\text{H}]^-$  589 peak was observed alongside the appearance of the expected peak with a  $[\text{M}-\text{H}]^-$  of 631 corresponding to the acylated sugar molecule (**3.9**). After 3 hrs the presence of the substrate (**3.8**) was no longer detectable by negative ESI LC-MS (**Figure 5.4b**). The heterogenous catalyst was removed by filtration and product lyophilised three times in water to remove excess acetic anhydride. Contrary to the previously published strategy,<sup>228</sup> this synthesis has shown that time-costly purification steps between enzymatic and chemical steps are unnecessary for substrate turnover. Additionally it was discovered that the amount of silver acetate catalyst required for efficient turnover was much less than previously published (1 equiv instead of 4 equiv)<sup>228</sup> bypassing the need for excess acetyl-CoA (**5.1**) and an extensive chemical synthesis to afford Pse5Ac7Ac (**1.13**). However despite advantages over previously published syntheses, the use of chemical acetylation within the enzymatic synthesis of Pse5Ac7Ac (**1.13**) does require two additional work-up procedures and no longer has the benefits of being a one-pot system.



**Figure 5.4** Negative ESI LC-MS demonstrating **a**) Enzymatic conversion of the PseC product (**3.8**) and **b**) chemical acetylation to afford the PseH (**3.9**) product.

### 5.2.2 *In situ* regeneration of acetyl-CoA using acetyl thiocholine iodide

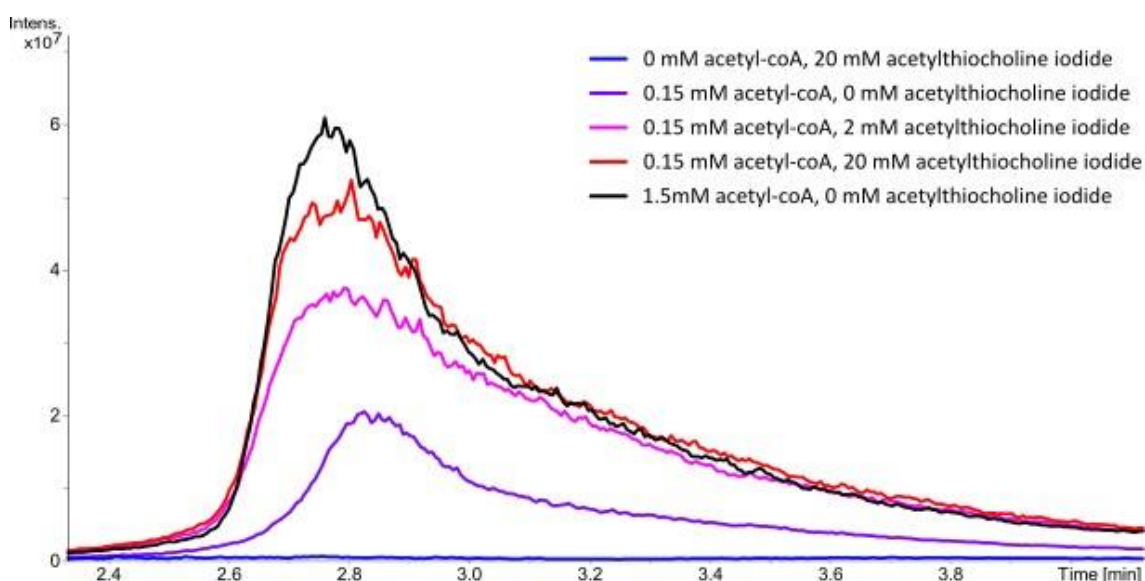
Regeneration of sub-stoichiometric amounts of acetyl-CoA (**5.1**) *in situ* during PseH turnover was also investigated as an acetylation strategy. Regeneration of acetyl-CoA (**5.1**) occurs specifically through attack of the CoA thiol (**5.2**) formed as a by-product during PseH turnover (**Scheme 5.9**). Therefore thioesters are ideal regeneration factors, whereby a thioester exchange could occur *in situ* to enable regeneration of acetyl-CoA (**5.1**). This could facilitate the enzymatic synthesis of Pse5Ac7Ac (**1.13**) in a one-pot procedure, while also enabling the reduction of the quantity of acetyl-CoA (**5.1**) initially required for optimal conversion. Acetyl thiocholine iodide (**5.5**) was investigated as a low cost, commercially available CoA (**5.1**) acetyl donor which is stable and soluble under aqueous conditions (**Scheme 5.9**). The electronegative nitrogen in the choline group was suggested to promote the desired reaction by increasing the electropositivity of the carbonyl carbon in the acetylated structure (**5.5**) and deter the reverse reaction by drawing electron density away from the resulting thiol (**5.6**).



**Scheme 5.9** Regeneration of acetyl-CoA (**5.1**) *in situ* during the enzymatic synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**).

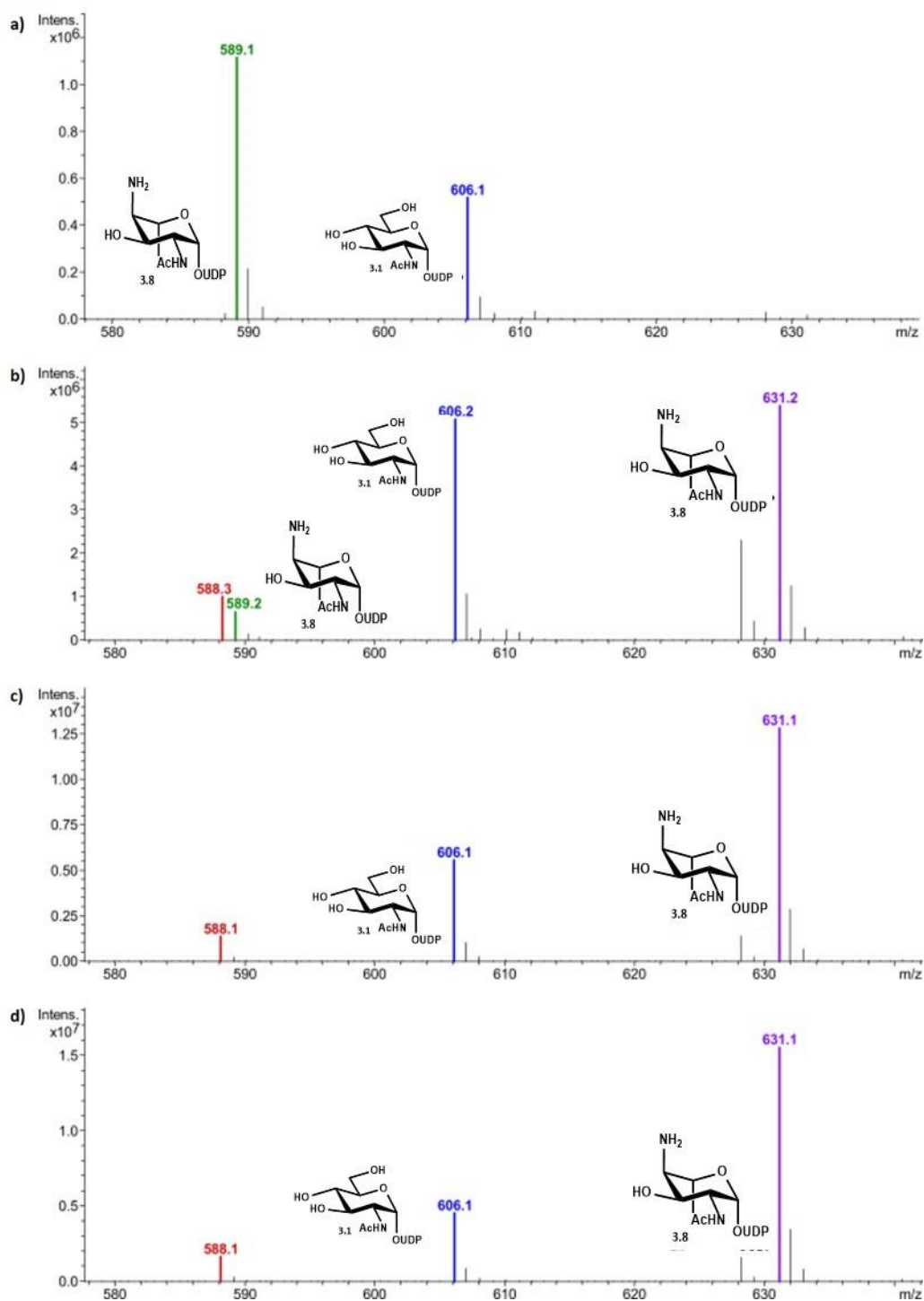
In order to ascertain the efficiency of acetyl-CoA (**5.1**) regeneration within this system, a number of small scale (0.1 mg UDP-GlcNAc) reaction mixtures were set up containing UDP-GlcNAc (1 mM), Pse B (25  $\mu$ M), Pse C (125  $\mu$ M) and their associated co-factors. PseH (50  $\mu$ M) was added to the reaction mixture and two controls were devised; a negative control containing 0 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide, and a positive control containing 1.5 mM acetyl-CoA (**5.1**). To the remaining reactions, increasing concentrations (0 mM to 20 mM) of acetyl thiocholine iodide (**5.5**) were added and all reactions were shaken (120 rpm) at 37  $^{\circ}$ C.

The reactions were monitored with negative ESI LC-MS and relative turnover to acetylated product (**3.9**) was calculated as a percentage of the total sugar compounds present in the reaction mixture. Qualitative conversion was observed when comparing the  $m/z$  631 ion trace (Pse H product) of all of the reactions after 4 hrs. It was observed that addition of acetylthiocholine iodide (**5.5**) to the reactions containing sub-stoichiometric amounts of acetyl-CoA (**5.1**) increased conversion to the PseH product (**3.9**) and hence acted as a co-factor regenerator, although this did not occur as efficiently compared to the reaction containing excess acetyl-CoA (**5.1**) (**Figure 5.5**).



**Figure 5.5** Overlay of the LC-MS negative ESI  $m/z$  631 extracted ion count traces after 4 hrs showing the effect of addition of the acetylthiocholine iodide on PseH catalysed acetyltransfer whilst using sub-stoichiometric amounts of acetyl-CoA (**5.1**).

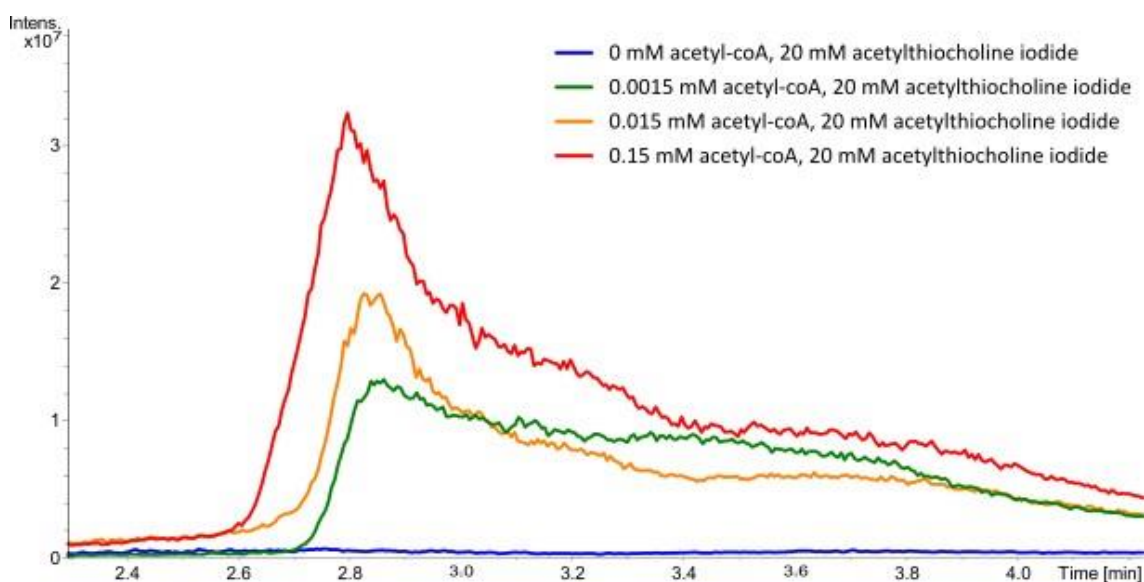
Analysis of the negative ESI LC-MS data showed that there was no detectable conversion to the acetylated sugar (**3.9**) in the reaction mixture with 0 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**) suggesting that it could not act as the co-factor itself, and nor could any of the other reaction constituents (**Figure 5.6a**). Decrease of the acetyl-CoA (**5.1**) concentration to 0.15 mM with no regeneration factor showed a marked decrease in the turnover to the PseH product (**3.9**); with only 44 % conversion (**Figure 5.6b**). However this conversion could be increased to 66 % with the addition of 2 mM acetylthiocholine iodide (**5.5**) (**Figure 5.6c**) and 72 % conversion when the reaction included 20 mM acetylthiocholine iodide (**5.5**) (**Figure 5.6d**).



**Figure 5.6** Negative ESI LC-MS analysis of conversion to the PseH product (**3.9**) from UDP-GlcNAc (**3.1**), investigating the use of acetylthiocholine iodide (**5.5**) as a regeneration factor with sub-stoichiometric amounts of acetyl-CoA (**5.1**) **a)** 0 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**), **b)** 0.15 mM acetyl-CoA (**5.1**) and 0 mM acetylthiocholine iodide (**5.5**), **c)** 0.15 mM acetyl-CoA (**5.1**) and 2 mM acetylthiocholine iodide (**5.5**), and **d)** 0.15 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**).

As the reaction containing 0.15 mM acetyl-CoA (**5.1**) and 20 mM acetylthiocholine iodide (**5.5**) displayed the highest level of conversion (72 %) to the PseH product, reactions using this

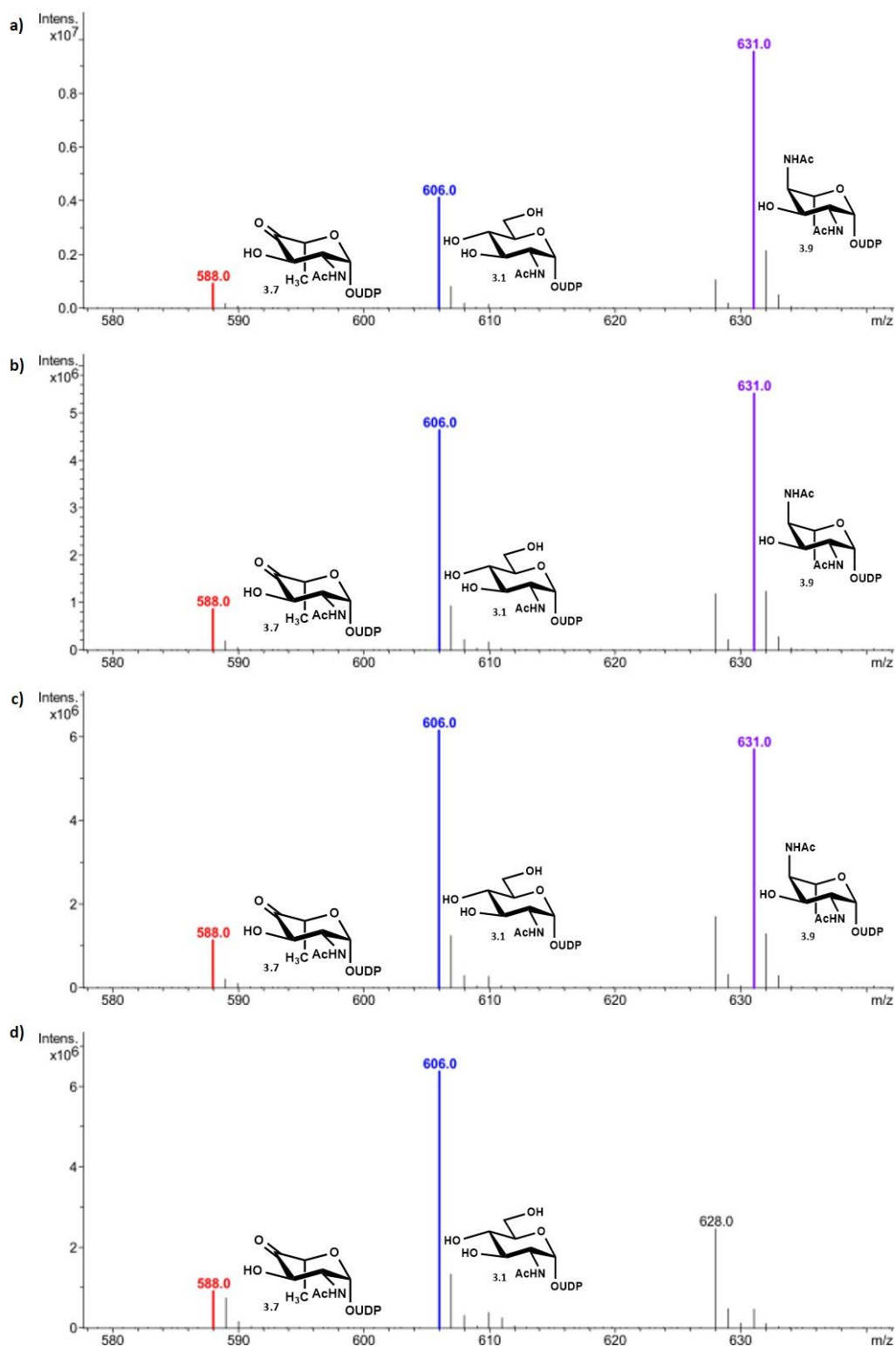
concentration of acetylthiocholine iodide (**5.5**) were further explored. Additionally, in order to further exploit this regeneration method, CoA (**5.2**) was trialled instead of acetyl-CoA (**5.1**) due to its lower cost. Similar reaction mixtures were constructed UDP-GlcNAc (1 mM), Pse B (25  $\mu$ M), Pse C (125  $\mu$ M), their associated co-factors and PseH (50  $\mu$ M). However the concentration of acetylthiocholine iodide (**5.5**) across reactions was kept constant (20 mM) and increasing concentrations of acetyl-CoA (**5.1**) (0 mM to 0.15 mM) were added. All reactions were shaken (120 rpm) at 37 °C and monitored with negative ESI LC-MS. After 4 hrs, negative LC-MS ESI showed that production of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) could be achieved in reaction mixtures containing 20 mM acetylthiocholine iodide (**5.5**) and all concentrations of CoA (**5.2**) screened (**Figure 5.7**, **5.8**). However, as anticipated, comparison of the m/z 631 ion trace showed that conversion decreased as the CoA (**5.1**) concentration was decreased from 0.15 mM to 0.0015 mM; suggestive that the supply of acetyl-coA (**5.1**) was limiting the rate of the reaction (**Figure 5.7**).



**Figure 4.7** Overlay of the LC-MS negative ESI m/z 631 extracted ion count traces after 4 hrs showing the effect of decreasing the concentration of CoA from sub-stoichiometric to catalytic amounts on PseH catalysed acetyltransfer in the presence of 20 mM regeneration factor; acetylthiocholine iodide (**5.5**).

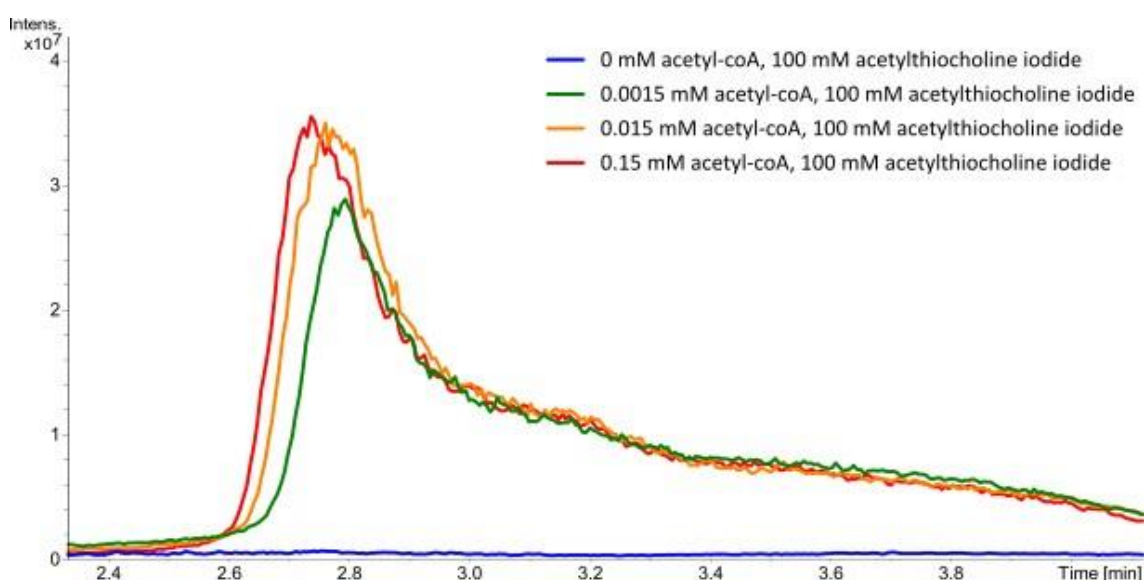
Comparison of the relative intensities of the product (**3.9**) peak, as a percentage of the total intensity of peaks corresponding to sugar products showed that utilising 20 mM regeneration factor (**5.5**) with substoichiometric amounts of CoA (**5.2**) generated turnover to the PseH product (**3.9**). A 65 % conversion was achieved when the concentration was ten-fold less than the original amount of acetyl-CoA (**5.1**) required (**Figure 5.8a**). Decreasing the concentration by another factor of ten (0.015 mM) reduced the conversion rate to 49 % (**Figure 5.8b**) and a 44 %

conversion was still achieved when catalytic amounts of CoA (**5.2**) were utilised (0.0015 mM) (**Figure 5.8c**).



**Figure 5.8** LC-MS negative ESI mass spectra demonstrating conversion to the acetylated PseH product (**3.9**) using 20 mM regeneration factor (**5.5**) and substoichiometric amounts of coA (**5.2**); **a**) 0.15 mM CoA (**5.2**), **b**) 0.015 mM CoA (**5.2**), **c**) 0.0015 mM CoA (**5.2**), **d**) 0mM CoA (**5.2**).

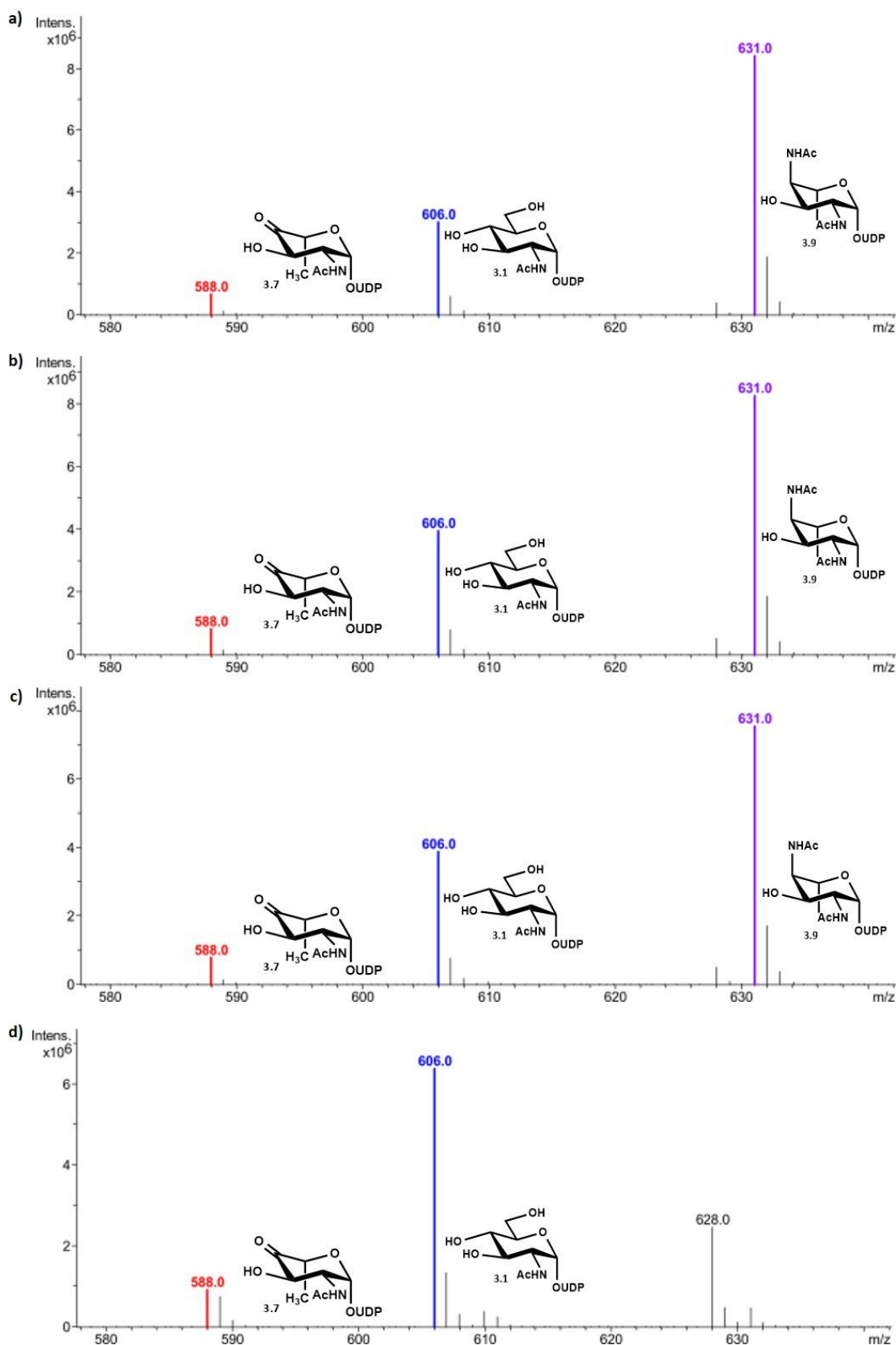
In order to practically use this co-factor regeneration strategy in the “one-pot” synthesis of Pse5Ac7Ac (**1.13**), further optimisation was required to increase the conversion rate. Therefore identical reactions mixtures were constructed as before, with varying concentrations of CoA (**5.2**), however in these reactions, the concentration of the regeneration factor (**5.5**) was increased to 100 mM. Notably, increasing the concentration of acetyl thiocholine iodide (**5.5**) from 20 mM to 100 mM drastically reduced the effect of decreasing the concentration of CoA (**5.2**) from 0.15 mM to 0.0015 mM (**Figure 5.9**).



**Figure 5.9** Overlay of the LC-MS negative ESI  $m/z$  631 extracted ion count traces after 4 hrs showing the effect of decreasing the concentration of CoA from sub-stoichiometric to catalytic amounts on PseH catalysed acetyltransfer in the presence of 100 mM regeneration factor; acetylthiocholine iodide (**5.5**).

Upon comparison of the relative intensities of the product (**3.9**) peak after 4 hrs in negative LC-MS analysis, as a percentage of the total sugars in the reaction mixture, the reactions had much more similar levels of conversion to the PseH product (**3.9**) when decreasing the CoA concentration; 68 %, 64 % and 61 % respectively (**Figure 5.10**).

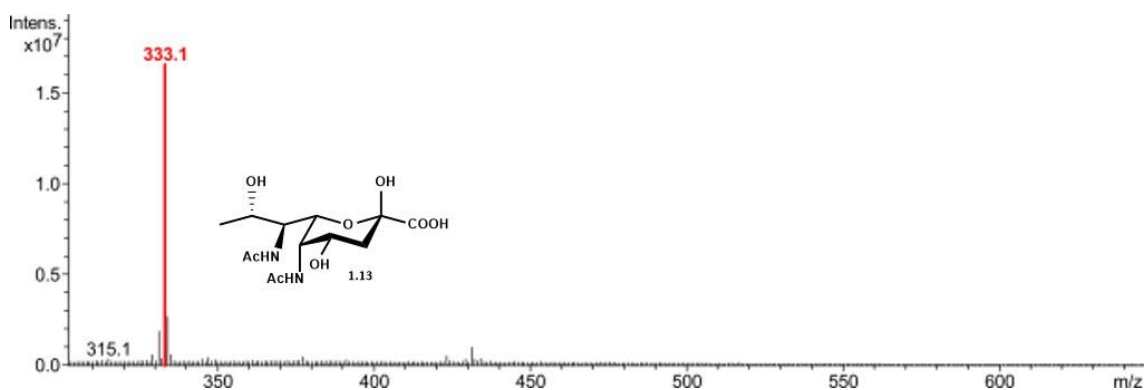
This strategy successfully demonstrated a synthetic route to UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**) in a “one-pot” enzymatic system where use of a regeneration co-factor could compensate for using a vastly reduced concentration of CoA (**5.2**) (1000 fold decrease compared to the original synthesis). Therefore these reaction conditions were deemed suitable for potential use in a large scale synthesis as the activity mainly retained but the cost of the reaction is drastically reduced. For example, using acetyl-CoA (**5.2**) at a concentration of 1.5 mM in a reaction containing 100 mg UDP-GlcNAc (**3.1**) the cost of the required quantity of co-factor would be £1864 (calculated using Sigma prices) but using 0.0015 mM CoA and 10 mM acetylthiocholine iodide (**5.5**) it was calculated to only cost £124 for the equivalent reaction.



**Figure 5.10** LC-MS negative ESI mass spectra demonstrating conversion to the acetylated PseH product (3.9) using 100 mM regeneration factor (5.5) and substoichiometric amounts of CoA (4.4) ; **a)** 0.15 mM CoA (5.2), **b)** 0.015 mM CoA (5.2), **c)** 0.0015 mM CoA (5.2), **d)** 0 mM CoA (5.2).



After the promising results utilising a regeneration factor for the synthesis of the PseH product (**3.9**) it was attempted to expand this reaction for the “one-pot” enzymatic synthesis of Pse5Ac7Ac (**1.13**). A reaction containing UDP-GlcNAc (1 mM), PLP (2.5 mM), L-Glu (10 mM), CoA (0.0015 mM), acetylthiocholine iodide (100 mM), PEP (1.5 mM) and the biosynthetic enzymes PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M) and PseI (25  $\mu$ M) in sodium phosphate buffer (50 mM, pH 7.4) was incubated, shaking (120 rpm) at 37 °C. Negative ESI LC-MS analysis indicated that that the addition of 100 mM acetylthiocholine iodide (**5.5**) in the reaction mixture did not have an adverse effect on the activities of the final biosynthetic enzymes; PseG and PseI as conversion to the Pse5Ac7Ac (**1.13**) product was observed. After 5 hrs negative ESI LC-MS showed complete conversion to Pse5Ac7Ac (**1.13**) using this new system; 100 mM acetylthiocholine iodide (**5.5**) and 0.0015 mM CoA (**5.2**), as determined by the lack of peaks corresponding to the Pse5Ac7Ac (**1.13**) biosynthetic starting material (**3.1**) and intermediates (**3.7-3.10**) (Figure 5.11).



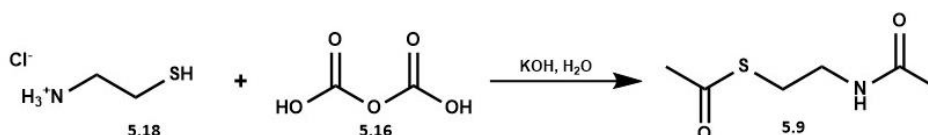
**Figure 5.11** LC-MS negative ESI demonstrating successful production of Pse5Ac7Ac (**3.1**) in the “one-pot” system when employing 0.0015 mM CoA (**5.2**) and 100 mM acetyl thiocholine iodide (**5.5**).

*In situ* co-factor regeneration during the PseH reaction makes the one-pot enzymatic synthesis of Pse5Ac7Ac (**1.13**) far more economically viable, potentially facilitating large scale production. However a potential limitation for the general use of acetyl thiocholine iodide (**5.5**) as a regeneration factor surrounds its anticipated reactivity as an acetyl transfer agent, which could result in non-selective acetylation during attempted synthesis of other pseudaminic acid derivatives. Therefore other alternatives for acetyl transfer regeneration factors could be investigated.

### 5.2.3 *N*-acetyl cysteamine thioacetate as a PseH co-factor substitute

Analysis of the binding of *C. jejuni* PseH to acetyl-CoA (**5.1**) provides a potentially powerful insight for the design of alternative co-factors. Briefly, the adenosine group of acetyl-CoA is surface bound, fairly flexible and thought to only marginally contribute to binding interactions with *C. jejuni* PseH. Contrastingly the pantetheine moiety and pyrophosphate group are more deeply inserted into the co-factor binding groove and participate in diverse interactions with nearby residues. Thus suggesting mimics of these portions of acetyl-CoA (**5.1**) could behave as substitute co-factors, potentially removing the requirement for even catalytic quantities of CoA (**5.2**) in the enzymatic synthesis of Pse5Ac7Ac (**1.13**).

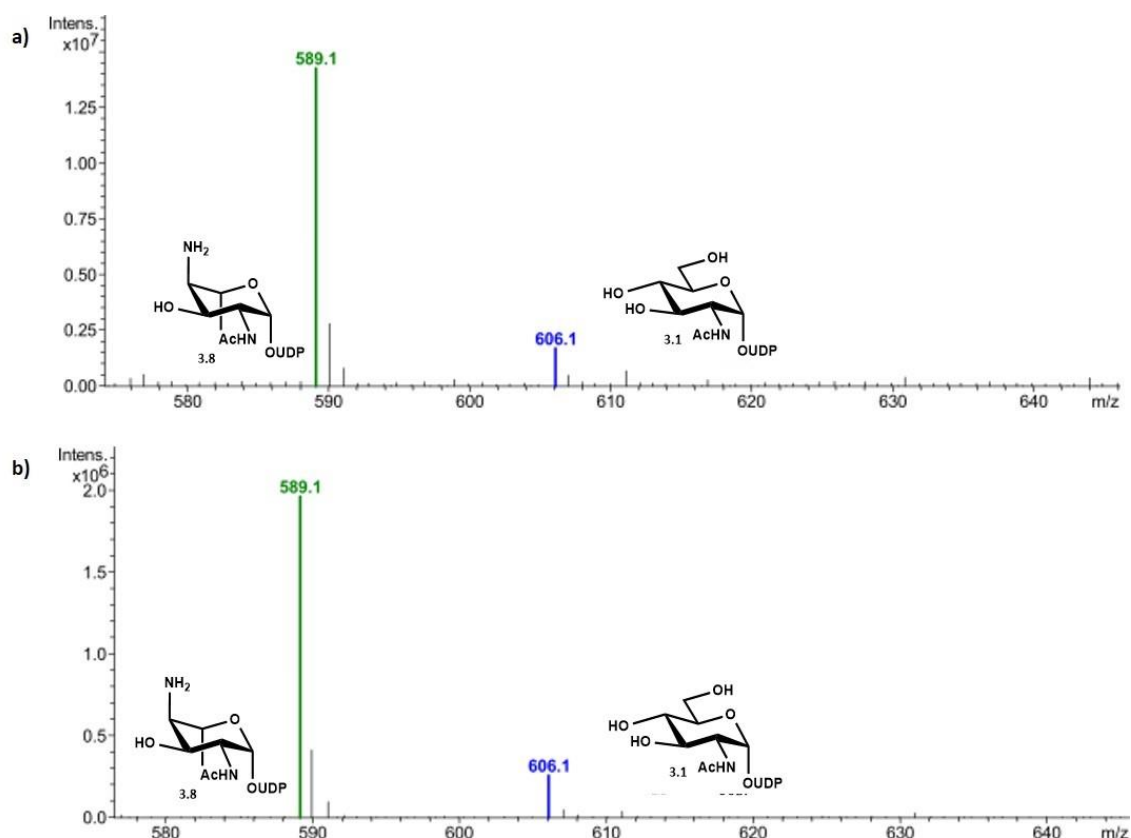
*N*-acetyl cysteamine thioacetate (**5.9**) was recognised as a short-chain acetyl donor that is structurally similar to part of the acetylated pantetheine arm of acetyl-CoA (**5.1**). The structure is composed of the thioacetate group which is susceptible to nucleophilic attack in acetyltransferase reactions and an *N*-acetylcysteamine that has been proposed to make key binding interactions with *C. jejuni* PseH.<sup>232</sup> It has recently been shown to act as an acetyl donor for PatB catalysed *O*-acetylation peptidoglycan synthesis.<sup>294</sup> This molecule was therefore also investigated as a co-factor substitute in PseH catalysed acetylation as it mimics the most tightly bound portion of the acetyl-CoA (**5.1**) structure and can be easily accessed from readily available reagents.<sup>297</sup>



**Scheme 5.10** Synthetic route to *N*-acetyl cysteamine thioacetate; addition of acetic anhydride to cysteamine HCl in H<sub>2</sub>O at pH 8.0, 0 °C, followed by stirring at pH 7.0 at 0 °C for 2 hrs.

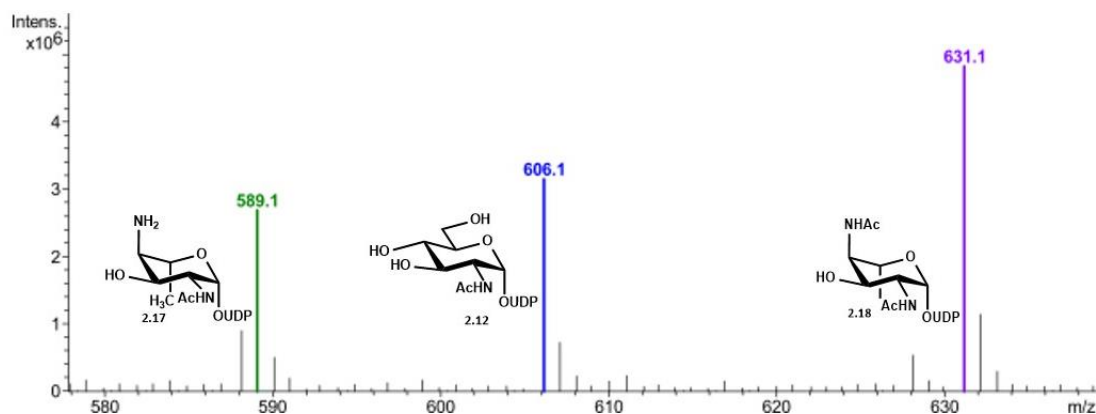
Following a literature procedure,<sup>297</sup> the *N*-acetyl cysteamine thioacetate (**5.9**) was successfully synthesised *via* acetyl transfer from acetic anhydride (**5.16**) to cysteamine HCl (**5.18**) (**Scheme 5.10**). Acetic anhydride (3 equivs) was added dropwise to an aqueous solution of cysteamine HCl (1 equiv) at pH 8.0, 0 °C. Following addition, the reaction was modified to pH 7.0 and stirred on ice for 2 hrs when the reaction was deemed to reach completion, as determined by TLC (**Scheme 5.10**). Spectroscopic analysis (NMR and MS) of this reaction after work-up confirmed formation of the desired product (**5.9**) and demonstrated that no further purification was required.

To ascertain whether the *N*-acetyl cysteamine thioacetate (**5.9**) was capable of acting as a co-factor substrate in PseH catalysed acetylation, “one-pot” synthesis of Pse5Ac7Ac (**1.13**) was attempted on a 0.1 mg scale. First, controls were set up by addition of UDP-GlcNAc (**3.1**) to either PseB, PseC, PseG, PseI (and their co-factors) and PseH (with no co-factor) (**Figure 5.12a**) or PseB, PseC, PseG, PseI (and their co-factors) and 100 mM *N*-acetyl cysteamine thioester (**5.9**) (with no PseH) (**Figure 5.12b**). Both reactions resulted in enzyme-catalysed conversion to the PseC product (**3.8**) but no production of the acetylated sugar intermediate (**3.9**) or any of the further intermediates in the Pse5Ac7Ac (**1.13**) biosynthetic pathway. Hence confirming that none of the molecules in the enzymatic reaction mixture can act as alternate PseH co-factors (**Figure 5.12a**) and that *N*-acetyl cysteamine thioester (**5.9**) cannot acetylate UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) without the PseH enzyme under these conditions (**Figure 5.12b**).



**Figure 5.12** LC-MS negative ESI analysis of the PseH controls demonstrates that the reaction halts at the PseC product intermediate (**3.8**) in the one-pot, enzymatic synthesis of Pse5Ac7Ac (**1.13**) containing **a**) PseB, PseC, PseG, PseI (their co-factors) and PseH or **b**) PseB, PseC, PseG, PseI (their co-factors) and *N*-acetyl cysteamine thioester (**5.9**).

Pleasingly upon incubation (shaking 120 rpm, 37 °C) of PseB (25  $\mu$ M), PseC (125  $\mu$ M), (and their co-factors), in addition to PseH (50  $\mu$ M) and 10 mM *N*-acetyl cysteamine thioacetate (**5.9**), the reaction displayed conversion to the acetylated sugar product (**3.9**) (**Figure 5.13**). Upon comparison of the intensity of the product associated with the PseH product ( $[M-H]^-$  631), a conversion of 45 % after 4 hrs was observed.

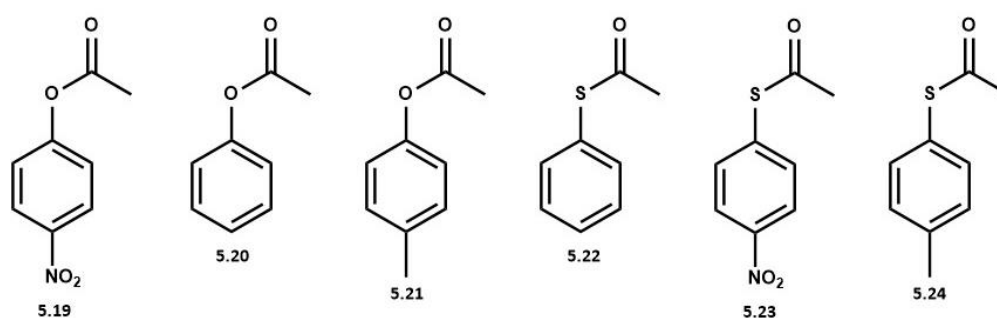


**Figure 5.13** LC-MS negative ESI analysis after 4 hrs, demonstrating conversion to the PseH product (**3.9**) using *N*-acetylcysteamine thioacetate (**5.9**) as a PseH co-factor substitute in a reaction consisting of UDP-GlcNAc (1 mM), PLP (1.5 mM), *L*-glutamate (10 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M) and PseH (50  $\mu$ M) in sodium phosphate buffer (50 mM, pH 7.4).

#### 5.2.4 Other alternate PseH co-factor substitutes

Following the positive results investigating co-factor substitutes with *N*-acetylcysteamine thioacetate (5.9), the promiscuity of the *C. jejuni* PseH (acetyl)-CoA (5.1/5.2) binding site was further probed. The acceptance of substitute acetyl donors varies greatly between different acetyltransferase enzymes and structural modifications distal to the thioacetate group can have great effects on the suitability as a co-factor. For example, this work has shown that *C. jejuni* PseH will accept *N*-acetylcysteamine thioacetate (5.9) as a co-factor substitute but not acetylthiocholine iodide (5.5). They both contain an ethane-thioacetate moiety which mimics part of the natural co-factor structure and has been shown to have favourable interactions with the PseH co-factor binding site. However the functionality at the other end of the molecules differs from an amide in *N*-acetylcysteamine thioacetate (5.9) to a tri-methylated positive amine group in acetylthiocholine iodide (5.5). Therefore it was postulated that this moiety may prevent acetylthiocholine iodide (5.5) from binding to PseH, either through steric bulk, charge repulsion or the change in hydrophilicity at the amine with the surrounding residues.

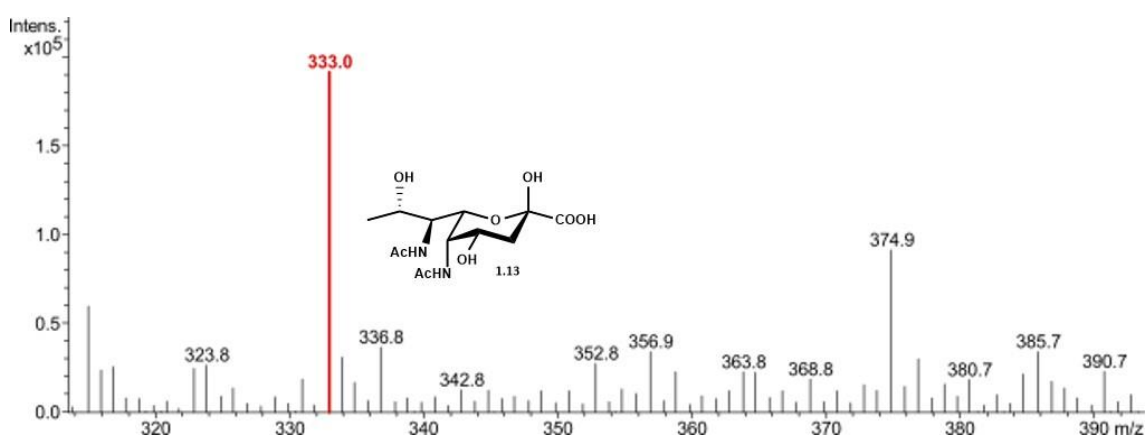
Additionally, some acetyltransferases have been shown to have a considerable flexibility in the structure of the acetyl-donor that they can utilise. Of particular interest was the identification that *p*-nitrophenyl acetate (5.19) can be employed as an acetyl donor in place of acetyl-CoA (5.1) during kinetic studies, to measure the rate of enzyme catalysed acetyl transfer.<sup>286, 298-301</sup> In order to investigate the acetyl donor promiscuity of *C. jejuni* PseH, its ability to catalyse acetyl-transfer reactions utilising phenolacetate and phenolthioacetate derivatives (Figure 5.15) in the absence of acetyl-CoA (5.1) was also investigated.



**Figure 5.15** Phenolacetate (5.19-5.21) and phenolthioacetate (5.22-5.24) structures proposed as putative acetyl-transferase co-factor alternatives.

Prior to incubation within the enzymatic reaction (for synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) or of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**)), solubility tests of the phenolacetate and phenolthioacetate derivatives were carried out under aqueous conditions. Three of the compounds; phenyl acetate (**5.20**), *p*-tolyl acetate (**5.21**) and *S*-phenyl thioacetate (**5.22**) were soluble at the stock (100 mM) and reaction (10 mM) concentration in sodium phosphate buffer (50 mM pH 7.4). However it was found that *p*-nitrophenyl acetate (**5.19**), *S*-(4-nitrophenyl)thioacetate (**5.23**), and *S*-(*p*-tolyl)thioacetate (**5.24**) required a solution of 10 % DMSO in sodium phosphate buffer (50 mM pH 7.4) to be soluble at the reaction concentration (10 mM).

A 5X reaction mixture stock was assembled for all of the following reactions containing UDP-GlcNAc (5 mM), PLP (7.5 mM) and L-glutamate (50 mM) in sodium phosphate buffer (50 mM pH 7.4). After addition of any other components, the required volume of sodium phosphate buffer (50 mM pH 7.4) was added to dilute the stock to the 1X reaction mixture previously used. An initial enzymatic reaction was set up for the synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**) from UDP-GlcNAc (**3.1**) to investigate the potential effect of DMSO in the reaction mixture. To an aliquot of the 5X stock, acetyl-CoA (1.5 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M), and PseI (25  $\mu$ M) were added and DMSO included to a final concentration of 10 % in sodium phosphate buffer (50 mM pH 7.4). The aliquot was incubated 37 °C, shaking (120 rpm) and monitored after 4 hrs with negative ESI LC-MS. Analysis showed conversion to the Pse5Ac7Ac (**1.13**) product, suggesting that the inclusion of 10 % DMSO did not affect the enzymatic activity (**Figure 5.16**).



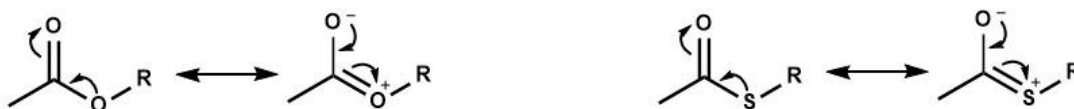
**Figure 5.16** Enzymatic synthesis of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**3.1**) in 10 % DMSO, sodium phosphate buffer (50 mM, pH 7.4) containing UDP-GlcNAc (1 mM), PLP (1 mM), L-glutamate (10 mM), acetyl-CoA (1.5 mM), PEP (1.5 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M) and PseI (25  $\mu$ M).

Negative controls for each of the potential substitute co-factors (**5.19-5.24**) were constructed whereby the 5X stock was incubated with PseB (25  $\mu\text{M}$ ), PseC (125  $\mu\text{M}$ ) and 10 mM acetyl-donor (**5.9** or **5.19-5.24**) in sodium phosphate buffer (50 mM pH 7.4) or 10 % DMSO in sodium phosphate buffer (50 mM pH 7.4) as required. To monitor PseH catalysed acetyl transfer the same reactions as above were assembled but with the addition of PseH (50  $\mu\text{M}$ ).

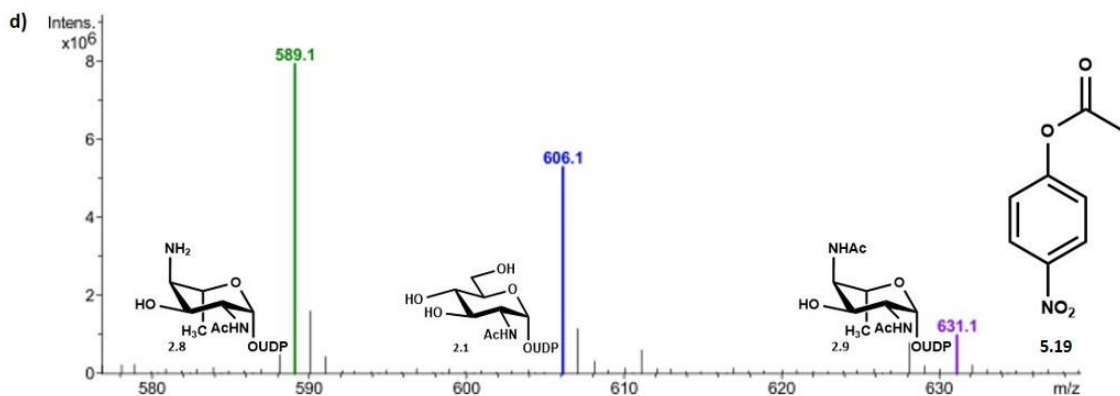
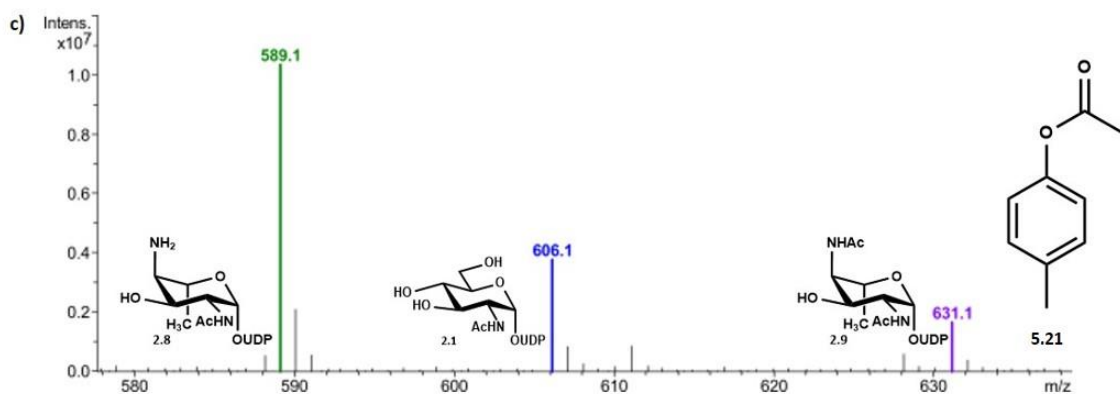
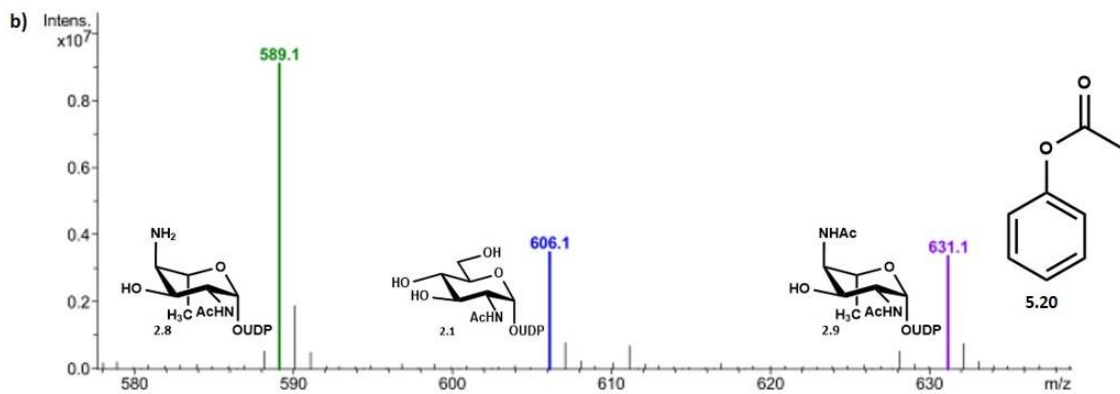
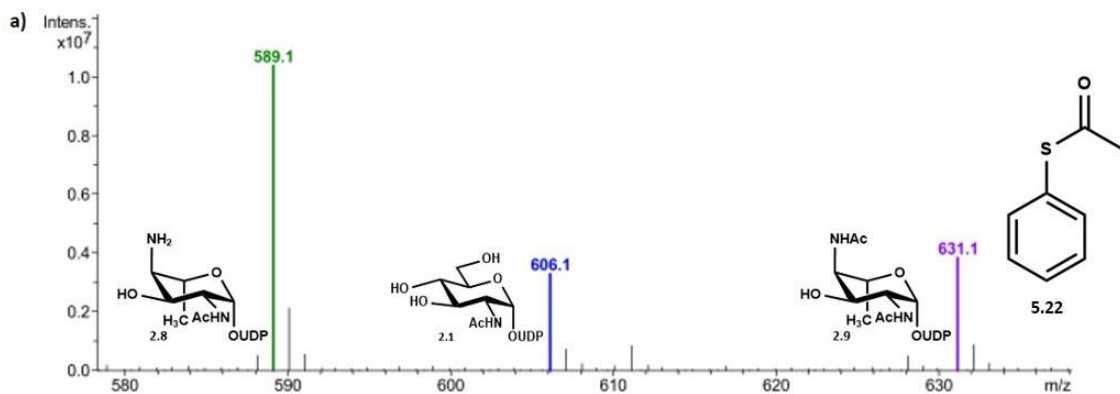
After 3 hrs of incubation (120 rpm, 37  $^{\circ}\text{C}$ ), negative ESI LC-MS analysis was used to monitor conversion to the acetylated product (**3.9**) *via* inspection of the  $[\text{M}-\text{H}]^{-}$  631 peak. In all reactions without PseH there was no observation of turnover to the PseH product, suggesting that the acetyl donors (**5.9**, **5.19-5.24**) are not reactive enough to undergo non-enzymatic acetyl transfer to the sugar amine (**3.8**) (**Appendix 8**).

All of the reaction mixtures containing phenolacetate and phenolthioacetate co-factors demonstrated PseH catalysed turnover to the PseH product, with a  $[\text{M}-\text{H}]^{-}$  631 peak observed in the negative ESI LC-MS traces. Comparison of the relative intensity of the  $[\text{M}-\text{H}]^{-}$  631 peak in these reactions showed that conversion was highest with *S*-phenyl thioacetate (**5.22**) and phenyl acetate (**5.20**) 22 % and 21 % respectively. Conversion to the PseH product (**3.9**) was found to decrease to 10 % with *p*-tolyl acetate (**5.21**) as the co-factor and just 7 % with *p*-nitrophenyl acetate (**5.19**). (**Figure 17**). Less than 5 % conversion to the PseH product (**3.9**) was detected in reactions containing *S*-(4-nitrophenyl)thioacetate (**5.23**) and *S*-(*p*-tolyl)thioacetate (**5.24**); 5 % and 3 % respectively, suggesting that PseH was not able to utilise these molecules as effective acetyl donors.

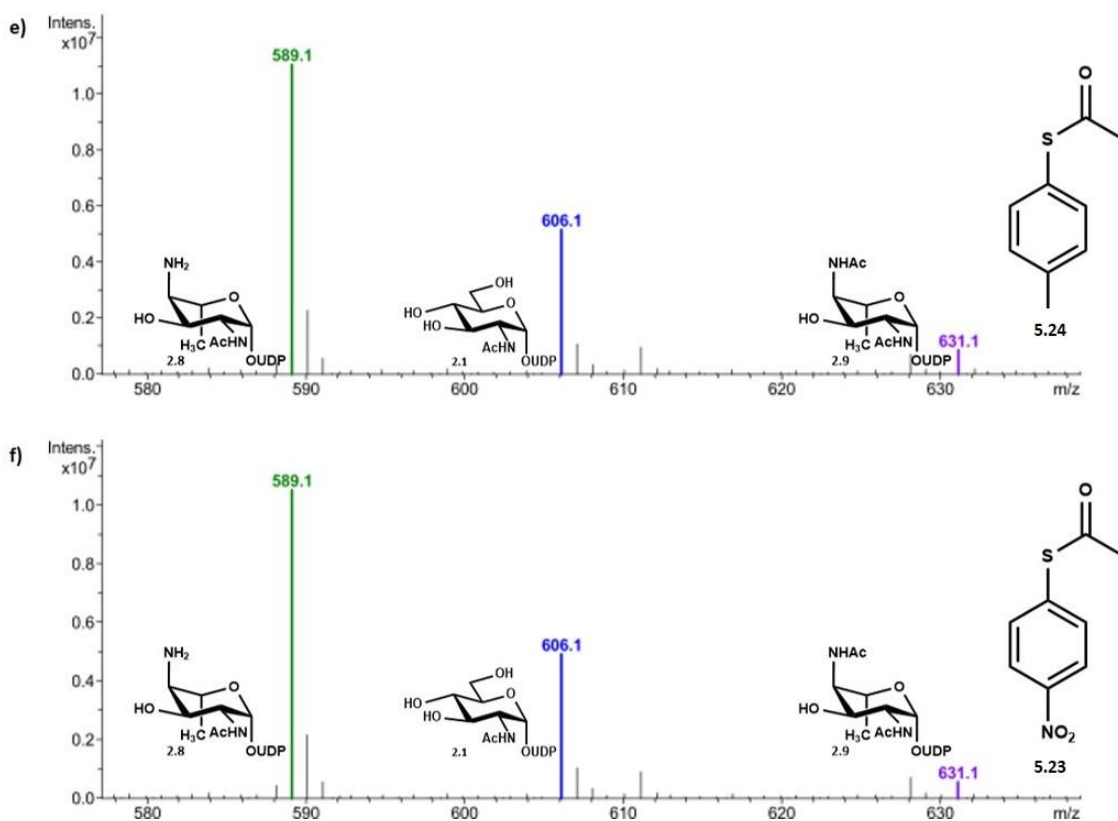
These results were somewhat unexpected when considering the reactivity of thioesters compared to esters and the electronic effects of the *para*-groups. Thioesters are more reactive than esters as there is less orbital overlap between the  $3p$  orbitals in sulphur and  $2p$  carbon orbitals than the  $2p$  oxygen and  $2p$  carbon orbitals. Therefore the sulphur lone pair is less delocalised compared to the oxygen lone pair and hence it less readily stabilised by resonance and the carbon in a thioester carbonyl is more electropositive than the carbon in an ester carbonyl (**Scheme 5.11**).<sup>302</sup>



**Scheme 5.11** The resonance forms of esters and thioesters.







**Figure 5.17** Negative ESI LC-MS investigating the ability of different PseH co-factors to promote conversion from UDP-GlcNAc (**3.1**) to the PseH product (**3.9**) in an enzymatic reaction; UDP-GlcNAc (1 mM), PLP (1.5 mM), L-glu (10 mM), PseB (25 M), PseC (125 M) and PseH (50 M) **a**) 10 mM *S*-phenyl thioacetate (**5.22**), **b**) 10 mM phenyl acetate (**5.20**), **c**) 10 mM *p*-tolyl acetate (**5.21**), **d**) 10 mM *p*-nitrophenyl acetate (**5.19**), **e**) 10 mM *S*-(*p*-tolyl)thioacetate (**5.24**) and **f**) 10 mM *S*-(4-nitrophenyl)thioacetate (**5.24**).

The expected increase in reactivity was observed to some extent when comparing reactivity of reactions containing *S*-phenyl thioacetate (**5.22**) and phenyl acetate (**5.20**). The reaction containing the thioester compound (**5.22**) showed a marginally higher conversion to the acetylated product, 22 %, compared to 21 % for the reaction containing phenyl acetate (**5.20**). However the opposite was found when comparing *S*-(*p*-tolyl)thioacetate (**5.24**) and *p*-tolyl acetate (**5.21**) with the former displaying almost no turnover (5 %) to the acetylated product (**2.9**) and the latter showing a small amount of PseH catalysed acetyl donation (10 % conversion). It was proposed that the difference in expected conversion could be due to the difference in solubility rather than the reactivity of the molecule. The thioester containing compound (**5.24**) required addition of DMSO to aid in solubility and solvation with this solvent may have prevented *S*-(*p*-tolyl)thioacetate (**5.24**) from being able to access the co-factor binding site.

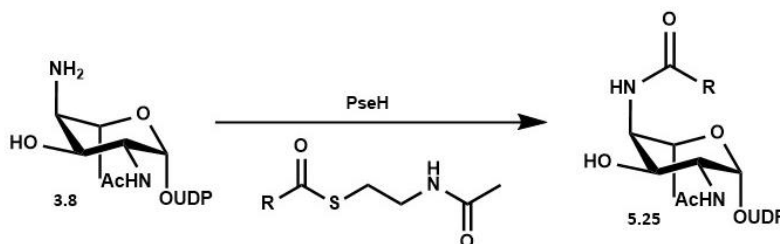
Upon comparing the different para-substituents it was expected that the nitro-containing compounds would be more efficient acetyl donors than their non-substituted counterparts and that the methyl derivatives would be less so. Evidence for the latter was demonstrated when comparing the relative conversions of phenyl acetate (**5.20**) and *p*-tolyl acetate (**5.19**), 21 % and 10 % respectively and, *S*-phenyl thioacetate (**5.22**) and *S*-(*p*-tolyl)thioacetate (**5.24**), 22 % and 5 % respectively. The positive inductive effect of the methyl group added electron density towards the benzene ring, resulting in a less electrophilic carbonyl carbon hence decreasing acetyl donation. Substitution with the *p*-methyl group appeared to attenuate acetyl transfer to a greater degree in the thioester compared to the ester. However solubility and solvation may also have been a factor in the thiophenyl acetates as *S*-(*p*-tolyl)thioacetate (**5.24**) required addition of DMSO to be soluble. It was surprising to observe that the electron-withdrawing nitro-substituted compounds demonstrated a reduced propensity for acetyl-transfer compared to their *p*-methyl substituted and non-substituted counterparts. It was proposed that, similarly to acetylthiocholine iodide (**5.5**), the steric bulk or the charge of the nitro group could be causing unfavourable interactions between the compound and the PseH co-factor binding site, and hence account for the reduced conversion to the PseH product (**3.9**) despite the favourable electronic effects.

*C. jejuni* PseH can readily accept some phenolacetate and phenolthioacetate derivatives as acetyl donors and utilise them in the enzymatic synthesis of UDP-4-acetamido--4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**). Unfortunately the level of conversion could not be improved upon from those established with *N*-acetylcysteamine thioacetate (**5.9**) however the difference in conversion levels has provided insight into the PseH co-factor binding site and the variety of molecules that it can accept. Importantly it has highlighted the slow rate of turnover with *p*-nitrophenyl acetate (**5.19**) which could aid in reaction design for the kinetic analysis of *C. jejuni* PseH.

## 5.3 Manipulation of the acetylation step for the synthesis of pseudaminic acid C7 derivatives

### 5.3.1 Utilisation of the PseH co-factor substitutes to introduce other C7 functionality

A strategy for production of pseudaminic acid C7 derivatives was devised utilising the Pse5Ac7Ac (**1.13**) biosynthetic enzymes under the optimised conditions previously discussed. It was decided to exploit the co-factor promiscuity of PseH to allow for introduction of a C4 *N*-acyl group during the synthesis of an acylated PseH product (**5.25**). Due to its cost and structural complexity it was undesired to attempt synthesis of acyl-CoA derivatives, so derivatives of the co-factor substitutes were considered. Derivatives of *N*-acetylcysteamine thioacetate (**5.9**) were focussed on due to their proposed facile synthesis and efficiency as acyl-donors when used with PseH (**Scheme 5.12**). In order to potentially allow for production of pseudaminic acids with manufactured C7 functionality in a one-pot chemo-enzymatic synthesis, without the need for expensive co-factors or time costly purification steps.



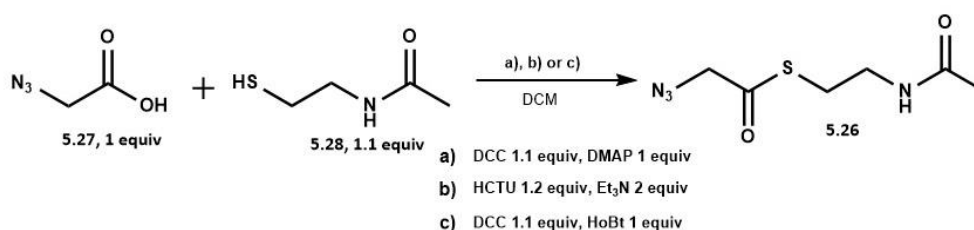
**Scheme 5.12** Introduction of desired C7 functionality in pseudaminic acids utilising a co-factor substitute.

Introduction of a bioorthogonal tag to Pse5Ac7Ac (**1.13**) allows for a handle for detection, purification and quantification of Pse5Ac7Ac (**1.13**) if desired. Based on observation of the active site structures of the biosynthetic enzymes, and considering the chemistry performed during the synthesis, it was hypothesised that small modifications to the C7 acetamido group may be accepted and remain in the sugar during enzymatic synthesis of Pse5Ac7Ac (**1.13**). Importantly a previously synthesised Pse5Ac7Az precursor (**5.14**) was shown to be taken up by *C. jejuni* and was accepted by the downstream biological machinery to display Pse5Ac7Az (**5.10**) on the flagellin.

It was proposed that an improved one-pot synthetic route to access Pse5Ac7Az (**5.10**) would aid in its use for labelling other Pse5Ac7Ac (**1.13**) containing bacterial surface structures. Therefore initial investigations focussed on the enzymatic synthesis of Pse5Ac7Az (**5.10**) with introduction of the azido group during the PseH catalysed acyl-transfer from *N*-acetylcysteamine thioazidoacetate (**5.26**).

### 5.3.2 Synthesis of *N*-acetylcysteamine thioazidoacetate

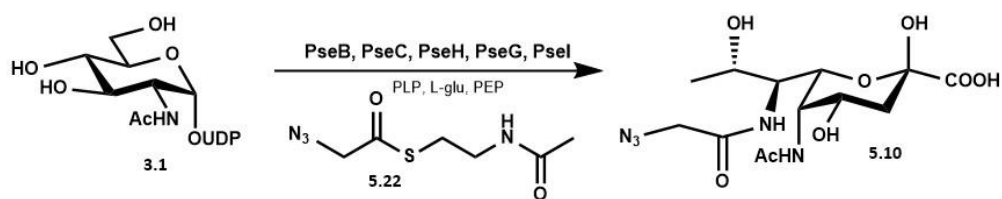
Three protocols were investigated for synthesis of *N*-acetylcysteamine thioazidoacetate (**5.26**) utilising amide synthesis strategies to form a thioester between azidoacetic acid (**5.27**) and *N*-acetylcysteamine (**5.28**). Each employed a combination of a coupling reagent to react with the azido acetic acid (DCC or HCTU) and further increasing its reactivity with the thiol, promoting synthesis of the thio-ester.



**Scheme 5.13** Three synthetic strategies for coupling azidoacetic acid and *N*-acetylcysteamine HCl to form *N*-acetylcysteamine thioazidoacetate (**5.26**).

DCC (0.84 mmol) (**a**, **c**) or HCTU (0.924 mmol) (**b**) and azidoacetic acid (0.76 mmol) were suspended in DCM and stirred on ice for five minutes. DMAP (0.76 mmol) (**a**), Et<sub>3</sub>N (1.54 mmol) (**b**) or HoBt (0.76 mmol) (**c**) and *N*-acetylcysteamine (0.84 mmol) were added and the reaction left to warm to room temperature overnight (**Scheme 5.13**). In reactions **a** and **b**, TLC analysis displayed depletion of the reagents after 16 hrs and formation of a spot with a *R<sub>f</sub>* of 0.28 (similar to that of the literature for *N*-acetylcysteamine thioazidoacetate (**5.26**)). However a number of other products were also present in the mixture suggesting that although there was no azidoacetic acid (**5.27**), the reaction may still contain some intermediate species. Reaction **c** appeared to react more slowly, with reagents still present after 72 hrs and hence this reaction was discontinued. Following column chromatography of **a** and **b**, a yield of 17 % product was achieved in crude **a** and 55 % gained from crude **b**. NMR analysis of combined *N*-acetylcysteamine thioazidoacetate (**5.26**) from both synthetic strategies (**a**, **b**) revealed peaks corresponding to the literature values for the desired product but also impurities (**Appendix 9**). However it was deemed unnecessary to purify *N*-acetylcysteamine thioazidoacetate (**5.26**) further for preliminary results as none of the remaining impurities were the aromatic structures and it was proposed that impurities of the other solutes would not affect the reaction.

### 5.3.3 Utilisation of *N*-acetylcysteamine thioazidoacetate in the enzymatic synthesis of a Pse5Ac7Az

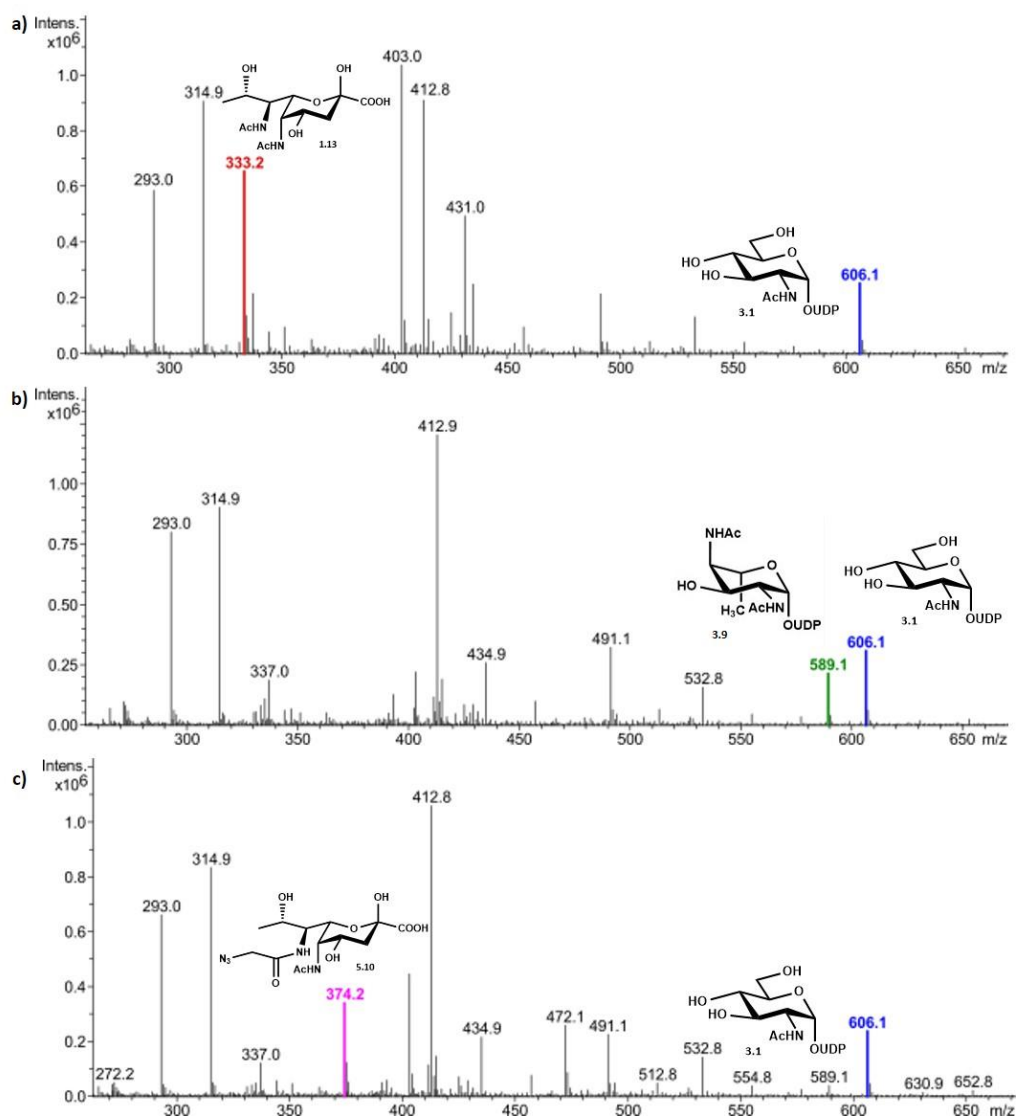


**Scheme 5.14** One-pot enzymatic route to Pse5Ac7Az (**5.10**), utilising the *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic enzymes and an azido-tagged PseH co-factor substitute (**5.26**).

A reaction mixture composed of UDP-GlcNAc (1 mM), PLP (1 mM), L-Glu (10 mM), PEP (1.5 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M), and PseI (25  $\mu$ M) in sodium phosphate (50 mM, pH 7.4) was assembled in order to analyse conversion to Pse5Ac (**1.13**) and Pse5Ac7az (**5.10**) using *N*-acetylcysteamine thioacyl derivatives. A positive control aliquot included the addition of *N*-acetylcysteamine thioacetate (10 mM) and the other aliquot included either *N*-acetylcysteamine thioazidoacetate (10 mM) or *N*-acetylcysteamine thioazidoacetate (10 mM) with CoA (0.015 mM) added. Negative ESI LC-MS analysis of this set of reactions displayed a number of additional *m/z* peaks in the mass spectra (discussed below) and hence it was deemed unsuitable to quantify the level of conversion. However analysis after 4 hrs displayed conversion to Pse5Ac7Ac (**1.13**) in the reaction containing *N*-acetylcysteamine thioacetate (10 mM), but with some remaining UDP-GlcNAc (**3.1**) also observed (**Figure 5.18a**). Only sugar peaks corresponding to UDP-GlcNAc (**3.1**) and the PseC product (**3.8**) were observed in the reaction mixture containing *N*-acetylcysteamine thioazidoacetate (10 mM) suggesting that it could not be utilised by PseH as a substitute co-factor (**Figure 5.18b**). However in the negative ESI LC-MS analysis of the reaction containing *N*-acetylcysteamine thioazidoacetate (10 mM) and CoA (0.015 mM), a peak corresponding to the desired Pse5Ac7Az (**5.10**) product was observed that was not present in the other spectra.

From this preliminary data it can be tentatively proposed that Pse5Ac7Az (**5.10**) can be produced in a one-pot enzymatic synthesis with the acetylazido group transferred during the PseH catalysed reaction in the presence of CoA. It was speculated that the steric hindrance or charge of the azido group may be less favourable for PseH binding compared to the acetyl group in *N*-acetylcysteamine thioacetate (**5.9**) and hence it is less effective as a PseH co-factor substitute, explaining why Pse5Ac7Az (**5.10**) was not observed without the addition of CoA (**5.2**). The occurrence of a number of other significant peaks during analysis of these reactions was unprecedented and especially surprising in the synthesis of Pse5Ac7Ac (**1.13**) as the same stock of *N*-acetylcysteamine thioacetate (10 mM) was used in this reaction as in previous reaction for

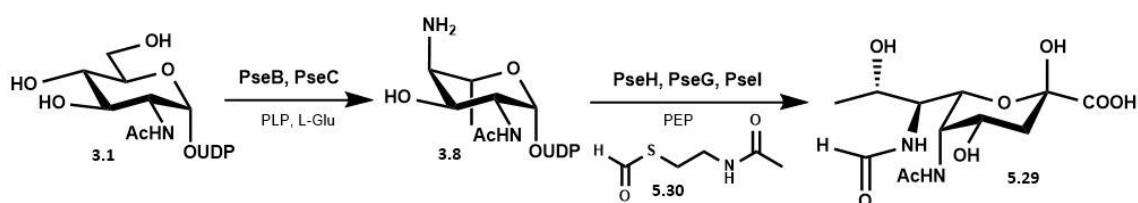
the synthesis of the PseH product (**3.9**). The additional peaks could not be assigned to any potential side reactions that were proposed and hence it was hypothesised that these molecules could have been contaminants from the LC-MS and noted that the quality of other data recorded had also decreased at this time. Additionally the chromatography column had demonstrated deterioration over the course of this project and could have been contaminated. In order to ascertain the presence of these peaks it was proposed to repeat this experiment after a full quality check of the equipment and on a large enough scale to allow for confirmation of production of Pse5Ac7Az (**5.10**) by NMR.



**Figure 5.18** LC-MS negative ESI analysis after 4 hrs, demonstrating conversion to **a)** Pse5Ac7Ac (**1.13**) using N-acetylcysteamine thioacetate (**5.9**) as a PseH co-factor substitute or **b)** Pse5Ac7Az (**5.10**) using N-acetylcysteamine thioazidoacetate (**5.22**) as a PseH co-factor substitute, **c)** Pse5Ac7Az (**5.10**) using N-acetylcysteamine thioazidoacetate (**5.22**) and 0.015 mM CoA (**5.2**) as a PseH co-factor substitute in a reaction consisting of UDP-GlcNAc (1 mM), PLP (1 mM), L-glutamate (10 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M), PseH (50  $\mu$ M), PseG (30  $\mu$ M) and PseI (25  $\mu$ M) in sodium phosphate buffer (50 mM, pH 7.4).

### 5.3.4 Extending the utilisation of *N*-acetylcysteamine thioacyl derivatives to access other pseudaminic acid derivatives

This method was further utilised in the Fascione lab (Matthew Best, MChem) in order to synthesise another prevalent pseudaminic acid; Pse5Ac7Fm (**5.29**). Following synthesis of *N*-acetylcysteamine thioformyl (**5.30**), Pse5Ac7Fm (**5.29**) production was attempted using the one-pot enzymatic methodology employed above but with inclusion of *N*-acetylcysteamine thioformyl (10 mM). Unfortunately conversion was heavily attenuated compared to synthesis of Pse5Ac7Ac (**1.13**) as *N*-acetylcysteamine thioformyl (**5.30**) appeared to react with a PseC co-factor. Therefore to obtain a higher yield, a two-step synthesis was employed (**Scheme 5.15**) whereby the PseB and PseC reaction with UDP-GlcNAc (**3.1**) was allowed to occur and then addition of PseH, PseG and PseI (and co-factors including *N*-acetylcysteamine thioformyl) produced Pse5Ac7Fm (**5.29**) from the PseC product (**3.8**).



**Scheme 5.15** Two-step enzymatic route to Pse5Ac7Fm (**5.29**), utilising the *C. jejuni* Pse5Ac7Ac (**1.13**) biosynthetic enzymes and a formyl PseH co-factor substitute (**5.30**).

## 5.4 Conclusions and future work

The requirement for the use of the acetyl-CoA (**5.1**) has been eliminated by the development of three strategies for the selective acetyl transfer to the amino group of UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**). These methods have been successfully used in the enzymatic production of Pse5Ac7Ac (**1.13**) from UDP-GlcNAc (**2.1**) elucidating an economically viable, efficient, one-pot synthesis.

Modifications to the previously published chemo-enzymatic synthesis of UDP-4-acetamido-4,6-dideoxy- $\beta$ -L-AltNAc (**3.9**), successfully avoided time costly purification steps and reduced the number of equivalents of catalyst required for synthesis. However there are two limitations that make this method unfavourable for the large scale synthesis of Pse5Ac7Ac (**1.13**) and derivatives compared to using acetyl-CoA (**5.1**). Firstly the ease of the onepot synthesis is lost and even though the reduction in cost compensates for this factor, it is still suboptimal. Additionally it was desired to develop a synthetic strategy that could be applicable to other pseudaminic acid derivatives, however the specificity for acetylation of the 4-amino group in this reaction relies on no other amino groups being present in the sugar structure.

Investigations regarding the *in situ* regeneration of the PseH co-factor using acetyl thiocholine iodide (**5.5**) and CoA (**5.2**) as reagents, re-established the one-pot synthesis of Pse5Ac7Ac (**1.13**) whilst drastically reducing the expense of this process compared to use of the natural co-factor (**5.1**). However to attain comparable turnover with catalytic amounts of CoA (**5.1**), high concentrations of the regeneration factor were required (**5.5**, 100 mM) which can be criticised for low atom economy and may invoke down-stream purification issues.

*N*-acetylcysteamine thioacetate (**5.9**), synthesised from readily available materials, was identified as an ideal direct replacement for acetyl-CoA (**5.1**) as a PseH co-factor and was successfully employed in acetyl-transfer to UDP-4-amino-4,6-dideoxy- $\beta$ -L-AltNAc (**3.8**). The added benefit of this system was in its predicted generalisation for the synthesis of other pseudaminic acid derivatives using *N*-acetylcysteamine thioacyl derivatives. Preliminary results utilising this system in this way resulted in production of Pse5Ac7Az (**5.10**) and Pse5Ac7Fm (**5.29**). However optimisation was required in both reactions; for the synthesis of Pse5Ac7Ac (**5.10**) CoA (**5.2**) was required in the reaction mixture and for the synthesis of Pse5Ac7Fm (**5.29**), the synthesis was carried out in two sequential steps. Despite these setbacks and potential need for optimisation with different functionalities, this strategy utilising *N*-acetylcysteamine thioacyl derivatives, provides a general method for the chemo-enzymatic production of pseudaminic acid C7 derivatives with significant advantages over existing syntheses.

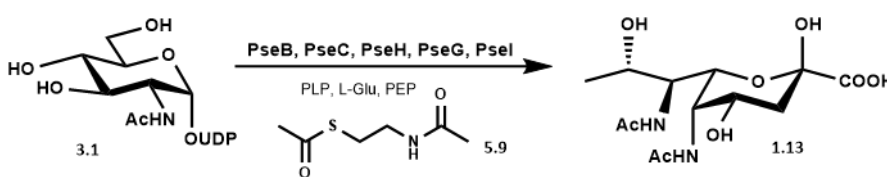


## **Chapter 6 Concluding remarks**

## 6.1 Conclusions and future work

Preliminary PA2794 characterisation to gain insight into the natural substrate unfortunately did not allow for the definitive allocation of the natural ligand. The crystal structure data was consistent with the previously published data,<sup>3</sup> suggestive that binding of a sugar with the Neu5Ac (**1.1**) C5 stereochemistry is unfavourable in the most energetically favourable conformation of the transition state structure (**Figure 2.13**). However non-enzyme electron density in the active site was attributed to solute molecules (**Figure 2.15**) rather than the Pse5Ac7Ac (**1.13**) ligand which could be attributed to; the higher concentration of solute molecules, the addition of Pse5Ac7Ac (**1.13**) being after crystallisation making the active site inaccessible, or a low binding affinity to free Pse5Ac7Ac (**1.13**). Future work could address such issues by establishing alternate PA2794 purification and crystallisation conditions with Pse5Ac7Ac (**1.13**) or by focussing on the synthesis of Pse5Ac7Ac (**1.13**) analogues that would have an expected higher affinity with this putative pseudaminidase. Further investigations could also include trialling different pseudaminic acid derivatives native to *Pseudomonas aeruginosa* such as Pse5hB7Fm.<sup>187</sup>

Due to the lack of availability of pseudaminic acid derivatives and chemical probes, full characterisation of any pseudaminic acid processing enzymes is yet to be reported. This project developed a strategy towards the synthesis of Pse5Ac7Ac (**1.13**) on a large scale in order to supply the quantity of material required for future studies (**Scheme 6.1**). This has elucidated a one-pot chemo-enzymatic pathway to Pse5Ac7Ac (**1.13**) that has eliminated any issues that previously occurred with the use of the enzymatic synthesis.



**Scheme 6.1** One-pot chemo-enzymatic pathway to Pse5Ac7Ac (**1.13**) utilising the five *C. jejuni* biosynthetic enzymes with substitution of the PseH co-factor acetyl-CoA (**5.1**) with *N*-acetyl cysteamine thioacetate (**5.9**).

The second step in the pathway is catalysed by PseC and this enzyme was observed to have solubility under the induction conditions used for optimal yields. Although this was successfully rectified by using different induction conditions, it is disappointing to lose yield of this enzyme in particular as it is needed at a higher concentration during Pse5Ac7Ac (**1.13**) synthesis than the other enzymes. This was discovered through monitoring conversion to the PseB and PseC

products in D<sub>2</sub>O, whereby overall conversion to the PseC product (**3.8**) was achieved by out-competing the secondary PseB reaction with increasing the concentration of PseC (**Figure 4.7**). Further exploration of the PseB mechanism could aid in design of mutants to prevent this secondary PseB catalysed epimerisation.<sup>242</sup> However it was also hypothesised that expressing the bound PseB and PseC enzymes together could also increase conversion to the PseC product as the catalyst for the second reaction (PseC) would be held in close proximity to the product of the first reaction (**3.7**). This could also have the additional benefit of promoting PseC solubility and allow for induction conditions to be utilised that allow for maximum enzymatic production.

The main limitation for using a purely enzymatic route for the synthesis of Pse5Ac7Ac (**1.13**) was the cost of the co-factor for the third step of the biosynthetic pathway; PseH catalysed acetyl transfer from acetyl-CoA (**5.1**).<sup>232</sup> Use of this co-factor was eliminated by the development of a strategy to generate acetyl-CoA (**5.1**) *in situ* via acetyl transfer by having high concentrations of a regeneration factor (acetylthiocholine iodide **5.5**) and sub-stoichiometric amounts of CoA (**5.2**). Additionally the necessity for acetyl-CoA was completely abolished when an alternate co-factor, *N*-acetylcysteamine thioacetate **5.9**, was successfully utilised in the one-pot synthesis of Pse5Ac7Ac (**1.13**) with comparable turnover (**Figure 5.13**). A significant advantage of this latter strategy was the ease of synthesis of other *N*-acetylcysteamine thioacyls allowing for the proposed applicability of these methods to introduce other desired C7 functionalities. However preliminary results with other *N*-acetylcysteamine thioacyl derivatives highlighted the need for considerations of the structural integrity of the PseH co-factor binding site and potential steric clashes with the acyl derivatives. Crystal structures in complex with *N*-acetylcysteamine thioacyl derivatives would provide further insight into the binding dynamics and interactions with the thioacyl groups in particular.

Development of this chemo-enzymatic synthesis of Pse5Ac7Ac (**1.13**) is primed for expansion to production on a large scale for the first time. After establishing purification techniques, further processing of this sugar, adapting methodology established for Neu5Ac (**1.1**), will produce the desired glycosides, inhibitors and inactivators. The characterisation of PA2794 can therefore then be re-visited, which it has not been done since its crystallisation in 2009,<sup>3</sup> and fully investigated with the predicted natural ligand. Knowledge of a synthetic route to pseudaminic acid chemical probes will surely facilitate the characterisation of other pseudaminic acid processing enzymes and promote investigation of a multitude of potential therapeutic targets.

## Chapter 7 Experimental

## 7.1 General methods

Unless otherwise specified, all chemical reagents were obtained from commercial sources and used without further purification. Pse5Ac7Ac was custom ordered from Sussex Research Laboratories Inc. (Ottawa, ON Canada) Recombinant plasmids were either kindly gifted (PseB, PseC Tanner Research Lab, University of British Columbia, Canada) or purchased from Genscript.

### *LB recipe;*

NaCl 1 % (w/v), Tryptone 1 % (w/v), Yeast 0.5 % (w/v) in dH<sub>2</sub>O and autoclave.

### *LB agar recipe;*

NaCl 1 % (w/v), Tryptone 1 % (w/v), Yeast 0.5 % (w/v), Agar 1.5 % (w/v) in dH<sub>2</sub>O and autoclave.

### *TYP recipe;*

Tryptone 1.6 % (w/v), Yeast extract 2.4 % (w/v), NaCl 0.5 % (w/v), K<sub>2</sub>HPO<sub>4</sub> 1 % (w/v) in dH<sub>2</sub>O and autoclave.

### *SOC recipe;*

Tryptone 2 % (w/v), Yeast extract 0.5 % (w/v), KCl 2.5 mM, NaCl 10 mM, MgCl<sub>2</sub> 10 mM, MgSO<sub>4</sub> 10 mM, Glucose 20 mM in dH<sub>2</sub>O and autoclave.

### *Electrocompetent E. coli BL21 DE3 cells;*

A single *E. coli* BL21 DE3 culture was inoculated in LB (10 mL) 225 rpm, 37 °C, overnight. 5 mL of culture was added to 1 L LB in a 2L baffled flask and incubated 180 rpm, 37 °C, until an OD<sub>600</sub> of 0.6 and placed immediately on ice. The cells were pelleted (1000 xg, 20 mins, 4 °C in pre-cooled tubes), supernatant removed and the cell pellets resuspended in sterile ice cold dH<sub>2</sub>O (800 mL). Cells were pelleted (1000 xg, 20 mins, 4 °C) with the dH<sub>2</sub>O being discarded and the cell pellets resuspended in sterile ice cold dH<sub>2</sub>O (400 mL). Cells were pelleted again (1000 xg, 20 mins, 4 °C) with the dH<sub>2</sub>O being discarded and the cell pellets resuspended in sterile ice cold 10 % glycerol (v/v) dH<sub>2</sub>O (40 mL). Cells were pelleted once more (1000 xg, 20 mins, 4 °C), supernatant removed and pellets resuspended in 10 % glycerol (v/v) dH<sub>2</sub>O (1 mL), aliquoted into 50 µL portions, flash frozen in liquid nitrogen and stored -80 °C.

### *Chemically competent E. coli BL21 DE3 cells;*

A single *E. coli* BL21 DE3 culture was inoculated in LB (10 mL) 225 rpm, 37 °C, overnight. 5 mL of culture was added to 1 L LB in a 2L baffled flask and incubated 180 rpm, 37 °C, until an OD<sub>600</sub> of 0.3-0.4 and placed immediately on ice for 20 minutes, swirling occasionally. The cells were pelleted (3000 xg, 15 mins, 4 °C in pre-cooled tubes), supernatant removed and the cell pellets resuspended in sterile ice cold 100 mM MgCl<sub>2</sub> (100 mL). Cells were pelleted (2000 xg, 15 mins,

4 °C) with the supernatant discarded and the cell pellets resuspended in sterile ice cold 100 mM CaCl<sub>2</sub> (200 mL) and left on ice for 20 minutes. Cells were pelleted again (2000 xg, 15 mins, 4 °C) with the supernatant discarded and the cell pellets resuspended in sterile ice cold 85 mM CaCl<sub>2</sub> 15 % glycerol (v/v) dH<sub>2</sub>O (40 mL). Cells were pelleted once more (1000 xg, 15 mins, 4 °C), supernatant removed and pellets resuspended in 85 mM CaCl<sub>2</sub> 15 % glycerol (v/v) dH<sub>2</sub>O (1 mL), aliquoted into 50 µL portions, flash frozen in liquid nitrogen and stored -80 °C.

#### *Electrotransformation;*

Electrocompetent cells (50 µL) and the plasmid (100 ng µL<sup>-1</sup>) were thawed on ice (20 mins) and an electroporation cuvette (0.2 cm) was placed on ice. 1 µL plasmid was added to the cells, gently mixed and placed on ice for 20 mins. This mixture was transferred to the electroporation cuvette and electroporated (BIORAD GENEPULSER II). TYP media (1 mL) was immediately added to the electroporation cuvette and incubated 180 rpm, 37 °C, 1 hour. Cells were centrifuged (10, 000 xg, 30 seconds), 600 µL of the supernatant was discarded and the pellet resuspended in the remaining media. Cells were streaked onto a LB agar plate containing the appropriate antibiotic resistance and incubated 37 °C, overnight then stored 6 °C.

#### *Heat shock;*

Chemically competent cells (50 µL) and the plasmid (100 ng µL<sup>-1</sup>) were thawed on ice (20 mins). 2 µL plasmid was added to the cells, gently mixed and placed on ice for 15 mins. This mixture was placed in a 42 °C water bath for 60 seconds then immediately placed on ice for 120 seconds. SOC media (1 mL) was added to the cells and incubated 180 rpm, 37 °C, 1 hour. Cells were centrifuged (10, 000 xg, 30 seconds), 600 µL of the supernatant was discarded and the pellet resuspended in the remaining media. Cells were streaked onto a LB agar plate containing the appropriate antibiotic resistance and incubated 37 °C, overnight then stored 6 °C.

#### *Glycerol stocks;*

LB (15 mL) containing the appropriate antibiotic resistance was inoculated with a single colony from a LB agar plate, and incubated 180 rpm, 37 °C, overnight. A 500 µL aliquot was mixed with 500 µL glycerol 50 % (v/v), flash frozen in liquid N<sub>2</sub> and immediately stored at -80 °C.

#### *Protein storage;*

Proteins were stored in their dialysis buffers at 6 °C for up to a week. Longer storage involved concentration to 1 mg mL<sup>-1</sup> and division into 1 mL aliquots, flash freezing in liquid N<sub>2</sub> and immediately storing at -80 °C.

*Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE);*

5 x SDS reducing sample buffer recipe;

SDS 10 % (w/v), Glycerol 20 % (w/v), Bromophenol blue 0.05 % (w/v), Tris-HCl 200 mM pH 6.8,  $\beta$ -mercaptoethanol 10 mM, in dH<sub>2</sub>O.

SDS resolving gel buffer;

SDS 0.4 % (w/v), Tris-HCl 1.5 M pH 8.8, in dH<sub>2</sub>O.

SDS stacking gel buffer;

SDS 0.4 % (w/v), Tris-HCl 0.5 M pH 6.8, in dH<sub>2</sub>O.

SDS running buffer 4X;

Glycine 160 mM, Tris 0.1 M

Staining solution;

0.1 % Coomassie Brilliant Blue R-250, 50 % MeOH, 10 % glacial acetic acid

Destaining solution

50 % MeOH, 10 % glacial acetic acid

12 % acrylamide gels were used throughout and made up as follows;

2.5 mL resolving buffer, 4.2 mL acrylamide and 3.2 mL dH<sub>2</sub>O were combined and gently inverted.

50  $\mu$ L APS 20 % (w/v) was added and the solution inverted, followed by addition of 16  $\mu$ L TEMED, a further inversion and pipetted into the gel plates. Once set, 1.3 mL stacking buffer, 0.5 mL acrylamide and 3.2 mL dH<sub>2</sub>O were combined and gently inverted. 25  $\mu$ L APS 20 % (w/v) was added and the solution inverted, followed by addition of 8  $\mu$ L TEMED a further inversion and pipetted onto the resolving gel and ladders placed in to form wells.

Protein samples contained 3  $\mu$ L 5 x SDS reducing sample buffer and 12  $\mu$ L sample, held at 95 °C for 5 minutes, 10  $\mu$ L of sample was loaded per well. Gels were subjected to electrophoresis for 30 mins (Flowgen Consort E734), then stained and destained.

*DNA gel;*

0.35 g of agarose was added to 50 mL TAE 1X and boiled (in the microwave, 3 x 1 minute), then set aside to cool. 5  $\mu$ L Sybasafe was added, the solution was poured into the gel mould and a comb added.

DNA samples consisted of 2  $\mu$ L 6x gel loading dye and 10  $\mu$ L sample, and 10  $\mu$ L of sample was added per well. Gels were subjected to electrophoresis for 60 mins (Flowgen Consort E734), then imaged (GelDoc).

*UV-vis;*

Bacterial growth curves were derived from measuring absorbance at a wavelength of 600 nm conducted on a Ultra-violet U1900 (Hitachi) spectrometer utilising UV solutions 2.2 software (Hitachi). Protein concentrations were calculated (**equation 1**) from absorbance measured at a wavelength of 280 nm on a nanodrop (Denovix DS-II FX Spectrometer/Fluorometer).

$$\text{Conc (M)} = \frac{A}{\varepsilon(\text{dm mol}^{-1} \text{cm}^{-1})l(\text{cm})} \quad \text{equation 1}$$

*Liquid Chromatography Mass Spectrometry (LC-MS);*

All mass spectrometry was carried out using negative ion mode electrospray ionisation, unless otherwise detailed, on a Bruker HCTultra ETD II system (Bruker Daltonics) mass spectrometer in The University of York Centre of Excellence in Mass Spectrometry (CoEMS). A Waters C18 column was fitted to a high performance Dionex UltiMate® 3000 LC system (ThermoScientific) equipped with an UltiMate® 3000 Diode Array Detector. Chromeleon® 6.80 SR12 software (ThermoScientific) combined with esquireControl version 6.2, Build 62.24 software (Bruker Daltonics), and Bruker compass Hystar 3.2-SR2, Hystar version 3.2, Build 44 software (Bruker Daltonics).

General procedure;

Solvent A - Water, Formic Acid 0.1 % (v/v)

Solvent B - Acetonitrile, Formic Acid 0.1 % (v/v)

The general method employed a flow-rate of 300  $\mu\text{L min}^{-1}$ , followed pre-equilibrium of the column in solvent A 95 % for 30 seconds followed by application of a linear gradient; solvent B 70 % to 30 % over 6 minutes. The column was washed in 95 % solvent B for 1 minute before solvent A 95 % was applied for 1 minute to re-equilibrate the column.

During analysis of mass spectra a buffer peak was observed ( $\sim 1$  min) as well as a peak during the final wash stage and as such data from this part of the run was emitted from spectra.

*NMR;*

NMR data was collected on a Jeol ECS400 NMR Spectrometer or the Bruker AVIIIHD500 FT-NMR Spectrometer made available at The University of York Centre for Magnetic Resonance.



## 7.2 PA2794 and PA2794 F129A mutant

### PA2794 F129A mutation

*P. aeruginosa* PA2794 recombinant plasmid was designed and purchased from GenScript. The sequence was *E. coli* codon optimised and ligated into a pET15b plasmid between the NdeI and BamHI restriction sites resulting in a N-terminal His<sub>6</sub> tagged protein. Site directed mutagenesis was carried out on the PA2794 pET15b recombinant plasmid to produce the phenylalanine129alanine mutant. Primers were designed using snap gene to replace the TTC phenylalanine codon with GCC.

Forward primer; 5'- GGC GCG GAT TAC AAC GCC GCG CAC GGC AAG AGC -3'

Reverse primer; 5'- ACC ACG CGC CAG GGT GAA -3'

PCR reaction mixtures were composed according to the 5 x Phusion GC buffer manufacturer conditions (**Table 1**) and subjected to thermocycling conditions as follows;

98 °C	30 seconds	
98 °C	10 seconds	x 35
55, 58.4, 61.8, 65.2, 68.6, 72.0 °C	30 seconds	
72 °C	210 seconds	
72 °C	10 minutes	
4 °C	Hold	

**Table 1;** Components of the PCR reaction mixture

Component	Per 20 µL reaction / µL
5 x Phusion GC buffer	4
DNTPs (10 mM)	0.4
Forward primer (10 µM)	1
Reverse primer (10 µM)	1
PA2794 recombinant plasmid (20 ng µL <sup>-1</sup> )	0.2
DMSO	0.6
Phusion DNA polymerase	0.2
dH <sub>2</sub> O	12.6

The resulting DNA was transformed *via* heat shock into *E. coli* Top-10 cells and after incubation (37 °C, 220 rpm, 45 minutes) the cell culture was spread onto LB agar plates containing ampicillin (100 µg mL<sup>-1</sup>) and incubated 37 °C, overnight. A single colony was incubated in LB (10 mL) containing ampicillin (100 µg mL<sup>-1</sup>) and incubated 37 °C, 180 rpm, overnight. The cells were centrifuged (12,000 x g, 5 mins, table top centrifuge), the supernatant was discarded and cell pellet subjected to a miniprep (Qiagen, QIAprep Spin, following the manufacturer's guidelines) to extract the DNA plasmid for sequencing (GATC Sanger sequencing).

#### *PA2794 and PA2794 F129A transformation*

Following confirmation that the mutation had been successful, the plasmid containing the *PA2794 F129A* and the native *P. aeruginosa* PA2794 pEt15b recombinant plasmid were transformed *via* heat shock into chemically competent *E. coli* BL21(DE3) cells, which had been previously prepared. Following incubation (37 °C, 220 rpm, 45 minutes) the cell culture was spread onto LB agar plates containing ampicillin (100 µg mL<sup>-1</sup>) and incubated 37 °C, overnight.

#### *PA2794 and PA2794 F129A expression*

A single colony was inoculated in liquid LB media containing ampicillin (100 µg mL<sup>-1</sup>) (15 mL) and incubated 180 rpm, 37 °C overnight. 2 mL of culture was added per litre of media in baffled flasks for aerobic growth at 180 rpm, 37 °C until an OD<sub>600</sub> of 0.6 was reached. At this point, the cells were induced with 0.5 mM IPTG and proteins were expressed under a variety of conditions during trial experiments and then for 18 hours at 16 °C in all subsequent grow ups. Centrifugation (6000 x g, 40 mins, 6 °C, Beckman Avanti centrifuge J-25) after the appropriate time afforded cell pellets which were stored at -80 °C until required or used straight away.

#### *PA2794 and PA2794 F129A purification*

Lysis buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 10 mM imidazole Benzonase (25 U/L grow up), protease inhibitor tablet (used as instructed).

Equilibrium buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 10 mM imidazole.

Elution buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 500 mM imidazole.

Size exclusion buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol.

Cell pellets were defrosted on ice and resuspended in cold lysis buffer. The solution was sonicated on ice 20 x 30 second intervals (Soniprep 150), centrifuged (20, 000 x g, 20 min, 6 °C, Beckman Avanti centrifuge J-25) and the resulting supernatant filtered (0.22 µm).

The supernatant was loaded onto a HisTrap HP Ni<sup>2+</sup> affinity column (5 mL), pre-equilibrated with equilibrium buffer (AKTA Start, GE Healthcare Life Technologies). The same programme was applied in each case and eluent monitored by UV Abs<sub>280</sub>;

Equilibrium buffer for 7 column volumes, linear gradient from 10 mM imidazole to 300 mM imidazole over 15 column volumes, elution buffer for 5 column volumes.

Fractions containing protein were analysed by SDS PAGE and those containing purified over-expressed protein were combined and concentrated to a volume of 2 mL using 30 kDa MWCO falcon tubes. The resulting 2 mL was centrifuged (12,000 x g, 2 mins, table top centrifuge) and the supernatant applied to the size exclusion column (120 mL, HiLoad 16/600 S200, GEHealthcare) and fractions collected when an increase in UV Abs<sub>280</sub> was detected. Fractions containing protein were analysed by SDS PAGE and those containing purified over-expressed protein were combined and concentrated 3.5 mg mL<sup>-1</sup> and stored.

#### *PA2794 crystallography*

PA2794 48 well sitting drop trays were setup to screen conditions based on previously published crystallography conditions (0.1 M bicine pH 5.0, PEG 6K 10 % (w/v) or 0.1 M imidazole pH 8.0, PEG 8K 10 % (w/v)) from the stock solutions below;

1 M bicine pH 4.44, pH 5.01 or pH 5.53, adjusted with concentrated NaOH and HCl

1 M imidazole pH 7.51, pH 8.02, pH 8.50, adjusted with concentrated NaOH and HCl

PEG 6K 50 % (w/v) (stirred to resuspend)

PEG 8K 25 % (w/v) (stirred to resuspend)

Mosquito robot was utilised to dispense either 0.3 µL or 0.5 µL PA2795 (3.5 mg mL<sup>-1</sup>) or PA2794 (3.5 mg mL<sup>-1</sup>), Pse5Ac7Ac (1 mM) for co-crystallisation attempts, into each drop followed by 0.5 µL well solution. The tray was sealed and left at room temperature for crystals to develop for at least 10 days. Pse5Ac7Ac (**1.13**) was introduced to crystals either as a solid or 0.5 µL Pse5Ac7Ac (1 mM in well solution).

### Conditions for *apo* PA2794

PEG 6K	Bicine 0.1 M pH 4.5		Bicine 0.1 M pH 5.0		Bicine 0.1 M pH 5.5	
14%						
16%						
18%						
20%						
22%						
24%						
26%						
28%						
	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L

### Alternative conditions for ligand bound PA2794 attempts

PEG 8K	Imidazole 0.1 M pH 7.5		Imidazole 0.1 M pH 8.0		Imidazole 0.1 M pH 8.5	
8%						
10%						
12%						
14%						
16%						
18%						
20%						
22%						
	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L	0.3 $\mu$ L: 0.5 $\mu$ L	0.5 $\mu$ L: 0.5 $\mu$ L

Regular crystals were fished from the solution and placed in a cryo solution (well solution containing 20 % glycerol) for 1 minute then flash frozen in N<sub>2</sub> (l) and stored in N<sub>2</sub> (l) prior to diffraction testing. X-ray diffraction was carried out in-house using facilities in the YSBL; a Rigaku MicroMAX 007HF generator, RAXIS IV++ imaging plate detector and an Actor robotic sample changer.

Crystals that successfully diffracted were stored in in N<sub>2</sub> (l) and delivered to the Diamond synchrotron facilities. Data was processed using CCP4i2, refined using REFMAC 5 and Coot, and graphics developed on CCP4MG.

### 7.3 *Campylobacter jejuni* Pse5Ac7Ac biosynthetic enzymes

#### *Transformation*

A Gene Pulser II Electroporation system, consisting of a Gene Pulser II Unit and a Capacitance Expander II Unit, was employed for electroporation of the PseB and PseC recombinant plasmids into electrocompetent *E. coli* BL21(DE3) cells, which had been previously prepared. The PseH, PseG and PseI recombinant plasmids were introduced into chemically competent *E. coli* BL21(DE3) cells, which had been previously prepared according to literature procedures, *via* heat shock.

#### *Expression*

All bacterial growths were performed in LB media containing antibiotic; Kanamycin (50  $\mu\text{g mL}^{-1}$ ) (PseB) or Ampicillin (100  $\mu\text{g mL}^{-1}$ ) (PseC, PseH, PseG, PseI). Freshly transformed cells (or glycerol stocks) were streaked onto LB agar plates containing appropriate antibiotics and incubated 37 °C, overnight. A single colony was inoculated in liquid media (15 mL) and incubated 180 rpm, 37 °C overnight. 2 mL of culture was added per litre of media in baffled flasks for aerobic growth at 180 rpm, 37 °C until an OD<sub>600</sub> of 0.6 was reached. At this point, protein expression was induced with IPTG using optimised conditions as discussed. Centrifugation (6000 x g, 40 mins, 6 °C, Beckman Avanti centrifuge J-25) after the appropriate time afforded cell pellets which were stored at -80 °C until required.

#### *Purification*

Lysis buffer; 50 mM sodium phosphate buffer, pH7.4, 400 mM NaCl, 10 mM imidazole, Benzonase (25 U/L grow up), protease inhibitor tablet (used as instructed).

Equilibrium buffer; 50 mM sodium phosphate buffer, pH7.4, 400 mM NaCl, 10 mM imidazole.

Elution buffer; 50 mM sodium phosphate buffer, pH7.4, 400 mM NaCl, 500 mM imidazole.

Desalting buffer; 25 mM sodium phosphate buffer, pH7.4, 50 mM NaCl.

Cell pellets were defrosted on ice and resuspended in cold lysis buffer. The solution was sonicated on ice 20 x 30 second intervals (Soniprep 150), centrifuged (20,000 x g, 20 min, 6 °C, Beckman Avanti centrifuge J-25) and the resulting supernatant filtered (0.22  $\mu\text{m}$ ). The supernatant was loaded onto a HisTrap HP Ni<sup>2+</sup> affinity column (5 mL), pre-equilibrated with equilibrium buffer (AKTA Start, GE Healthcare Life Technologies). The following programme was applied in each case and the eluent monitored by UV Abs<sub>280</sub>;

Equilibrium buffer for 7 column volumes, linear gradient from 10 mM imidazole to 300 mM imidazole over 15 column volumes, elution buffer for 5 column volumes.

Fractions containing protein were analysed by SDS PAGE and those containing purified over-expressed protein were combined and buffer exchanged into dialysis buffer using a desalting column (HiPrep 26/10 Desalting, GEHealthcare) and either used straight away or prepped for storage.

## 7.4 Chemo-enzymatic syntheses

### *C. jejuni* enzymatic synthesis of Pse5Ac7Ac (2.1)

Initial reaction mixtures were composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, 1.5 mM acetyl-coA, 1.5 mM phosphoenolpyruvate in either Tris-HCl buffer (50 mM, pH 7.4) or sodium phosphate buffer (50 mM, pH 7.4). Addition of the *C. jejuni* biosynthetic enzymes each to a final concentration of 0.38 mg mL<sup>-1</sup> was used to initiate the reaction and the mixture was incubated 180 rpm, 37 °C. Small scale reactions were carried out on < 1mg scale.

### *A. cavaie* enzymatic synthesis of Pse5Ac7Ac (2.1)

Reaction mixtures were composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, 1.5 mM acetyl-coA, 1.5 mM phosphoenolpyruvate in sodium phosphate buffer (50 mM, pH 7.4). Biosynthetic enzymes were added to final concentration of (0.38 mg mL<sup>-1</sup>) in reactions with a total volume of 1 mL as follows:

Experiment	Enzyme composition
A	<i>C. jejuni</i> PseB, PseC, PseH, PseG, PseI
B	<i>C. jejuni</i> PseB, PseC, PseH, PseG, <i>A. cavaie</i> NeuB
C	<i>C. jejuni</i> PseB, PseC, PseH, PseI, <i>A. cavaie</i> FlmD
D	<i>C. jejuni</i> PseB, PseH, PseG, PseI, <i>A. cavaie</i> FlmB
E	<i>C. jejuni</i> PseC, PseH, PseG, PseI, <i>A. cavaie</i> FlmA
F	<i>A. cavaie</i> FlmA, FlmB, FlmD, NeuB

All reactions were incubated 37 °C overnight, with aliquots taken at regular intervals for testing.

### *PseB and PseC experiments in deuterium oxide*

PseB and PseC were dialysed into deuterated Tris-HCl buffer (50 mM, pH 7.4) by repeated concentrating to  $\frac{1}{10}$  of their original volumes and addition of the deuterated buffer. Reaction mixtures were composed of final concentrations of 1 mM UDP-GlcNAc, 10 mM L-Glu and 1.5 mM PLP in deuterated Tris- HCl buffer (50 mM, pH 7.4). PseB was added to a final concentration of 25 µM for experiments only containing PseB and the reaction incubated 120 rpm, 37 °C, after 2 hours PseC was added to a final concentration of 0.1 mg mL<sup>-1</sup>. For experiments containing both PseB and PseC, PseB was added to a final concentration of 25 µM to the reaction mixture

after PseC had been added to a final concentration of 25  $\mu\text{M}$  or 125  $\mu\text{M}$  and the reaction incubated 120 rpm, 37  $^{\circ}\text{C}$ .

#### *Chemical acetylation*

PseC to a final concentration of 25  $\mu\text{M}$  followed by PseB to a final concentration of 25  $\mu\text{M}$ , were added to a reaction mixture composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid in Tris-HCl buffer (50 mM, pH 7.4), with a final volume of 16.5 mL and incubated 180 rpm, 37  $^{\circ}\text{C}$ . Aliquots were removed at regular intervals and monitored with LC-MS negative ESI. After 4.5 hours, enzymes were removed using molecular weight cut off falcon tubes (30 kDa) and the resulting solution lyophilised.

Crude PseC product was resuspended in MeOH (5 mL), acetic anhydride (100  $\mu\text{L}$ , 11.5 mmol) and silver acetate (20 mg, 0.12 mmol) were added and the reaction mixture stirred for 4 hours. The reaction was filtered through celite, concentrated *in vacuo*, and resuspended in dH<sub>2</sub>O (5 mL) and lyophilised three times in dH<sub>2</sub>O to remove remaining acetic anhydride.

#### *Acetyl-coA regeneration*

PseC to a final concentration of 125  $\mu\text{M}$  and PseH to a final concentration of 50  $\mu\text{M}$  followed by PseB to a final concentration of 25  $\mu\text{M}$ , were added to reaction mixtures composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid in sodium phosphate buffer (50 mM, pH 7.4). An initial trial had the following added and reactions were incubated 120 rpm, 37  $^{\circ}\text{C}$  and monitored with LC-MS negative ESI after 3 hours;

1. 0 mM acetyl-coA, 20 mM acetylthiocholine iodide
2. 0.15 mM acetyl-coA, 0 mM acetylthiocholine iodide
3. 0.15 mM acetyl-coA, 2 mM acetylthiocholine iodide
4. 0.15 mM acetyl-coA, 20 mM acetylthiocholine iodide
5. 1.5 mM acetyl-coA, 0 mM acetylthiocholine iodide

Further reactions (of the core mixture above) consisted of acetylthiocholine iodide to a final concentration of 20 mM or 100 mM with increasing amounts of CoA (0 mM, 0.0015 mM, 0.015 mM, 0.15 mM) and incubated 120 rpm 37  $^{\circ}\text{C}$ . The reaction mixture was monitored with LC-MS negative ESI after 4 hours.

A reaction mixture composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, 0.0015 mM CoA, 100 mM acetylthiocholine iodide, 1.5 mM phosphoenolpyruvate in sodium phosphate buffer (50 mM, pH 7.4) was incubated 120 rpm 37  $^{\circ}\text{C}$  with PseB (25  $\mu\text{M}$ ), PseC (125  $\mu\text{M}$ ), PseH (50  $\mu\text{M}$ ), PseG (30  $\mu\text{M}$ ) and Pse I (25  $\mu\text{M}$ ). Aliquots were taken at regular intervals and monitored with LC-MS negative ESI for production of Pse5Ac7Ac.



### *Acetyl-coA co-factor substitution*

Control reaction mixtures composed of 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, 1.5 mM phosphoenolpyruvate in sodium phosphate buffer (50 mM, pH 7.4), with PseB (25 µM), PseC (125 µM), PseG (30 µM), Pse I (25 µM) and either PseH (50 µM) or 10 mM *N*-acetyl cysteamine thioacetate. As well as a DMSO control reaction that contained the core mixture above with PseH (50 µM), 1.5 mM acetyl-CoA and 10 % DMSO. Reaction mixtures were incubated 120 rpm, 37 °C and aliquots monitored with LC-MS negative ESI for production of Pse5Ac7Ac.

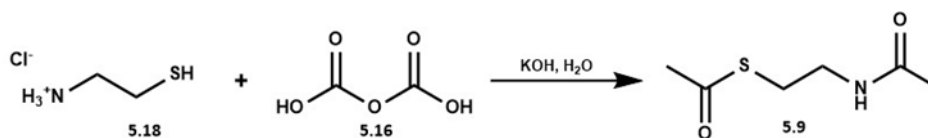
A reaction mixture containing 1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, in sodium phosphate buffer (50 mM, pH 7.4), with PseB (25 µM), PseC (125 µM), PseH (50 µM) with 10 mM *N*-acetyl cysteamine thioacetate, 10 mM phenyl acetate, 10 mM *p*-tolyl acetate, 10 mM *S*-phenyl thioacetate, 10 mM *p*-nitrophenyl acetate, 10 mM *S*-(4-nitrophenyl)thioacetate, or 10 mM *S*-(*p*-tolyl)thioacetate (with the latter three suspended as a 10X stock in DMSO) were incubated 120 rpm, 37 °C. Aliquots were taken at regular intervals and monitored with LC-MS negative ESI.

### *Enzymatic synthesis of Pse5Ac7Az*

1 mM UDP-GlcNAc, 1 mM pyridoxal 5' phosphate, 10 mM L-Glutamic acid, 1.5 mM phosphoenolpyruvate in sodium phosphate buffer (50 mM, pH 7.4), were incubate (120 rpm, 37 °C) with PseB (25 µM), PseC (125 µM), PseH (50 µM), PseG (30 µM), Pse I (25 µM) and either 10 mM *N*-acetyl cysteamine thioacetate or 10 mM *N*-acetyl cysteamine thioazidoacetate. Aliquots were monitored with LC-MS negative ESI.

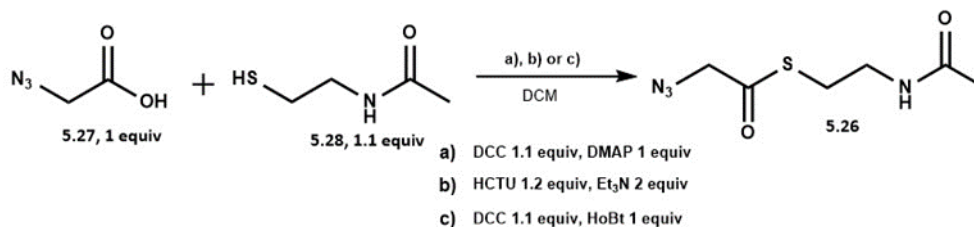
## 7.5 Chemical synthesis

### *S*-(2-acetamidoethyl) ethanethioate (SNAc)



Cysteamine HCl (1 g, 8.8 mmol) was suspended in dH<sub>2</sub>O (8 mL) and the pH was adjusted to 8.0 using 4 N KOH on ice. Whilst maintaining the pH at 8 with 4 N KOH, acetic anhydride (2.5 mL, 24.5 mmol) was added dropwise. 1 N HCl was added to adjust the pH to 7.0 and the reaction mixture was stirred at 0 °C for 2 hours. Brine (10 mL) was added to the reaction mixture and the product was extracted with DCM (3 x 10 mL). The organic layer was dried (MgSO<sub>4</sub>) and condensed azeotroping with toluene. The resulting white solid appeared pure by mass spec and NMR. *R<sub>f</sub>* 0.56 (5:1 (v/v) DCM:MeOH). <sup>1</sup>H NMR (500 MHz, Chloroform-*d*) δ 3.23 (q, *J* = 6.5 Hz, 2H), 2.87 (t, *J* = 6.8 Hz, 2H), 2.19 (s, 3H), 1.83 (s, 3H) (**Appendix 9**).

### *S*-(2-(2-azido)acetamido)ethyl ethanethioate (SNAz)



- A) Azidoacetic acid (57 μL, 0.77 mmol) was added to DCC (0.173 g, 0.84 mmol) suspended in anhydrous DCM (11 mL) and stirred for 5 minutes under a N<sub>2</sub> atm in a water/ice bath. DMAP (0.093 g, 0.76 mmol) and N-acetyl cysteamine (89 μL, 0.84 mmol) were added and the reaction mixture was left stirring to warm to r.t overnight.
- B) Azidoacetic acid (57 μL, 0.77 mmol) was added to HCTU (0.382 g, 0.924 mmol) suspended in anhydrous DCM (11 mL) and stirred at 0 °C for 5 minutes under a N<sub>2</sub> atm in a water/ice bath. N-acetyl cysteamine (89 μL, 0.84 mmol) and Et<sub>3</sub>N (215 μL, 1.54 mmol) were added and the reaction mixture was left stirring to warm to r.t overnight.

C) Azidoacetic acid (57  $\mu$ L, 0.77 mmol) was added to a DCC (0.173 g, 0.84 mmol) suspended in anhydrous DCM (11 mL) and stirred for 5 minutes under a N<sub>2</sub> atm in a water/ice bath. N-acetyl cysteamine (89  $\mu$ L, 0.84 mmol) and HoBt (0.104 g, 0.77 mmol) were added and the reaction mixture was left stirring to warm to r.t overnight.

Reaction mixtures A) and B) were filtered *in vacuo*, condensed and resuspended in ethyl acetate (20 mL). After filtering the filtrate was washed with brine (20 mL), dried (NaSO<sub>4</sub>) and condensed *in vacuo* to afford a syrup. The crude oils were purified by flash column chromatography (silica gel; step-wise gradient of hexane and ethyl acetate; hexane, hexane-ethyl acetate 3:1 (v/v), hexane-ethyl acetate 2:1 (v/v), hexane-ethyl acetate 1:1 (v/v), hexane-ethyl acetate 1:2 (v/v), hexane-ethyl acetate 1:3 (v/v), ethyl acetate 1, pure fractions were combined and condensed to afford a syrup. R<sub>F</sub> 0.28 (1:1 (v/v) ethyl acetate:hexane). <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  4.12 (s, 2H), 3.33 (t, *J* = 6.6 Hz, 2H), 3.06 (t, *J* = 6.6 Hz, 2H), 1.90 (s, 3H) (**Appendix 9**).

## **Appendix**

## 1. PA2794 sequences

PA2794 DNA coding sequence, optimised for *E. coli*.

```
ATGAACACCTATTTTGTATATCCGCATCGCCTGGTGGGCAAAGCGCTGTATGAAAGCTAT
TATGATCATTTTGGCCAGATGGATATCTGAGCGATGGCAGCCTGTATCTGATTTATCGC
CGCGCGACCGAACATGTGGGCGGCAGCGATGGCCGCGTGGTGTTTAGCAAACCTGGAAGGC
GGCATTGGAGCGCGCCGACCATTGTGGCGCAGGCGGGCGCCAGGATTTTCGCGATGTG
GCGGGCGGCACCATGCCGAGCGGCCGCATTTGTGGCGGCAGCACCGTGTATGAAACCGGC
GAAGTGAAAGTGTATGTGAGCGATGATAGCGGCGTGACCTGGGTGCATAAAATTTACCCTG
GCGCGCGCGGCGCGGATATAACTCCGCGCATGGCAAAAGCTTTTCAGGTGGGCGCGCGC
TATGTGATTCGCTGTATGCGGCGACCGCGTGAACATGAACTGAAATGGCTGGAAAGC
AGCGATGGCGGCGAAACCTGGGGCGAAGGCAGCACCATTTATAGCGGCAACACCCCGTAT
AACGAAACCAGCTATCTGCCGTTGGCGATGGCGTGATCTGGCGGTGGCGCGCGTGGGC
AGCGGCGCGGCGCGCGCTGCGCCAGTTTATTAGCCTGGATGATGGCGGCACCTGGACC
GATCAGGGCAACGTGACCGCGCAGAACGGCGATAGCACCGATATTCTGGTGGCGCCGAGC
CTGAGCTATATTTATAGCGAAGGCGGCACCCCGCATGTGGTGTCTGTATACCAACCGC
ACCACCATTTTTGCTATATCGCACCATTTCTGCTGGCGAAAGCGGTGGCGGGCAGCAGC
GGCTGGACCGAACGCGTGGCGGTGTATAGCGCGCCGCGCGGCGAGCGGCTATACCAGCCAG
GTGGTGTGGGCGCGCCCGCATTTCTGGGCAACCTGTTTCGCGAAACCAGCAGCACCCACC
AGCGGCGGTATCAGTTTGAAGTGTATCTGGGCGGCGTGGCGGATTTTGAAGCGATTGG
TTTAGCGTGAGCAGCAACAGCCTGTATACCCTGAGCCATGGCCTGCAGCGCAGCCCGCGC
CGCGTGGTGGTGAATTTGCGCGCAGCAGCAGCCGAGCACCTGGAACATTGTGATGCCG
AGCTATTTAACGATGGCGGCCATAAAGGCAGCGGCGCGCAGGTGGAAGTGGGCGAGCCTG
AACATTCGCTGGGCACCGCGCGGCGGTGTGGGGCACCGGCTATTTTGGCGGCATTGAT
AACAGCGCGACACCCGCTTTGCGACCGGCTATTATCGCGTGGCGCGGTGGATT
```

PA2794 amino acid sequence expressed by a pET15b recombinant plasmid, highlighting the N-terminal hexa-his tag (red) and thrombin cleavage site (blue) MW 47151 g mol<sup>-1</sup>).

```
MGSSHHHHHSSGLVPRGSHMNTYFDI PHRLVKGALYESYYDHFQMDILSDGSLYLIYR
RATEHVGGSDGRVVF SKLEGGIWSAPTIVAQAGGQDFRDVAGGTMPSGRIVAASTVYETG
EVKVYVSDDSGVTWVHKFTLARGGADYNFAHGKSFQVGARYVIPLYAATGVNYELKWLES
SDGGETWEGSTIYSGNTPYNETSYLPVGDGVILAVARVGSAGGALRQFISLDDGGTWT
DQGNVTAQNGDSTDILVAPSLSYIYSEGGTPHVLLYTNRTTHFCYYRTILLAKAVAGSS
GWTERPVPVYSAPAASGYTSQVVLGGRRILGNLFRETSSTTS GAYQFEVYLGGVPDFESDW
FSVSSNSLYTL SHGLQRSPRRVVVEFARSSSPSTWNIVMPSYFNDGGHKGSGAQVEVGS L
NIRLGTGAAVWGTGYFGGIDNSATTFATGYRVRRAWI
```

## 2. PA2794 F129A sequences

PA2794 F129A DNA coding sequence, with the mutation highlighted in **green**.

```
ATGAACACCTATTTTGGATATTCGCGATCGCCTGGTGGGCAAAGCGCTGTATGAAAGCTAT
TATGATCATTTTGGCCAGATGGATATTTCTGAGCGATGGCAGCCTGTATCTGATTTATCGC
CGCGCGACCGAACATGTGGGCGGCAGCGATGGCCGCGTGGTGTTTAGCAAACCTGGAAGGC
GGCATTGGAGCGCGCCGACCATTTGTGGCGCAGGCGGGCGGCCAGGATTTTCGCGATGTG
GCGGGCGGCACCATGCCGAGCGGCCGCATTTGTGGCGGCAGCACCGTGTATGAAACCGGC
GAAGTGAAAGTGTATGTGAGCGATGATAGCGGCGTGACCTGGGTGCATAAATTTACCCTG
GCGCGCGCGCGCGCGGATATAACGCCGCGCATGGCAAAAGCTTTTCAGGTGGGCGCGCGC
TATGTGATTCGCTGTATGCGGCGACCGCGTGAACATGAACTGAAATGGCTGGAAAGC
AGCGATGGCGGCGAAACCTGGGGCGAAGGCAGCACCATTTATAGCGGCAACACCCCGTAT
AACGAAACCAGCTATCTGCCGGTGGGCGATGGCGTGATTTCTGGCGGTGGCGCGCGTGGG
AGCGGCGCGGGCGCGCGCTGCGCCAGTTTATTAGCCTGGATGATGGCGGCACCTGGACC
GATCAGGGCAACGTGACCGCGCAGAACGGCGATAGCACCGATATTCTGGTGGCGCCGAGC
CTGAGCTATATTTATAGCGAAGGCGGCACCCCGCATGTGGTGTCTGTATACCAACCGC
ACCACCATTTTTGCTATTATCGCACCATTTCTGCTGGCGAAAGCGGTGGCGGGCAGCAGC
GGCTGGACCGAACGCGTGGCGGTGTATAGCGCGCCGGCGGCGAGCGGCTATACCAGCCAG
GTGGTGTGGGCGCGCCCGCATTCTGGGCAACCTGTTTCGCGAAACCAGCAGCACCCACC
AGCGGCGGTATCAGTTTGAAGTGTATCTGGGCGGCGTGGCGGATTTTGAAGCGATTGG
TTTAGCGTGAGCAGCAACAGCCTGTATACCCTGAGCCATGGCCTGCAGCGCAGCCCGCGC
CGCGTGGTGGTGAATTTGCGCGCAGCAGCAGCCGAGCACCTGGAACATTGTGATGCCG
AGCTATTTAACGATGGCGGCCATAAAGGCAGCGGCGCGCAGGTGGAAGTGGGCGAGCCTG
AACATTCGCTGGGCACCGCGCGCGGTGTGGGGCACCGGCTATTTTGGCGGCATTGAT
AACAGCGCACCCCGCTTTGCGACCGGCTATTATCGCGTGGCGCGTGGATT
```

PA2794 F129A codon sequence expressed by a pET15b recombinant plasmid, highlighting the mutation (**green**), N-terminal hexa-his tag (**red**) and thrombin cleavage site (**blue**) MW 47151 g mol<sup>-1</sup>).

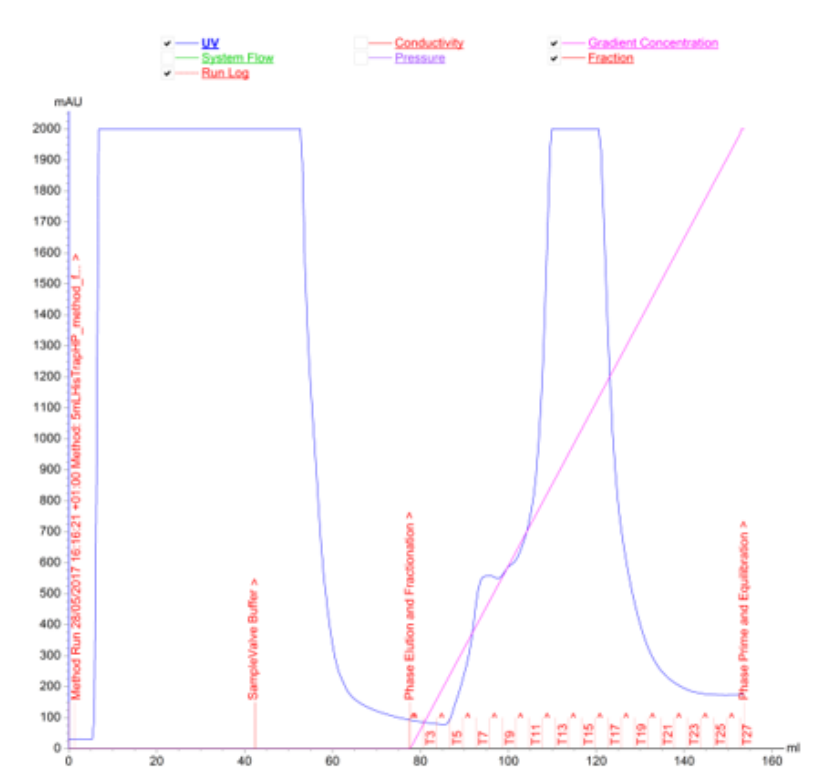
```
MGSSHHHHHHSSGLLVPRGSHMNTYFDI PHRLVGKALYESYYDHFQMDILSDGSLYLIYR
RATEHVGGSDGRVVF SKLEGGIWSAPTIVAQAGGQDFRDVAGGTMPSGRIVAASTVYETG
EVKVYVSDDSGVTWVHKFTLARGGADYNAAHGKSFQVGARYVIPLYAATGVNYELKWLES
SDGGETWEGSTIYSGNTPYNETSYLPVGDGVI LAVARVGSAGGALRQFISLDDGGTWT
DQGNVTAQNGDSTDILVAPSLSYIYSEGGTPHVLLYTNRTTHFCYYRTILLAKAVAGSS
GWTERVPVYSAPAASGYTSQVVLGGRRILGNLFRETSSTTSGAYQFEVYLGGVPDFESDW
FSVSSNSLYTL SHGLQRS PRRVVVEFARSSSPSTWNI VMPSYFNDGGHKGSGAQVEVGS L
NIRLGTGA AVWGTYFGGIDNSATTRFATGYRVR AWI
```

### 3. PA2794 and PA2794 F129A Akta trace for Ni-His trap purification

Equilibrium buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 10 mM imidazole.

Elution buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 500 mM imidazole.

Lysate (55 mL) loaded onto the Ni-His trap column, followed by washing with equilibrium buffer (25 mL). Fractions (3 mL) were collected during application of a linear gradient of 0-100 % elution buffer over 80 mL.

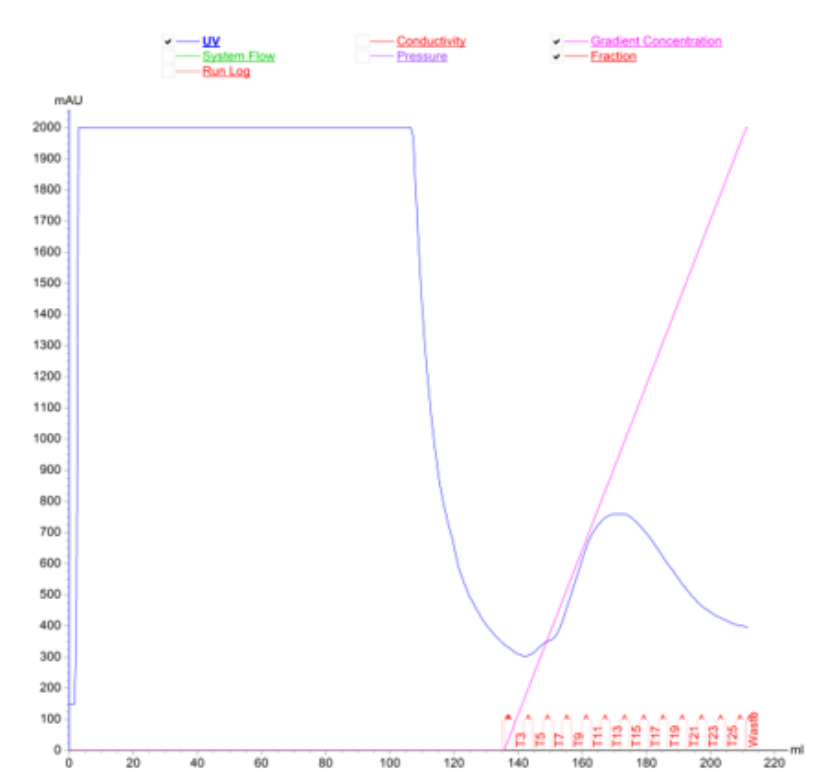


3.1 PA2794 Akta trace; demonstrating binding and elution of a hexa-his tagged protein.

Equilibrium buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 10 mM imidazole.

Elution buffer; 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol, 500 mM imidazole.

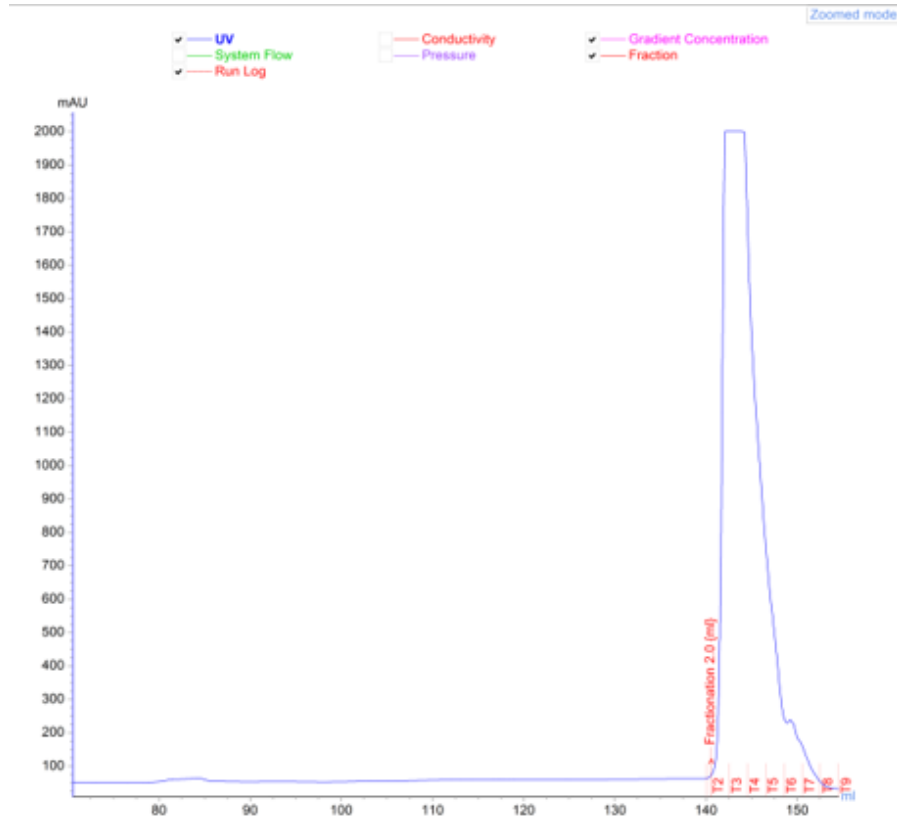
Lysate (110 mL) loaded onto the Ni-His trap column, followed by washing with equilibrium buffer (25 mL). Fractions (3 mL) were collected during application of a linear gradient of 0-100 % elution buffer over 80 mL.



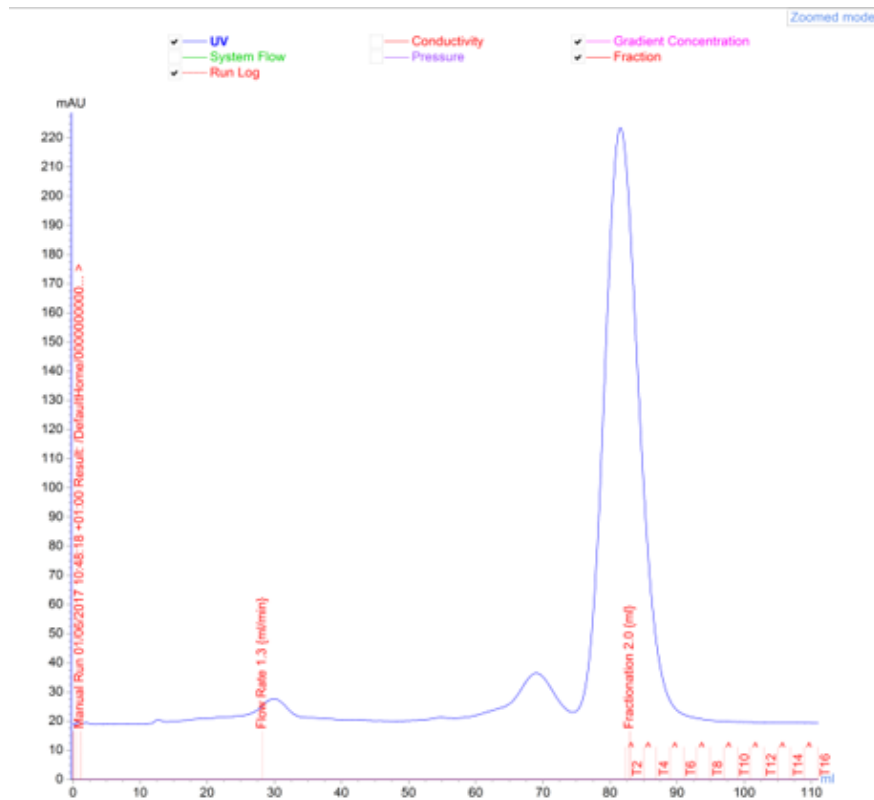
3.2 PA2794 F129A Akta trace; demonstrating binding and elution of a hexa-his tagged protein.



#### 4. PA2794 and PA2794 F129A Akta trace for size exclusion purification



4.1 PA2794 size exclusion Akta trace in 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol.



4.2 PA2794 F129A size exclusion Akta trace in 50 mM Tris-HCl buffer, pH7.5, 10 % glycerol.

## 5. Crystallography statistics

	PA2794 apo	PA2794 ligand
Data collection		
Space group	<i>P</i> 213	<i>P</i> 213
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	126.6, 126.6, 126.6	127.7, 127.7, 127.7
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	89.52-1.22	73.70-1.94
R <sub>sym</sub> or R <sub>merge</sub>	0.096 (2.104)	0.083 (1.159)
R <sub>pim</sub>	0.023 (0.521)	0.020 (0.284)
CC <sub>1/2</sub>	0.999 (0.538)	1.000 (0.838)
<i>I</i> / $\sigma$ <i>I</i>	15.7 (1.5)	26.1 (3.6)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Redundancy	19.5 (18.2)	19.8 (18.8)
Refinement		
No. reflections	55643	54719
R <sub>work</sub> / R <sub>free</sub>	0.133/0.152	0.200/0.227
No. atoms		
Protein	3432	3350
Ligand/ion	41	30
Water	479	114
<i>B</i> -factors		
Protein	16.4	42.7
Ligand/ion	38.3	57.0
Water	29.6	38.8
R.m.s deviations		
Bond lengths (Å)	0.021	0.0106
Bond angles (°)	2.23	1.734

$$R_{\text{merge}} = \frac{\sum_h \sum_i (|<I_h> - I_{h,i}|)}{\sum_h \sum_i I_{h,i}}$$

*h* = the unique reflections and *i* = their symmetry-equivalent contributors

$$R_{\text{pim}} = \frac{\sum_{hkl} \sqrt{\frac{1}{n-1} \sum_{j=1}^n |I_{hklj} - I_{hkl}|}}{\sum_{hkl} \sum_j I_{hklj}}$$

$$R_{\text{free}} = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum F_{\text{obs}}}$$

*F*<sub>obs</sub> = experimentally obtained structure factor amplitudes and *F*<sub>calc</sub> = calculated structure factor amplitudes

## 6. *Campylobacter jejuni* biosynthetic enzymes.

*E. coli* codon optimised DNA sequences for purchased genes and amino acid sequences of all enzymes.

### 6.1 *PseB* (*Cj1293*) MW 37372 g mol<sup>-1</sup>

MFNGKNILITGGTGSFGKTYTKVLLENYKPNKIIIIYSRDELKQFEMSSIFNSNCMRYFIG  
DVRDKERLSVAMRDVDFVIHAAAMKHVPVAEYNPMECIKTNIHGAQNVIDACFENGVKKC  
IALSTDKACNPVNLYGATKCLASDKLFVAANNIAGNKQTRFSVTRYGNVVGSRGSVVPFFK  
KLI AQGSKELPITDTRMTRFWISLEDGKVFVLSNFERMHGGEIFIPKIPSMKITNLAAHAL  
APNLSHKIIGIRAGEKLHEIMISSDDSHLTYEFENYYAISPSIKLVDQESDFSINALGEK  
GQKVKDGFSSYSSDNNPQWASEKELLDIINHTEGF

### 6.2 *PseC* (*Cj1294*) MW 42317 g mol<sup>-1</sup>

MLTYSHQNI DQSDIDTLTKALKDEILTGGKKVNEFEALCEYMGVKHACVLNSATSALHL  
AYTALGVQEKIVLTTPLTFAATANAAALMAGAKVEFIDIKNDGNI DEKKLEARLLKESENI  
GAISVVDVFAGNSVEMDEISNLTKKYNIPLI DDASHALGALYKSEKVGKKADLSIFS FHPV  
KPIITTFEGGAVVSDNEELIDKIKLLRSHGIVKKRLWDS DMVELGYNRYRLSDVACALGINQ  
LKKLDHNLEKREEIANFYDKEFEKNPYFSTIKIKDYKSSRHLYPILLFPEFYCQKEELF  
ESLLHAGIGVQVHYKPTYEF SFYKLLGEIKLQ NADNFYKAELSIPCHQEMNLKDAKFVK  
DTLFSILEKVKKGYCG

### 6.3 *PseH* (*Cj1313*) MW 15103 g mol<sup>-1</sup>

ATGCTGATTA AACTGAAAACTTCGCGAACTGAATAGCCAGGAAATTA AACTGATCTTCAAATGGCGTAACCACC  
CGGACATTAGCCAATTCATGAAGACCAAACACATCGACTTCGAGGAACACCTGCGTTTTATTTCGTAACCTGCACCA  
GGATAGCAACAAGAAATACTTCCTGGTGTTCAGGACGAGCAAATCATTGGTGTGATCGATTCGTTAACATTACC  
ACCAAAGCTGCGAATTTGGCCTGTATGCGATCCCGGACCTGAAGGGTGTGGGCCAAGTTCTGATGAACGAGATC  
AAGAAATACGCGTTCGAAATTCTGAAGGTGGACACCCTGAAAGCGTATGTTTTAAGGATAACCACAAGGCGCTG  
AAACTGTACCAGCAAACCACTTTACCATTTATGATGAGGACAAGGACTTTTATTATGTGTGCCTGAAACAGAGCC  
ACTGCAAGGCGCTGCCGAGCTAA

MIKLNFTLNSQEIELIFKWRNHPDINQFMKTKYIDFEEHLRFLKKLHQDSSKKYFLVF  
QDEQIIGVIDFVNITTKSCEFGLYAKPNLKGVGQILMNEIIKYAFENLKVNTLKAYVFKD  
NRKALKLYQQNHFTIYDEDKDFYHICLKQSDCKALPS

#### 6.4 PseG (*Cj1312*) MW 31319 g mol<sup>-1</sup>

ATGAAGGTGCTGTTCCGTAGCGACAGCAGCAGCCAGATCGGTTTTGGCCACATTAAGCGTGACCTGGTGCTGGCG  
AAACAATACAGCGATGTTAGCTTTGCGTGCCTGCCGCTGGAGGGTAGCCTGATCGACGAAATCCGTACCCGGTT  
TATGAACTGAGCAGCGAGAGCATCTACGAACTGATCAACCTGATTAAGAGAGAAATTCGAGCTGCTGATCATT  
GATCACTATGGTATCAGCGTGGACGATGAGAAGCTGATTAAGCTGAAACCGGCGTTAAGATCCTGAGCTTTGAC  
GATGAAATTAACCGCACCCTGCGACATCCTGCTGAACGTGAACGCGTACGCGAAGGCGAGCGATTATGAGGG  
TCTGGTGCCGTTCAAATGCGAAGTTCGTTGCGGCTTAGCTACGCGCTGATCCGTGAGGAATTCTATCAGGAAGC  
GAAGGAAAACCGTGAGAAGAAATACGACTTCTTTATTTGCATGGGTGGCACCGATATCAAAAACCTGAGCCTGCA  
GATTGCGAGCGAGCTGCCGAAGACCAAAATCATTAGCATCGCGACCAGCAGCAGCAACCCGAACCTGAAGAAAC  
TGCAAAAGTTCGCGAAACTGCACAACAACATCCGTCTGTTTATTGATCACGAGAACATTGCGAAGCTGATGAACG  
AAAGCAACAACTGATCATTAGCGCGAGCAGCCTGGTGAACGAGGCGCTGCTGCTGAAGGCGAACTTTAAAGCG  
ATCTGCTACGTTAAGAACCAAGAAAGCACCGCGACCTGGCTGGCGAAGAAAGGCTATGAGGTTGAATACAAATAT  
TAA

MLDKILYFKTLIRADSGSKIIGHGHVRRDLILAKNFKDVSFACIDLPGSLTGEI PCPVFTL  
KSADINELVNLIKEHKFELLIIDHYGISAADEKLIKEQTNVKILCFDDNYKEHFCDYLLN  
VNIYAQPQKYVNLV PANCELVFSPLVRSEFYDEAKIKREKKFDCFIALGGTDALNLTAKI  
ASNLLAKNKKVAAITTSANANLANLQNLADSESNFSLFINSNEVARLMNESEILVISASS  
LVNEALVVLGAKFKAVRVADNQNEMAQWLAANGREIYEADEICLNL

#### 6.5 Psel (*Cj1317*) MW 38647 g mol<sup>-1</sup>

ATGCATATGCAAATTGGTAACTTTAACACCGACAAGAAGGTTTTTATCATTGCGGAACTGAGCG  
CGAATCATGCGGGTAGCCTGGAGATGGCGCTGAAGAGCATCAAAGCGGCGAAGAAAGCGGGT  
GCGGACGCGATCAAGATTCAGACCTACACCCCGGATAGCCTGACCCTGAACAGCGACAAAGAG  
GACTTCATCATTAAAGGTGGCCTGTGGGACAAGCGTAAACTGTACGAACTGTATGAGAGCGCG  
AAAACCCCGTATGAATGGCACAGCCAGATCTTCGAAACCGCGCAAACGAGGGTATTCTGTGCT  
TCAGCAGCCCGTTTTGCGAAGGAAGACGTGGAGTTCCTGAAACGTTTTGATCCGATCGCGTACAA  
GATTGCGAGCTTCGAAGCGAACGATGAGAAGCTTTGTGCGTCTGATTGCGAAAGAGAAGAAACC  
GACCATCGTTAGCACCGGCATTGCGACCGAGGAAGAGCTGTTCAAGATCTGCGAAATTTTTAAG  
GAAGAGAAAAACCCGGACCTGGTGTTCCTGAAGTGCACCAGCACCTATCCGACCGCGATCGAG  
GATATGAACCTGAAAGGTATTGTTAGCCTGAAGGAAAAATTTAACGTTGAGGTGGGTCTGAGC  
GACCACAGCTTCGGCTTTCTGGCGCCGGTGATGGCGGTTGCGCTGGGTGCGCGTGTTATCGAA  
AAGCACTTCATGCTGGACAAAAGCATTGAAAGCGAGGATAGCAAGTTTAGCCTGGACTTCGAT  
GAATTTAAAGCGATGGTGGATGCGGTTTCGTAAGCGGAGAGCGCGCTGGGTGATGGCAAGCT  
GGACCTGGATGAAAAGGTGCTGAAAACCGTGTTCGCGCGTAGCCTGTACGCGAGCAAAGA  
TATCAAGAAAGGCGAGATGTTTAGCGAAGAGAACGTGAAGAGCGTTCGTCGAGCTTCGGTCT  
GCACCCGAAATTTTATCAAGAACTGCTGGGCAAGAAGGCGAGCAAGGACATCAAGTTCGGTGA  
CGCGCTGAAGCAAGGCGATTTCCAATAA

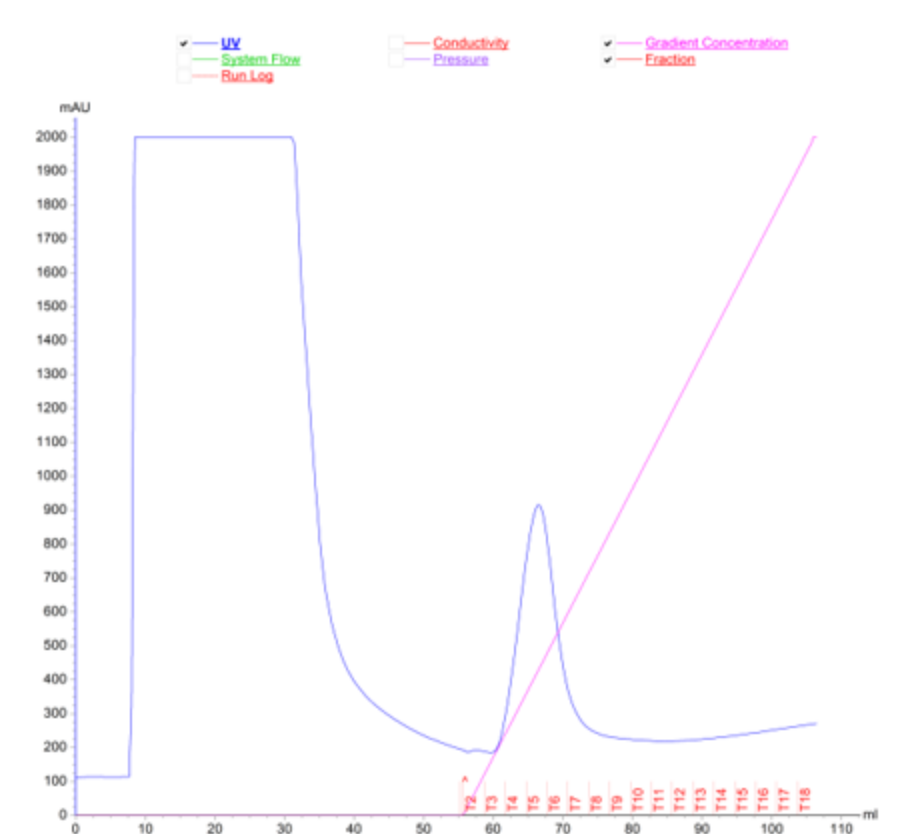
MQIGNFNNTDKKVFIIAELSANHAGSLEMALKSIKAAKKAGADAIKIQTYPDSLTLNSDK  
EDFIIKGGWLDKRKLYELYESAKTPYEWHSQIFETAQNEGILCFSSPFAKEDVEFLKRFD  
PIAYKIASFEANDENFVRLIAKEKKPTIVSTGIATEEELFKICEIFKEEKNPDLVFLKCT  
STYPTAIEDMNLKGI VSLKEKFNVEVGLSDHSFGFLAPVMAVALGARVIEKHFMLDKSIE  
SEDSKFSLDFDEFKAMVDAVRQAESALGDGKLDLDEKVLKNRVFARSLYASKDIKKGEMF  
SEENVKSVRPSFGLHPKFYQELLGKKASKDIKFGDALKQGDFQ

## 7. *Campylobacter jejuni* enzymes Akta trace for Ni-His trap purification.

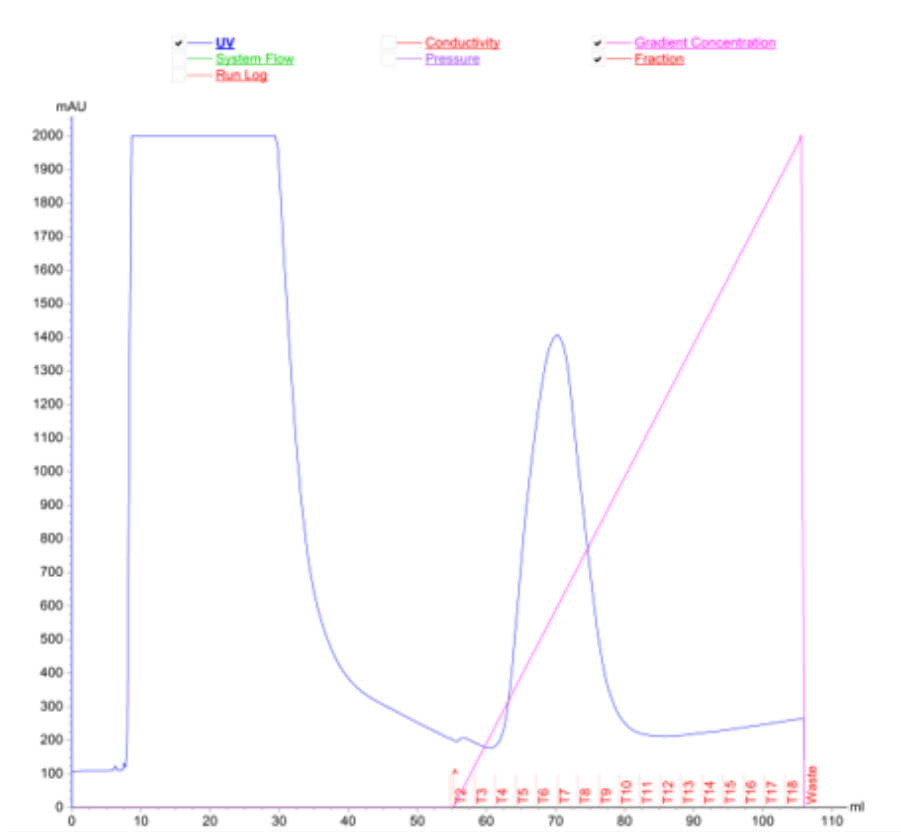
Equilibrium buffer; 50 mM sodium phosphate buffer, pH7.4, 400 mM NaCl, 10 mM imidazole.

Elution buffer; 50 mM sodium phosphate buffer, pH7.4, 400 mM NaCl, 500 mM imidazole.

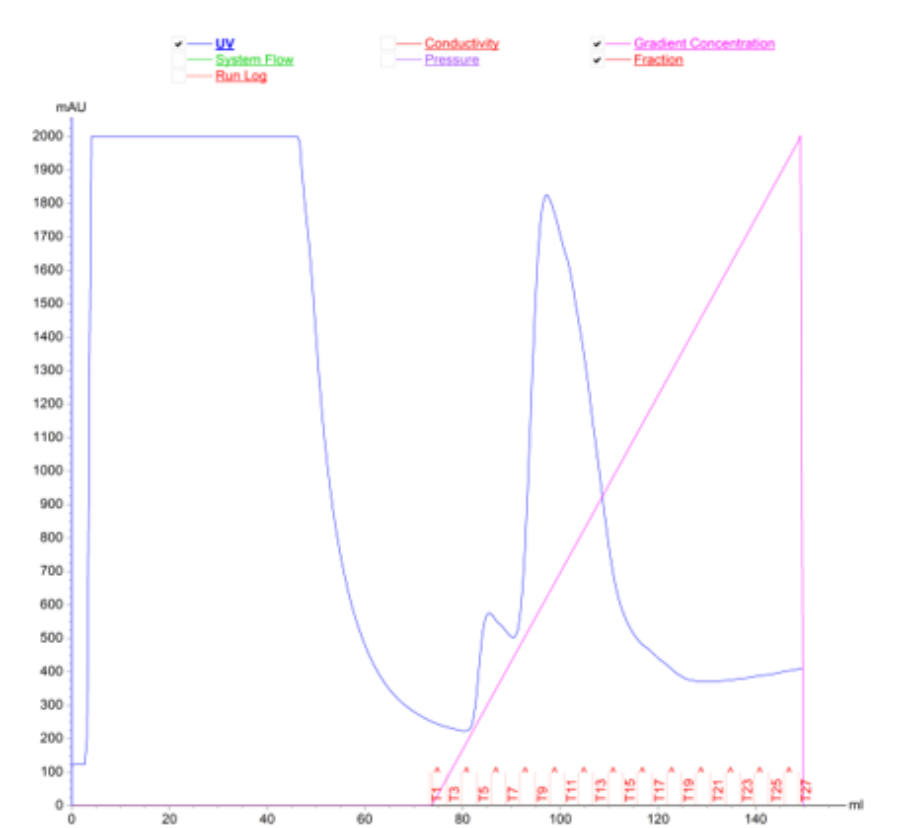
Lysate (30-50 mL) loaded onto the Ni-His trap column, followed by washing with equilibrium buffer (25 mL). Fractions (3 mL) were collected during application of a linear gradient of 0-100 % elution buffer over 50 or 75 mL.



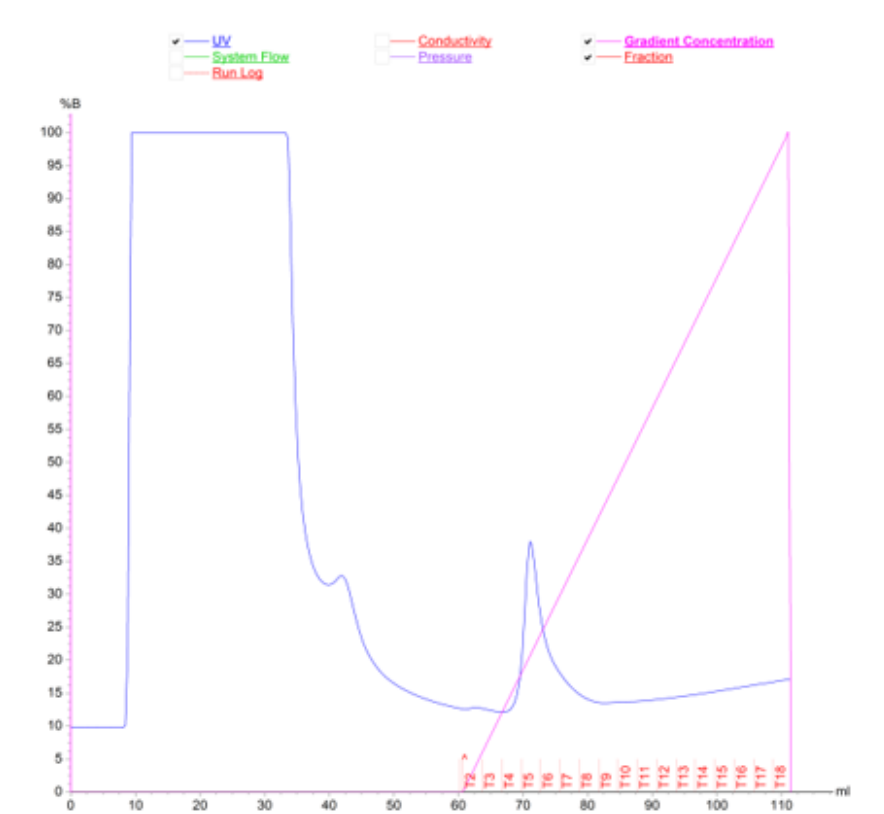
7.1 PseB Akta trace; demonstrating binding and elution of a hexa-his tagged protein.



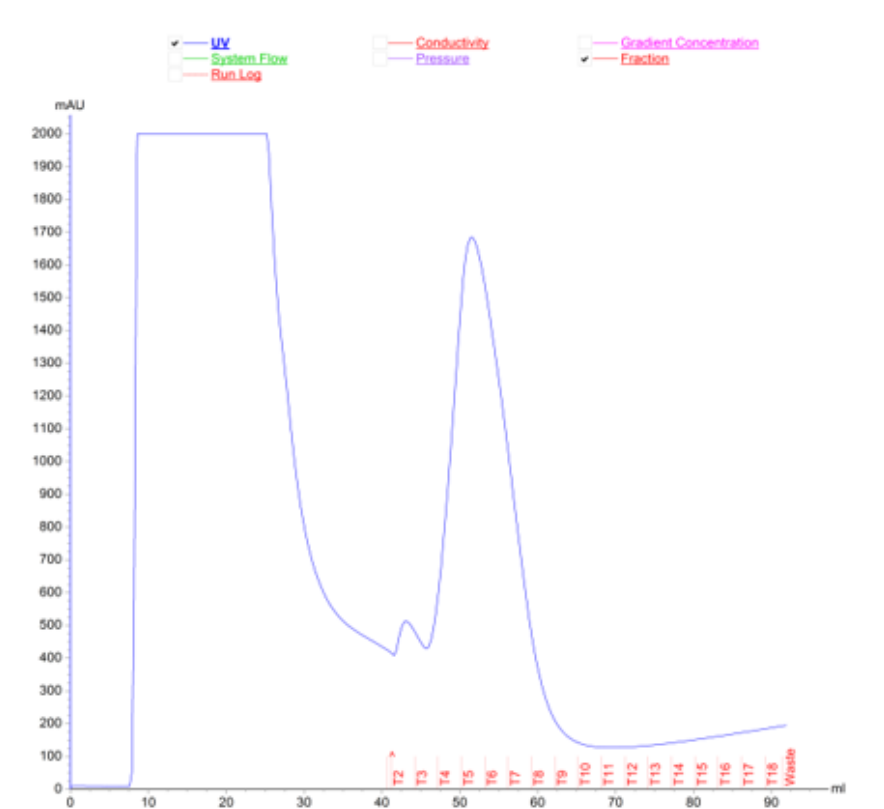
7.2 PseC Akta trace; demonstrating binding and elution of a hexa-his tagged protein.



7.3 PseH Akta trace; demonstrating binding and elution of a hexa-his tagged protein.



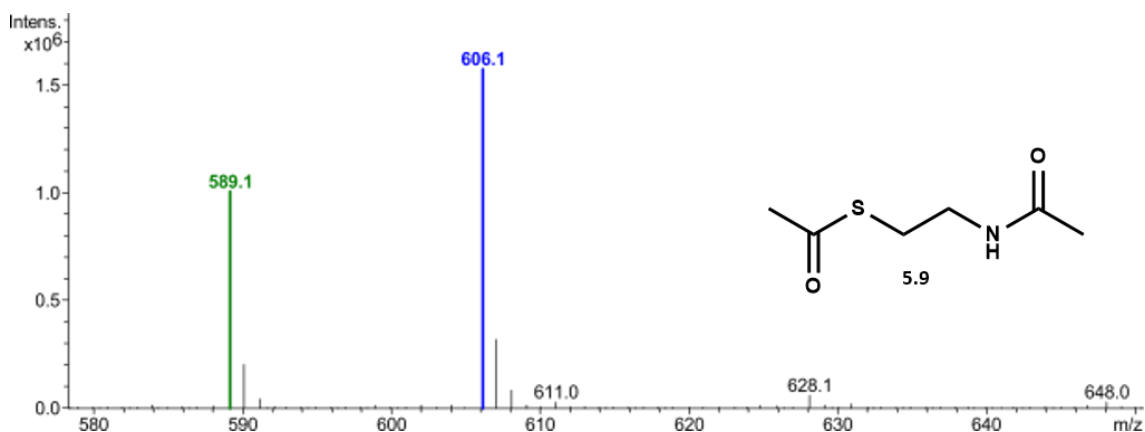
7.4 PseG Akta trace; demonstrating binding and elution of a hexa-his tagged protein.



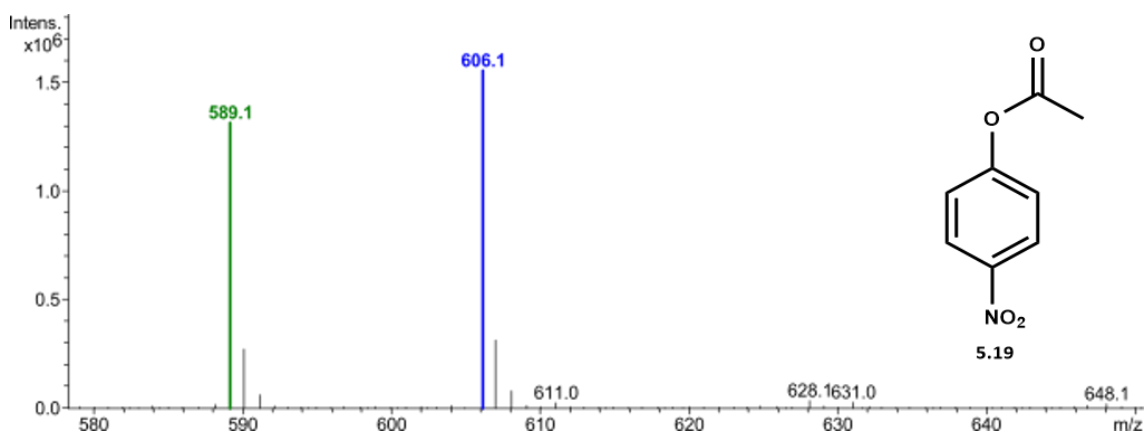
7.5 PseI Akta trace; demonstrating binding and elution of a hexa-his tagged protein.

## 8. LC-MS negative ESI of controls for co-factor substitute reactions

Reaction mixtures containing UDP-GlcNAc (1 mM), PLP (1.5 mM) and L-glutamate (10 mM), PseB (25  $\mu$ M), PseC (125  $\mu$ M) and 10 mM acetyl-donor (**5.9** or **5.19-5.24**) in sodium phosphate buffer (50 mM pH7.4) or 10 % DMSO in sodium phosphate buffer (50 mM pH7.4) as required, were monitored by LC-MS negative ESI after > 3 hours for conversion to the PseH product (**3.9**) as estimated by appearance of a  $[M-H]^-$  631 peak.

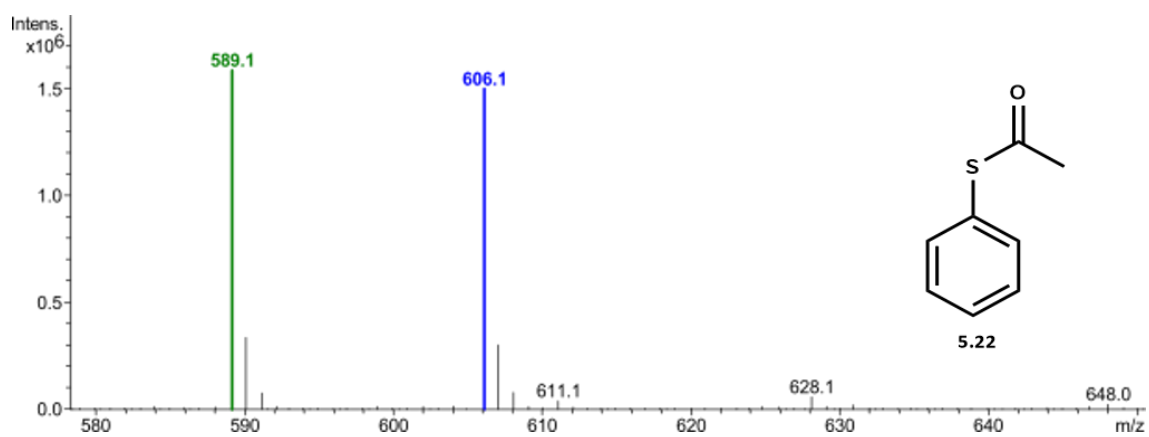


### 8.1 *N*-acetylcysteamine thioacetate (**4.9**) 10 mM

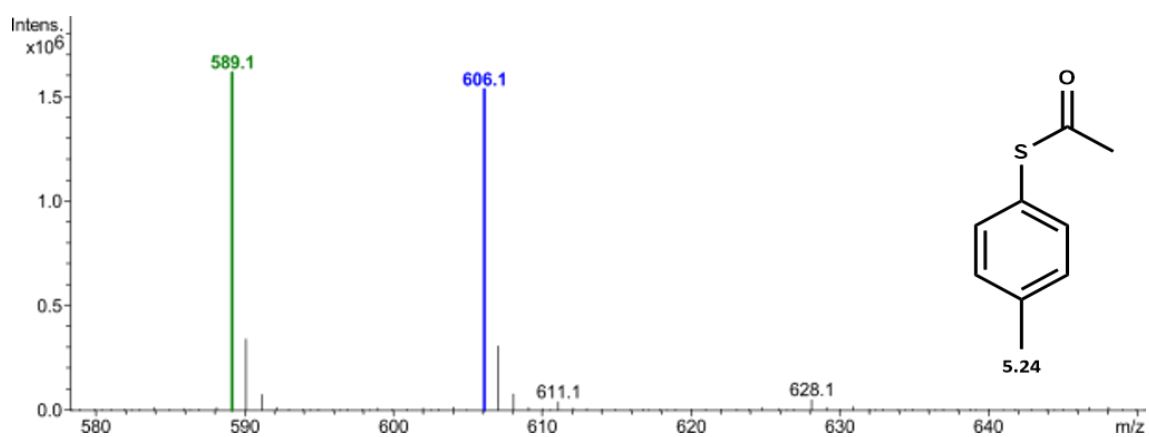


### 8.2 *p*-nitrophenyl acetate (**4.19**) 10 mM

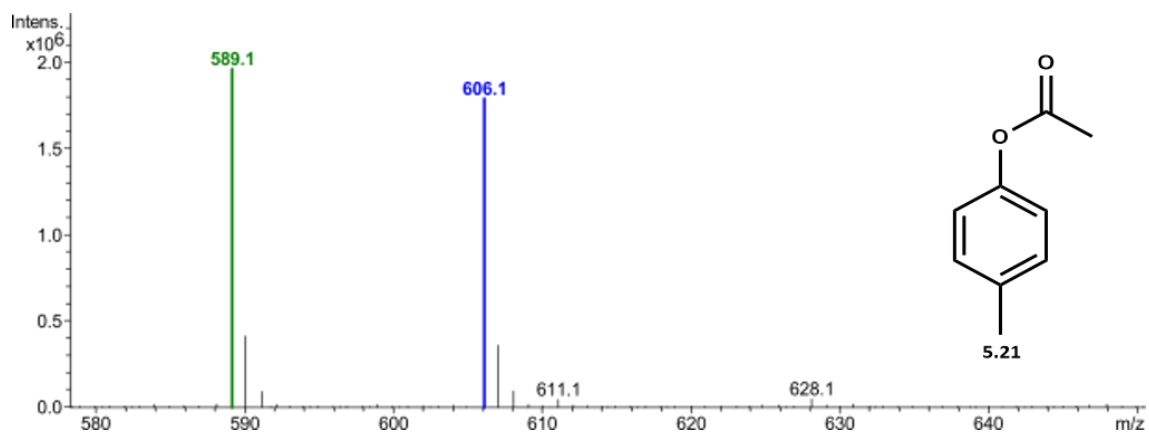




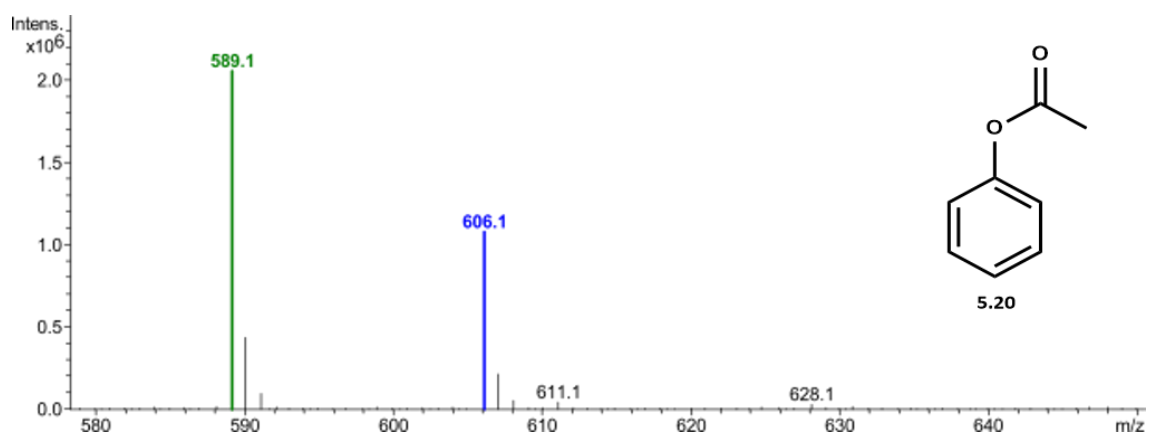
8.3 S-phenyl thioacetate (4.22) 10 mM



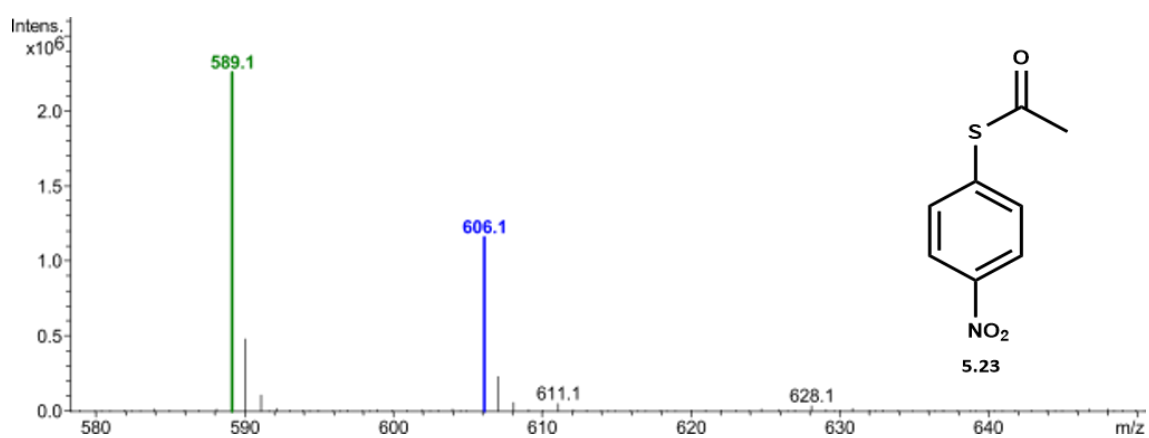
8.4 S-(p-tolyl)thioacetate 10 mM



8.5 p-tolyl acetate (4.21) 10 mM

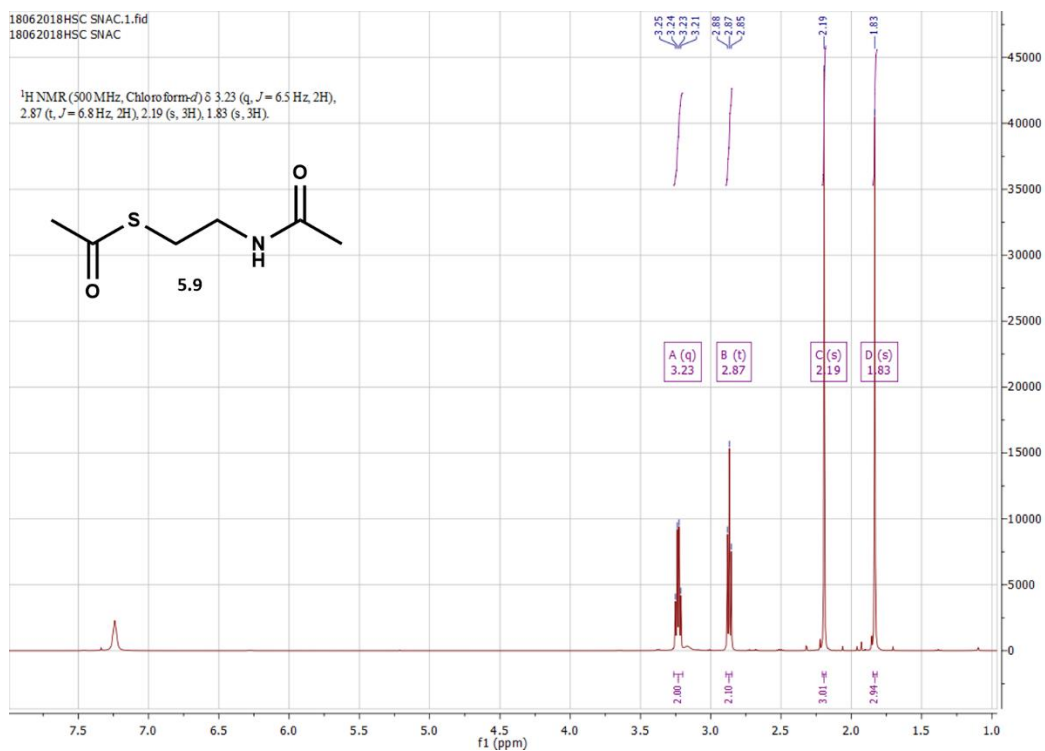


8.6 Phenyl acetate (4.20) 10 mM

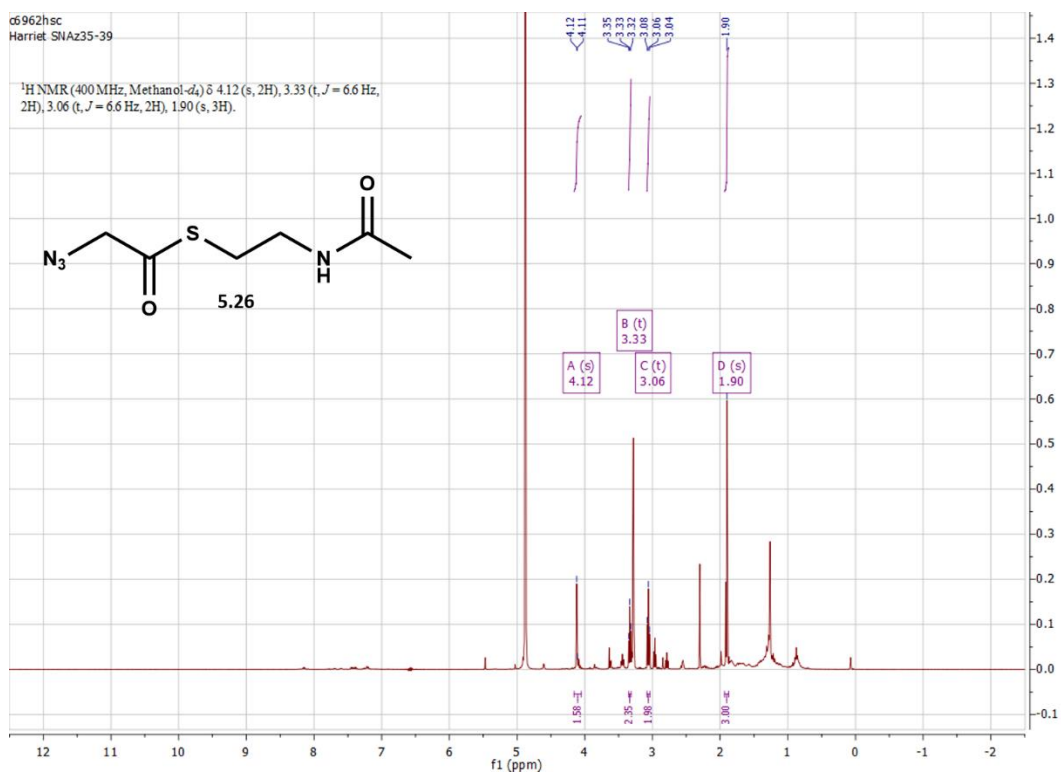


8.7 S-(4-nitrophenyl)thioacetate (4.24) 10 mM

## 9. NMR



### 9.1 *N*-acetylcysteamine thioacetate (5.9) proton NMR



### 9.2 *N*-acetylcysteamine thioazidoacetate (5.26) proton NMR

## List of Abbreviations

<b>Abs</b>	Absorbance
<b>Acetyl-CoA</b>	Acetyl-Coenzyme A
<b>ACS</b>	Acetyl-CoA synthetase
<b>Acyl-CoA</b>	Acyl-Coenzyme A
<b>ATP</b>	Adenosine Triphosphate
<b>cAMP</b>	Cyclic Adenosine Monophosphate
<b>CAT</b>	Carnitine acetyltransferase
<b>cDNA</b>	Complementary DNA
<b>CF</b>	Cystic Fibrosis
<b>CFG</b>	Consortium for Functional Glycomics
<b>CFTR</b>	Cystic Fibrosis Transmembrane conductance Regulator
<b>CMP-Neu5Ac</b>	Cytidine Monophosphate <i>N</i> -acetylneuraminic acid
<b>CMP-Pse5Ac7Ac</b>	Cytidine Monophosphate <i>N, N'</i> -diacetylpsseudaminic acid
<b>CoA</b>	Coenzyme-A
<b>CPS</b>	Capsular polysaccharide
<b>CTP</b>	Cytidine Triphosphate
<b>D<sub>2</sub>O</b>	Deuterium oxide
<b>DCC</b>	<i>N, N'</i> -dicyclohexylcarbodiimide
<b>DCM</b>	Dichloromethane
<b>DMAP</b>	4-Dimethylaminopyridine
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Deoxyribose nucleic acid
<b>ESI</b>	Electrospray Ionisation
<b>Et<sub>3</sub>N</b>	Triethylamine
<b>Equiv</b>	Equivalents
<b>GH</b>	Glycosyl hydrolase
<b>GlcN</b>	Glucosamine
<b>GlcN-1-P</b>	Glucosamine-1-phosphate
<b>GlcNAc</b>	<i>N</i> -acetylglucosamine
<b>GlcNAc-1-P</b>	<i>N</i> -acetylglucosamine-1-phosphate
<b>GNAT</b>	GCN5-related <i>N</i> -acetyltransferase
<b>GT</b>	Glycosyltransferase
<b>HCTU</b>	2-(6-Chloro-1 <i>H</i> -benzotriazole-1-yl)-1,1,3,3-tetramethylammonium hexafluorophosphate
<b>HOBt</b>	Hydroxybenzotriazole
<b>Hrs</b>	Hours
<b>IPTG</b>	Isopropyl β-D-1-thiogalactopyranoside
<b>Kb</b>	Kilobases
<b>kDa</b>	Kilodaltons
<b>KDN</b>	2-keto-3-deoxy- <i>D-glycero-D-galacto</i> -2-nonulosonic acid
<b>LB</b>	Lysogeny Broth
<b>LC-MS</b>	Liquid Chromatography-Mass Spectrometry
<b>Leg5Ac7Ac</b>	<i>N, N'</i> -diacetyl-legionaminic acid
<b>L-Glu</b>	L-Glutamate
<b>LPS</b>	Lipopolysaccharide
<b>ManNAc</b>	<i>N</i> -acetylmannosamine

<b>ManNAc-6-P</b>	<i>N</i> -acetylmannosamine-6-phosphate
<b>ManNAz</b>	<i>N</i> -azidoacetylmannosamine
<b>MeOH</b>	Methanol
<b>Mins</b>	Minutes
<b>μL</b>	Microlitre
<b>μM</b>	Micromolar
<b>mM</b>	Millimolar
<b>MS</b>	Mass spectrometry
<b>Mu-Neu5Ac</b>	4-Methylumbelliferone- $\alpha$ - <i>N</i> -acetylneuraminic acid
<b>MurNAc</b>	<i>N</i> -acetylmuramic acid
<b>NAD<sup>+</sup></b>	Nicotine Adenine Dinucleotide
<b>NADP<sup>+</sup></b>	Nicotine Adenine Dinucleotide Phosphate
<b>Neu5Ac</b>	5-acetamido-3,5-D- <i>glycero-D-galacto-2</i> -nonulosonic acid, <i>N</i> -acetylneuraminic acid
<b>Neu5Ac2en</b>	2-deoxy-2,3-dehydro- <i>N</i> -acetylneuraminic acid
<b>Neu5Az</b>	<i>N</i> -azidoacetylneuraminic acid
<b>Neu5Gc</b>	<i>N</i> -glycolylneuraminic acid
<b>NMR</b>	Nuclear Magnetic Resonance
<b>PCR</b>	Polymerase chain reaction
<b>PEG</b>	Poly(ethylene glycol)
<b>PEP</b>	Phosphoenolpyruvate
<b>PLP</b>	Pyridoxal 5'-phosphate
<b>PMP</b>	Pyridoxamine 5'-phosphate
<b>Pse5Ac7Ac</b>	5,7-diacetamido-3,5,7,9-tetra-deoxy-L- <i>glycero-L-manno-2</i> -nonulosonic acid, <i>N,N'</i> -diacetylpsseudaminic acid
<b>Pse5Ac7Az</b>	<i>N</i> -acetyl- <i>N'</i> -azidoacetylpsseudaminic acid
<b>Pse5Ac7RHb</b>	<i>N</i> -acetyl- <i>N'</i> -(3-hydroxybutanoyl)psseudaminic acid
<b>Pse5RHb7Fm</b>	<i>N</i> -(3-hydroxybutanoyl)- <i>N'</i> -formylpsseudaminic acid
<b>PTA</b>	Phosphotransacetylase
<b>R<sub>f</sub></b>	<i>Retention factor</i>
<b>Rpm</b>	Revolutions per minute
<b>SDR</b>	Short Chain Dehydrogenase/Reductase family
<b>SDS-PAGE</b>	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
<b>TLC</b>	Thin layer chromatography
<b>Tris</b>	Tris(hydroxymethyl)aminomethane
<b>UDP</b>	Uridine Diphosphate
<b>UDP-diNAcBac</b>	Uridine Diphosphate <i>N, N'</i> -diacetylbacillosamine
<b>UDP-GlcNAc</b>	Uridine Diphosphate <i>N</i> -acetylglucosamine
<b>UV</b>	Ultraviolet

## References

1. Gaskell, A.; Crennell, S.; Taylor, G., The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure* **1995**, *3* (11), 1197-205.
2. Zunk, M.; Kiefel, M. J., The occurrence and biological significance of the  $\alpha$ -keto-sugars pseudaminic acid and legionaminic acid within pathogenic bacteria. *RSC Adv.* **2014**, *4* (7), 3413-3421.
3. Xu, G.; Ryan, C.; Kiefel, M. J.; Wilson, J. C.; Taylor, G. L., Structural Studies on the *Pseudomonas aeruginosa* Sialidase-Like Enzyme PA2794 Suggest Substrate and Mechanistic Variations. *J. Mol. Biol.* **2009**, *386* (3), 828-840.
4. Schoenhofen, I. C.; Lunin, V. V.; Julien, J.-P.; Li, Y.; Ajamian, E.; Matte, A.; Cygler, M.; Brisson, J.-R.; Aubry, A.; Logan, S. M.; Bhatia, S.; Wakarchuk, W. W.; Young, N. M., Structural and Functional Characterization of PseC, an Aminotransferase Involved in the Biosynthesis of Pseudaminic Acid, an Essential Flagellar Modification in *Helicobacter pylori*. *J. Biol. Chem.* **2006**, *281* (13), 8907-8916.
5. Varki A, C. R., Esko JD, Freeze H, Hart GW, Marth JD., *Essentials of Glycobiology*. First ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 1999.
6. Vocadlo, D. J.; Davies, G. J., Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* **2008**, *12* (5), 539-555.
7. Laine, R. A., Invited Commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields  $1.05 \times 10^{12}$  structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **1994**, *4* (6), 759-767.
8. Krasnova, L.; Wong, C.-H., Understanding the Chemistry and Biology of Glycosylation with Glycan Synthesis. *Annu. Rev. Biochem* **2016**, *85* (1), 599-630.
9. Davies, G. J.; Gloster, T. M.; Henrissat, B., Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr. Opin. Struct. Biol.* **2005**, *15* (6), 637-645.
10. Pries, A. R.; Secomb, T. W.; Gaehtgens, P., The endothelial surface layer. *Pflügers Archiv* **2000**, *440* (5), 653-666.
11. Devuyst, O., Glycocalyx: The Fuzzy Coat Now Regulates Cell Signaling. *Peritoneal Dialysis International : Journal of the International Society for Peritoneal Dialysis* **2014**, *34* (6), 574-575.
12. Kornfeld, S.; Li, E.; Tabas, I., The synthesis of complex-type oligosaccharides. II. Characterization of the processing intermediates in the synthesis of the complex oligosaccharide units of the vesicular stomatitis virus G protein. *J. Biol. Chem.* **1978**, *253* (21), 7771-8.
13. Varki A, C. R., Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. , *Essentials of Glycobiology*. Second ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2009.
14. Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegthart, J. F. G.; Lütkeke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S., Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323-1324.

15. Nikaido, H., Structure and Functions of the Cell Envelope of Gram-Negative Bacteria. *Rev. Infect. Dis.* **1988**, *10*, S279-S281.
16. Austrian, R., THE GRAM STAIN AND THE ETIOLOGY OF LOBAR PNEUMONIA, AN HISTORICAL NOTE. *Bacteriol. Rev.* **1960**, *24* (3), 261-265.
17. Malanovic, N.; Lohner, K., Gram-positive bacterial cell envelopes: The impact on the activity of antimicrobial peptides. *Biochim. Biophys. Acta* **2016**, *1858* (5), 936-946.
18. LABISCHINSKI, H.; BARNICKEL, G.; BRADACZEK, H.; GIESBRECHT, P., On the Secondary and Tertiary Structure of Murein. *Eur. J. Biochem.* **1979**, *95* (1), 147-155.
19. Weidenmaier, C.; Peschel, A., Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature Reviews Microbiology* **2008**, *6*, 276.
20. Vollmer, W.; Bertsche, U., Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochim. Biophys. Acta* **2008**, *1778* (9), 1714-1734.
21. Abreu, M. T., Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. *Nature Reviews Immunology* **2010**, *10*, 131.
22. Ingalls, R. R.; Monks, B. G.; Golenbock, D. T., Membrane Expression of Soluble Endotoxin-binding Proteins Permits Lipopolysaccharide Signaling in Chinese Hamster Ovary Fibroblasts Independently of CD14. *J. Biol. Chem.* **1999**, *274* (20), 13993-13998.
23. Leive, L., THE BARRIER FUNCTION OF THE GRAM-NEGATIVE ENVELOPE. *Ann. N.Y. Acad. Sci.* **1974**, *235* (1), 109-129.
24. Nikaido, H., Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **2003**, *67* (4), 593-656.
25. MacGregor, D. R.; Elliker, P. R., A COMPARISON OF SOME PROPERTIES OF STRAINS OF PSEUDOMONAS AERUGINOSA SENSITIVE AND RESISTANT TO QUATERNARY AMMONIUM COMPOUNDS. *Can. J. Microbiol.* **1958**, *4* (5), 499-503.
26. Tamaki, S.; Sato, T.; Matsushashi, M., Role of Lipopolysaccharides in Antibiotic Resistance and Bacteriophage Adsorption of *Escherichia coli* K-12. *J. Bacteriol.* **1971**, *105* (3), 968-975.
27. Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B., The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37* (suppl\_1), D233-D238.
28. Henrissat, B., Glycosidase families. *Biochem. Soc. Trans.* **1998**, *26* (2), 153-156.
29. McCarter, J. D.; Stephen Withers, G., Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4* (6), 885-892.
30. Dwivedi, P.; Alavalapati, J. R. R.; Lal, P., Cellulosic ethanol production in the United States: Conversion technologies, current production status, economics, and emerging developments. *Energy for Sustainable Development* **2009**, *13* (3), 174-182.
31. Javier A. Linares-Pasten, M. A. a. E. N. K., Thermostable Glycoside Hydrolases in Biorefinery Technologies. *Current Biotechnology* **2014**, *3* (1), 26-44.
32. Viikari, L.; Kantelinen, A.; Sundquist, J.; Linko, M., Xylanases in bleaching: From an idea to the industry. *FEMS Microbiol. Rev.* **1994**, *13* (2-3), 335-350.

33. Reitinger, S.; Yu, Y.; Wicki, J.; Ludwiczek, M.; D'Angelo, I.; Baturin, S.; Okon, M.; Strynadka, N. C. J.; Lutz, S.; Withers, S. G.; McIntosh, L. P., Circular Permutation of *Bacillus circulans* Xylanase: A Kinetic and Structural Study. *Biochemistry* **2010**, *49* (11), 2464-2474.
34. Schwarzkopf, M.; Knobloch, K.-P.; Rohde, E.; Hinderlich, S.; Wiechens, N.; Lucka, L.; Horak, I.; Reutter, W.; Horstkorte, R., Sialylation is essential for early development in mice. *Proceedings of the National Academy of Sciences* **2002**, *99* (8), 5267-5270.
35. Crocker, P. R.; Paulson, J. C.; Varki, A., Siglecs and their roles in the immune system. *Nature Reviews Immunology* **2007**, *7*, 255.
36. Ströh, L. J.; Stehle, T., Glycan Engagement by Viruses: Receptor Switches and Specificity. *Annual Review of Virology* **2014**, *1* (1), 285-306.
37. Traving, C.; Schauer, R., Structure, function and metabolism of sialic acids. *Cellular and Molecular Life Sciences CMLS* **1998**, *54* (12), 1330-1349.
38. Varki, A., Diversity in the sialic acids. *Glycobiology* **1992**, *2* (1), 25-40.
39. Angata, T.; Varki, A., Chemical Diversity in the Sialic Acids and Related  $\alpha$ -Keto Acids: An Evolutionary Perspective. *Chem. Rev.* **2002**, *102* (2), 439-470.
40. Kim, J.-H.; Resende, R.; Wennekes, T.; Chen, H.-M.; Bance, N.; Buchini, S.; Watts, A. G.; Pilling, P.; Streltsov, V. A.; Petric, M.; Liggins, R.; Barrett, S.; McKimm-Breschkin, J. L.; Niikura, M.; Withers, S. G., Mechanism-Based Covalent Neuraminidase Inhibitors with Broad-Spectrum Influenza Antiviral Activity. *Science* **2013**, *340* (6128), 71-75.
41. Morley, T. J.; Willis, L. M.; Whitfield, C.; Wakarchuk, W. W.; Withers, S. G., A New Sialidase Mechanism: BACTERIOPHAGE K1F ENDO-SIALIDASE IS AN INVERTING GLYCOSIDASE. *J. Biol. Chem.* **2009**, *284* (26), 17404-17410.
42. Crennell, S.; Garman, E.; Laver, G.; Vimr, E.; Taylor, G., Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. *Structure* **1994**, *2* (6), 535-544.
43. Roggentin, P.; Rothe, B.; Kaper, J. B.; Galen, J.; Lawrisuk, L.; Vimr, E. R.; Schauer, R., Conserved sequences in bacterial and viral sialidases. *Glycoconjugate J.* **1989**, *6* (3), 349-353.
44. Crennell, S. J.; Garman, E. F.; Laver, W. G.; Vimr, E. R.; Taylor, G. L., Crystal structure of a bacterial sialidase (from *Salmonella typhimurium* LT2) shows the same fold as an influenza virus neuraminidase. *Proceedings of the National Academy of Sciences* **1993**, *90* (21), 9852-9856.
45. CHONG, A. K. J.; PEGG, M. S.; TAYLOR, N. R.; ITZSTEIN, M., Evidence for a sialosyl cation transition-state complex in the reaction of sialidase from influenza virus. *Eur. J. Biochem.* **1992**, *207* (1), 335-343.
46. Watts, A. G.; Damager, I.; Amaya, M. L.; Buschiazzi, A.; Alzari, P.; Frasch, A. C.; Withers, S. G., *Trypanosoma cruzi* Trans-sialidase Operates through a Covalent Sialyl-Enzyme Intermediate: Tyrosine Is the Catalytic Nucleophile. *J. Am. Chem. Soc.* **2003**, *125* (25), 7532-7533.
47. Amaya, M. a. F.; Watts, A. G.; Damager, I.; Wehenkel, A.; Nguyen, T.; Buschiazzi, A.; Paris, G.; Frasch, A. C.; Withers, S. G.; Alzari, P. M., Structural Insights into the Catalytic Mechanism of *Trypanosoma cruzi* trans-Sialidase. *Structure* **2004**, *12* (5), 775-784.
48. Dennis, J. W.; Granovsky, M.; Warren, C. E., Glycoprotein glycosylation and cancer progression. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1999**, *1473* (1), 21-34.



49. Hinek, A.; Pshezhetsky, A. V.; von Itzstein, M.; Starcher, B., Lysosomal Sialidase (Neuraminidase-1) Is Targeted to the Cell Surface in a Multiprotein Complex That Facilitates Elastic Fiber Assembly. *J. Biol. Chem.* **2006**, *281* (6), 3698-3710.
50. U. GRATA-BORKOWSKA, A. S., M. POKORSKI, J. DROBNIK, K. GŹSIOROWSKI, I. PIROGOWICZ, E. CIEGLAR-MARCZAK, EFFECTS OF NEURAMINIDASE ON APOPTOSIS OF BLOOD LYMPHOCYTES IN RATS WITH IMPLANTED MORRIS TUMOR. *J. Physiol. Pharmacol.* **2007**, *58* (Suppl 5), 253-262.
51. Prolo, L. M.; Vogel, H.; Reimer, R. J., The lysosomal sialic acid transporter sialin is required for normal CNS myelination. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **2009**, *29* (49), 15355.
52. Miyagi, T., Aberrant expression of sialidase and cancer progression. *Proceedings of the Japan Academy, Series B* **2008**, *84* (10), 407-418.
53. Bonardi, D.; Papini, N.; Pasini, M.; Dileo, L.; Orizio, F.; Monti, E.; Caimi, L.; Venerando, B.; Bresciani, R., Sialidase NEU3 Dynamically Associates to Different Membrane Domains Specifically Modifying Their Ganglioside Pattern and Triggering Akt Phosphorylation. *PLoS One* **2014**, *9* (6), e99405.
54. Miyagi, T.; Konno, K.; Emori, Y.; Kawasaki, H.; Suzuki, K.; Yasui, A.; Tsuik, S., Molecular cloning and expression of cDNA encoding rat skeletal muscle cytosolic sialidase. *J. Biol. Chem.* **1993**, *268* (35), 26435-26440.
55. Bonten, E.; van der Spoel, A.; Fornerod, M.; Grosveld, G.; d'Azzo, A., Characterization of human lysosomal neuraminidase defines the molecular basis of the metabolic storage disorder sialidosis. *Genes Dev.* **1996**, *10* (24), 3156-3169.
56. Rodriguez-Walker, M.; Daniotti, J. L., Human Sialidase Neu3 is S-Acylated and Behaves Like an Integral Membrane Protein. *Sci. Rep.* **2017**, *7* (1), 4167.
57. Yamaguchi, K.; Hata, K.; Koseki, K.; Shiozaki, K.; Akita, H.; Wada, T.; Moriya, S.; Miyagi, T., Evidence for mitochondrial localization of a novel human sialidase (NEU4). *Biochem. J* **2005**, *390* (1), 85-93.
58. Valaperta, R.; Chigorno, V.; Basso, L.; Prinetti, A.; Bresciani, R.; Preti, A.; Miyagi, T.; Sonnino, S., Plasma membrane production of ceramide from ganglioside GM3 in human fibroblasts. *The FASEB Journal* **2006**, *20* (8), 1227-1229.
59. Yogalingam, G.; Bonten, E. J.; van de Vlekkert, D.; Hu, H.; Moshich, S.; Connell, S. A.; d'Azzo, A., Neuraminidase 1 Is a Negative Regulator of Lysosomal Exocytosis. *Dev. Cell* **2008**, *15* (1), 74-86.
60. Andrews, N. W., Membrane repair and immunological danger. *EMBO reports* **2005**, *6* (9), 826.
61. Lillehoj, E. P.; Hyun, S. W.; Feng, C.; Zhang, L.; Liu, A.; Guang, W.; Nguyen, C.; Luzina, I. G.; Atamas, S. P.; Passaniti, A.; Twaddell, W. S.; Puché, A. C.; Wang, L.-X.; Cross, A. S.; Goldblum, S. E., NEU1 Sialidase Expressed in Human Airway Epithelia Regulates Epidermal Growth Factor Receptor (EGFR) and MUC1 Protein Signaling. *J. Biol. Chem.* **2012**, *287* (11), 8214-8231.
62. Duca, L.; Blanchevove, C.; Cantarelli, B.; Ghoneim, C.; Dedieu, S.; Delacoux, F.; Hornebeck, W.; Hinek, A.; Martiny, L.; Debelle, L., The Elastin Receptor Complex Transduces Signals through the Catalytic Activity of Its Neu-1 Subunit. *J. Biol. Chem.* **2007**, *282* (17), 12484-12491.

63. Seyrantepe, V.; Iannello, A.; Liang, F.; Kanshin, E.; Jayanth, P.; Samarani, S.; Szewczuk, M. R.; Ahmad, A.; Pshezhetsky, A. V., Regulation of Phagocytosis in Macrophages by Neuraminidase 1. *J. Biol. Chem.* **2010**, *285* (1), 206-215.
64. Stamatou, N. M.; Carubelli, I.; van de Vlekkert, D.; Bonten, E. J.; Papini, N.; Feng, C.; Venerando, B.; d'Azzo, A.; Cross, A. S.; Wang, L.-X.; Gornall, P. J., LPS-induced cytokine production in human dendritic cells is regulated by sialidase activity. *J. Leukocyte Biol.* **2010**, *88* (6), 1227-1239.
65. Chen, X. P.; Enioutina, E. Y.; Daynes, R. A., The control of IL-4 gene expression in activated murine T lymphocytes: a novel role for neu-1 sialidase. *The Journal of Immunology* **1997**, *158* (7), 3070-3080.
66. Seyrantepe, V.; Poupetova, H.; Froissart, R.; Zobot, M.-T.; Maire, I.; Pshezhetsky, A. V., Molecular pathology of NEU1 gene in sialidosis. *Hum. Mutat.* **2003**, *22* (5), 343-352.
67. Thomas, P. K.; Abrams, J. D.; Swallow, D.; Stewart, G., Sialidosis type 1: cherry red spot-myoclonus syndrome with sialidase deficiency and altered electrophoretic mobilities of some enzymes known to be glycoproteins. 1. Clinical findings. *Journal of Neurology, Neurosurgery & Psychiatry* **1979**, *42* (10), 873-880.
68. d'Azzo, A.; Machado, E.; Annunziata, I., Pathogenesis, Emerging therapeutic targets and Treatment in Sialidosis. *Expert opinion on orphan drugs* **2015**, *3* (5), 491-504.
69. d'Azzo, A.; Bonten, E., Molecular Mechanisms of Pathogenesis in a Glycosphingolipid and a Glycoprotein Storage Disease. *Biochem. Soc. Trans.* **2010**, *38* (6), 1453-1457.
70. Oliveira-Ferrer, L.; Legler, K.; Milde-Langosch, K., Role of protein glycosylation in cancer metastasis. *Semin. Cancer Biol.* **2017**, *44*, 141-152.
71. Rodrigues, E.; Macauley, M. S., Hypersialylation in Cancer: Modulation of Inflammation and Therapeutic Opportunities. *Cancers (Basel)* **2018**, *10* (6), 207.
72. Kato, K.; Shiga, K.; Yamaguchi, K.; Hata, K.; Kobayashi, T.; Miyazaki, K.; Saijo, S.; Miyagi, T., Plasma-membrane-associated sialidase (NEU3) differentially regulates integrin-mediated cell proliferation through laminin- and fibronectin-derived signalling. *Biochem. J* **2006**, *394* (Pt 3), 647-656.
73. Takahashi, K.; Hosono, M.; Sato, I.; Hata, K.; Wada, T.; Yamaguchi, K.; Nitta, K.; Shima, H.; Miyagi, T., Sialidase NEU3 contributes neoplastic potential on colon cancer cells as a key modulator of gangliosides by regulating Wnt signaling. *Int. J. Cancer* **2015**, *137* (7), 1560-1573.
74. King, S. J., Pneumococcal modification of host sugars: a major contributor to colonization of the human airway? *Mol. Oral Microbiol.* **2010**, *25* (1), 15-24.
75. PELTOLA, V. T.; MCCULLERS, J. A., Respiratory viruses predisposing to bacterial infections: role of neuraminidase. *The Pediatric Infectious Disease Journal* **2004**, *23* (1), S87-S97.
76. Manco, S.; Hernon, F.; Yesilkaya, H.; Paton, J. C.; Andrew, P. W.; Kadioglu, A., Pneumococcal Neuraminidases A and B Both Have Essential Roles during Infection of the Respiratory Tract and Sepsis. *Infect. Immun.* **2006**, *74* (7), 4014-4020.
77. Mally, M.; Shin, H.; Paroz, C.; Landmann, R.; Cornelis, G. R., Capnocytophaga canimorsus: A Human Pathogen Feeding at the Surface of Epithelial Cells and Phagocytes. *PLoS Path.* **2008**, *4* (9), e1000164.
78. Shtyrya, Y. A.; Mochalova, L. V.; Bovin, N. V., Influenza Virus Neuraminidase: Structure and Function. *Acta Naturae* **2009**, *1* (2), 26-32.

79. Matrosovich, M. N.; Matrosovich, T. Y.; Gray, T.; Roberts, N. A.; Klenk, H.-D., Neuraminidase Is Important for the Initiation of Influenza Virus Infection in Human Airway Epithelium. *J. Virol.* **2004**, *78* (22), 12665-12667.
80. Bouvier, N. M.; Palese, P., THE BIOLOGY OF INFLUENZA VIRUSES. *Vaccine* **2008**, *26* (Suppl 4), D49-D53.
81. Lipinski, C. A., Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies* **2004**, *1* (4), 337-341.
82. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Journal of Pharmaceutical Sciences* **1997**, *86* (1), 3-26. The article was originally published in *Advanced Drug Delivery Reviews* **1997**, *23* (1), 3-25. *Adv. Drug Del. Rev.* **2001**, *46* (1), 3-26.
83. EDMOND, J. D.; JOHNSTON, R. G.; KIDD, D.; RYLANCE, H. J.; SOMMERVILLE, R. G., THE INHIBITION OF NEURAMINIDASE AND ANTIVIRAL ACTION. *Br. J. Pharmacol. Chemother.* **1966**, *27* (2), 415-426.
84. Meindl, P.; Bodo, G.; Palese, P.; Schulman, J.; Tuppy, H., Inhibition of neuraminidase activity by derivatives of 2-deoxy-2,3-dehydro-N-acetylneuraminic acid. *Virology* **1974**, *58* (2), 457-463.
85. von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R., Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **1993**, *363*, 418.
86. Varghese, J. N.; Colman, P. M., Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 Å resolution. *J. Mol. Biol.* **1991**, *221* (2), 473-486.
87. Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C., Influenza Neuraminidase Inhibitors Possessing a Novel Hydrophobic Interaction in the Enzyme Active Site: Design, Synthesis, and Structural Analysis of Carbocyclic Sialic Acid Analogues with Potent Anti-Influenza Activity. *J. Am. Chem. Soc.* **1997**, *119* (4), 681-690.
88. Mendel, D. B.; Tai, C. Y.; Escarpe, P. A.; Li, W.; Sidwell, R. W.; Huffman, J. H.; Sweet, C.; Jakeman, K. J.; Merson, J.; Lacy, S. A.; Lew, W.; Williams, M. A.; Zhang, L.; Chen, M. S.; Bischofberger, N.; Kim, C. U., Oral Administration of a Prodrug of the Influenza Virus Neuraminidase Inhibitor GS 4071 Protects Mice and Ferrets against Influenza Infection. *Antimicrob. Agents Chemother.* **1998**, *42* (3), 640-646.
89. Mittal, R.; Aggarwal, S.; Sharma, S.; Chhibber, S.; Harjai, K., Urinary tract infections caused by *Pseudomonas aeruginosa*: A minireview. *Journal of Infection and Public Health* **2009**, *2* (3), 101-111.
90. Vincent, J.; Bihari, D. J.; Suter, P. M.; et al., The prevalence of nosocomial infection in intensive care units in Europe: Results of the European prevalence of infection in intensive care (epic) study. *JAMA* **1995**, *274* (8), 639-644.
91. Bodey, G. P.; Bolivar, R.; Fainstein, V.; Jadeja, L., Infections Caused by *Pseudomonas aeruginosa*. *Rev. Infect. Dis.* **1983**, *5* (2), 279-313.

92. Alvarez-Ortega, C.; Harwood, C. S., Responses of *Pseudomonas aeruginosa* to low oxygen indicate that growth in the cystic fibrosis lung is by aerobic respiration. *Mol. Microbiol.* **2007**, *65* (1), 153-165.
93. Wu, M.; Guina, T.; Brittnacher, M.; Nguyen, H.; Eng, J.; Miller, S. I., The *Pseudomonas aeruginosa* Proteome during Anaerobic Growth. *J. Bacteriol.* **2005**, *187* (23), 8185-8190.
94. Gaby, W. L., Study of Dissociative Behavior of *Pseudomonas aeruginosa*. *J. Bacteriol.* **1946**, *51* (2), 217-234.
95. Lister, P. D.; Wolter, D. J.; Hanson, N. D., Antibacterial-Resistant *Pseudomonas aeruginosa*: Clinical Impact and Complex Regulation of Chromosomally Encoded Resistance Mechanisms. *Clin. Microbiol. Rev.* **2009**, *22* (4), 582-610.
96. Reyes, E. A.; Bale, M. J.; Cannon, W. H.; Matsen, J. M., Identification of *Pseudomonas aeruginosa* by pyocyanin production on Tech agar. *J. Clin. Microbiol.* **1981**, *13* (3), 456-458.
97. Sousa, A. M.; Pereira, M. O., *Pseudomonas aeruginosa* Diversification during Infection Development in Cystic Fibrosis Lungs—A Review. *Pathogens* **2014**, *3* (3), 680-703.
98. Bendig, J. W.; Kyle, P. W.; Giangrande, P. L.; Samson, D. M.; Azadian, B. S., Two neutropenic patients with multiple resistant *Pseudomonas aeruginosa* septicaemia treated with ciprofloxacin. *J. R. Soc. Med.* **1987**, *80* (5), 316-317.
99. Lyczak, J. B.; Cannon, C. L.; Pier, G. B., Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist\*Address for correspondence: Channing Laboratory, 181 Longwood Avenue, Boston, MA 02115, USA. *Microb. Infect.* **2000**, *2* (9), 1051-1060.
100. Tramper-Stranders, G. A.; van der Ent, C. K.; Molin, S.; Yang, L.; Hansen, S. K.; Rau, M. H.; Ciofu, O.; Johansen, H. K.; Wolfs, T. F. W., Initial *Pseudomonas aeruginosa* infection in patients with cystic fibrosis: characteristics of eradicated and persistent isolates. *Clin. Microbiol. Infect.* **2012**, *18* (6), 567-574.
101. Allen, L.; Dockrell, D. H.; Pattery, T.; Lee, D. G.; Cornelis, P.; Hellewell, P. G.; Whyte, M. K. B., Pyocyanin Production by *Pseudomonas aeruginosa* Induces Neutrophil Apoptosis and Impairs Neutrophil-Mediated Host Defenses In Vivo. *The Journal of Immunology* **2005**, *174* (6), 3643-3649.
102. Schelstraete, P.; Haerynck, F.; Van daele, S.; Deseyne, S.; De Baets, F., Eradication therapy for *Pseudomonas aeruginosa* colonization episodes in cystic fibrosis patients not chronically colonized by *P. aeruginosa*. *J. Cyst. Fibros.* **2013**, *12* (1), 1-8.
103. and, M. B. M.; Bassler, B. L., Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.* **2001**, *55* (1), 165-199.
104. Smith, E. E.; Buckley, D. G.; Wu, Z.; Saenphimmachak, C.; Hoffman, L. R.; D'Argenio, D. A.; Miller, S. I.; Ramsey, B. W.; Speert, D. P.; Moskowitz, S. M.; Burns, J. L.; Kaul, R.; Olson, M. V., Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl. Acad. Sci.* **2006**, *103* (22), 8487-8492.
105. Hentzer, M.; Teitzel, G. M.; Balzer, G. J.; Heydorn, A.; Molin, S.; Givskov, M.; Parsek, M. R., Alginate overproduction affects *Pseudomonas aeruginosa* biofilm structure and function. *J. Bacteriol.* **2001**, *183* (18), 5395-5401.
106. Ramsey, D. M.; Wozniak, D. J., Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis. *Mol. Microbiol.* **2005**, *56* (2), 309-322.

107. DeVries, C. A.; Ohman, D. E., Mucoïd-to-nonmucoïd conversion in alginate-producing *Pseudomonas aeruginosa* often results from spontaneous mutations in algT, encoding a putative alternate sigma factor, and shows evidence for autoregulation. *J. Bacteriol.* **1994**, *176* (21), 6677-6687.
108. de la Fuente-Núñez, C.; Reffuveille, F.; Fernández, L.; Hancock, R. E. W., Bacterial biofilm development as a multicellular adaptation: antibiotic resistance and new therapeutic strategies. *Curr. Opin. Microbiol.* **2013**, *16* (5), 580-589.
109. Kirov, S. M.; Webb, J. S.; apos; May, C. Y.; Reid, D. W.; Woo, J. K. K.; Rice, S. A.; Kjelleberg, S., Biofilm differentiation and dispersal in mucoïd *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis. *Microbiology* **2007**, *153* (10), 3264-3274.
110. Hassett, D. J.; Korfhagen, T. R.; Irvin, R. T.; Schurr, M. J.; Sauer, K.; Lau, G. W.; Sutton, M. D.; Yu, H.; Hoiby, N., *Pseudomonas aeruginosa* biofilm infections in cystic fibrosis: insights into pathogenic processes and treatment strategies. *Expert Opin. Ther. Targets* **2010**, *14* (2), 117-130.
111. Kus, J. V.; Tullis, E.; Cvitkovitch, D. G.; Burrows, L. L., Significant differences in type IV pilin allele distribution among *Pseudomonas aeruginosa* isolates from cystic fibrosis (CF) versus non-CF patients. *Microbiology* **2004**, *150* (5), 1315-1326.
112. Oliver, A.; Cantón, R.; Campo, P.; Baquero, F.; Blázquez, J., High Frequency of Hypermutable *Pseudomonas aeruginosa* in Cystic Fibrosis Lung Infection. *Science* **2000**, *288* (5469), 1251-1253.
113. Maciá, M. D.; Blanquer, D.; Togores, B.; Sauleda, J.; Pérez, J. L.; Oliver, A., Hypermutation Is a Key Factor in Development of Multiple-Antimicrobial Resistance in *Pseudomonas aeruginosa* Strains Causing Chronic Lung Infections. *Antimicrob. Agents Chemother.* **2005**, *49* (8), 3382-3386.
114. Kiewitz, C.; Tümmler, B., Sequence Diversity of *Pseudomonas aeruginosa*: Impact on Population Structure and Genome Evolution. *J. Bacteriol.* **2000**, *182* (11), 3125-3135.
115. Frimmersdorf, E.; Horatzek, S.; Pelnikovich, A.; Wiehlmann, L.; Schomburg, D., How *Pseudomonas aeruginosa* adapts to various environments: a metabolomic approach. *Environ. Microbiol.* **2010**, *12* (6), 1734-1747.
116. Barth, A. L.; Pitt, T. L., Auxotrophic variants of *Pseudomonas aeruginosa* are selected from prototrophic wild-type strains in respiratory infections in patients with cystic fibrosis. *J. Clin. Microbiol.* **1995**, *33* (1), 37-40.
117. Xu, K. D.; Stewart, P. S.; Xia, F.; Huang, C.-T.; McFeters, G. A., Spatial Physiological Heterogeneity in *Pseudomonas aeruginosa* Biofilm Is Determined by Oxygen Availability. *Appl. Environ. Microbiol.* **1998**, *64* (10), 4035-4039.
118. Hassett, D. J.; Sokol, P. A.; Howell, M. L.; Ma, J. F.; Schweizer, H. T.; Ochsner, U.; Vasil, M. L., Ferric uptake regulator (Fur) mutants of *Pseudomonas aeruginosa* demonstrate defective siderophore-mediated iron uptake, altered aerobic growth, and decreased superoxide dismutase and catalase activities. *J. Bacteriol.* **1996**, *178* (14), 3996-4003.
119. Sugawara, E.; Nestorovich, E. M.; Bezrukov, S. M.; Nikaido, H., *Pseudomonas aeruginosa* Porin OprF Exists in Two Different Conformations. *J. Biol. Chem.* **2006**, *281* (24), 16220-16229.
120. Elmer, H. L.; Brady, K. G.; Drumm, M. L.; Kelley, T. J., Nitric oxide-mediated regulation of transepithelial sodium and chloride transport in murine nasal epithelium. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **1999**, *276* (3), L466-L473.

121. Amitani, R.; Wilson, R.; Rutman, A.; Read, R.; Ward, C.; Burnett, D.; Stockley, R. A.; Cole, P. J., Effects of Human Neutrophil Elastase and *Pseudomonas aeruginosa* Proteinases on Human Respiratory Epithelium. *Am. J. Respir. Cell Mol. Biol.* **1991**, *4* (1), 26-32.
122. Kim, D.-M.; Swartz, J. R., Regeneration of adenosine triphosphate from glycolytic intermediates for cell-free protein synthesis. *Biotechnol. Bioeng.* **2001**, *74* (4), 309-316.
123. Liu, B.; Knirel, Y. A.; Feng, L.; Perepelov, A. V.; Senchenkova, S. y. N.; Wang, Q.; Reeves, P. R.; Wang, L., Structure and genetics of Shigella O antigens. *FEMS Microbiol. Rev.* **2008**, *32* (4), 627-653.
124. Anderson, M.; Rich, D.; Gregory, R.; Smith, A.; Welsh, M., Generation of cAMP-activated chloride currents by expression of CFTR. *Science* **1991**, *251* (4994), 679-682.
125. Widdicombe, J. H.; Welsh, M. J.; Finkbeiner, W. E., Cystic fibrosis decreases the apical membrane chloride permeability of monolayers cultured from cells of tracheal epithelium. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (18), 6167-6171.
126. Cai, L.; Guan, W.; Kitaoka, M.; Shen, J.; Xia, C.; Chen, W.; Wang, P. G., A chemoenzymatic route to N-acetylglucosamine-1-phosphate analogues: substrate specificity investigations of N-acetylhexosamine 1-kinase. *Chem. Commun.* **2009**, (20), 2944-2946.
127. Quinton, P. M., Cystic fibrosis: a disease in electrolyte transport. *The FASEB Journal* **1990**, *4* (10), 2709-17.
128. Rommens, J. M.; Iannuzzi, M. C.; Kerem, B.-s.; Drumm, M. L.; Melmer, G.; Dean, M.; Rozmahel, R.; Cole, J. L.; Kennedy, D.; Hidaka, N.; Zsiga, M.; Buchwald, M.; Riordan, J. R.; Tsui, L.-C.; Collins, F. S., Identification of the Cystic Fibrosis Gene: Chromosome Walking and Jumping. *Science* **1989**, *245* (4922), 1059-1065.
129. Riordan, J.; Rommens, J.; Kerem, B.; Alon, N.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.; Plavsic, N.; Chou, J.; al., e., Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **1989**, *245* (4922), 1066-1073.
130. Kerem, B.-s.; Rommens, J. M.; Buchanan, J. A.; Markiewicz, D.; Cox, T. K.; Chakravarti, A.; Buchwald, M.; Tsui, L.-C., Identification of the Cystic Fibrosis Gene: Genetic Analysis. *Science* **1989**, *245* (4922), 1073-1080.
131. Tsui, L.; Buchwald, M.; Barker, D.; Braman, J.; Knowlton, R.; Schumm, J.; Eiberg, H.; Mohr, J.; Kennedy, D.; Plavsic, N.; et, a., Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science* **1985**, *230* (4729), 1054-1057.
132. Gerlach, J. H.; Endicott, J. A.; Juranka, P. F.; Henderson, G.; Sarangi, F.; Deuchars, K. L.; Ling, V., Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance. *Nature* **1986**, *324* (6096), 485-489.
133. Bear, C. E.; Duguay, F.; Naismith, A. L.; Kartner, N.; Hanrahan, J. W.; Riordan, J. R., Cl-channel activity in *Xenopus* oocytes expressing the cystic fibrosis gene. *J. Biol. Chem.* **1991**, *266* (29), 19142-5.
134. Thomas, P. J.; Shenbagamurthi, P.; Sondek, J.; Hulihan, J. M.; Pedersen, P. L., The cystic fibrosis transmembrane conductance regulator. Effects of the most common cystic fibrosis-causing mutation on the secondary structure and stability of a synthetic peptide. *J. Biol. Chem.* **1992**, *267* (9), 5727-30.
135. Tsui, L.-C., The spectrum of cystic fibrosis mutations. *Trends Genet.* **1992**, *8* (11), 392-398.

136. Drumm, M.; Wilkinson, D.; Smit, L.; Worrell, R.; Strong, T.; Frizzell, R.; Dawson, D.; Collins, F., Chloride conductance expressed by delta F508 and other mutant CFTRs in *Xenopus* oocytes. *Science* **1991**, *254* (5039), 1797-1799.
137. Li, C.; Ramjeesingh, M.; Reyes, E.; Jensen, T.; Chang, X.; Rommens, J. M.; Bear, C. E., The cystic fibrosis mutation ([Delta]F508) does not influence the chloride channel activity of CFTR. *Nat. Genet.* **1993**, *3* (4), 311-316.
138. Decaestecker, K.; Decaestecker, E.; Castellani, C.; Jaspers, M.; Cuppens, H.; De Boeck, K., Genotype/phenotype correlation of the G85E mutation in a large cohort of cystic fibrosis patients. *Eur. Respir. J.* **2004**, *23* (5), 679-684.
139. Brodlie, M.; Haq, I. J.; Roberts, K.; Elborn, J. S., Targeted therapies to improve CFTR function in cystic fibrosis. *Genome Med.* **2015**, *7*, 101.
140. Ramsey, B. W.; Dorkin, H. L., Consensus conference: Practical applications of pulmozyme® september 22, 1993. *Pediatr. Pulmonol.* **1994**, *17* (6), 404-408.
141. Elkins, M. R.; Robinson, M.; Rose, B. R.; Harbour, C.; Moriarty, C. P.; Marks, G. B.; Belousova, E. G.; Xuan, W.; Bye, P. T. P., A Controlled Trial of Long-Term Inhaled Hypertonic Saline in Patients with Cystic Fibrosis. *New Engl. J. Med.* **2006**, *354* (3), 229-240.
142. Landau, L. I.; Phelan, P. D., The variable effect of a bronchodilating agent on pulmonary function in cystic fibrosis. *The Journal of Pediatrics* **1973**, *82* (5), 863-868.
143. Borowitz, D. S.; Grand, R. J.; Durie, P. R., Use of pancreatic enzyme supplements for patients with cystic fibrosis in the context of fibrosing colonopathy. *The Journal of Pediatrics* **1995**, *127* (5), 681-684.
144. Rich, D. P.; Anderson, M. P.; Gregory, R. J.; Cheng, S. H.; Paul, S.; Jefferson, D. M.; McCann, J. D.; Klinger, K. W.; Smith, A. E.; Welsh, M. J., Expression of cystic fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells. *Nature* **1990**, *347* (6291), 358-363.
145. Maeder, M. L.; Thibodeau-Beganny, S.; Osiaik, A.; Wright, D. A.; Anthony, R. M.; Eichinger, M.; Jiang, T.; Foley, J. E.; Winfrey, R. J.; Townsend, J. A.; Unger-Wallace, E.; Sander, J. D.; Müller-Lerch, F.; Fu, F.; Pearlberg, J.; Göbel, C.; Dassie, J. P.; Pruett-Miller, S. M.; Porteus, M. H.; Sgroi, D. C.; lafrate, A. J.; Dobbs, D.; McCray, P. B.; Cathomen, T.; Voytas, D. F.; Joung, J. K., Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell* **2008**, *31* (2), 294-301.
146. Van Goor, F.; Hadida, S.; Grootenhuys, P. D. J.; Burton, B.; Cao, D.; Neuberger, T.; Turnbull, A.; Singh, A.; Joubran, J.; Hazlewood, A.; Zhou, J.; McCartney, J.; Arumugam, V.; Decker, C.; Yang, J.; Young, C.; Olson, E. R.; Wine, J. J.; Frizzell, R. A.; Ashlock, M.; Negulescu, P., Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (44), 18825-18830.
147. Orenstein, D. M.; Kaplan, R. M., MEasuring the quality of well-being in cystic fibrosis and lung transplantation. the importance of the area under the curve. *Chest* **1991**, *100* (4), 1016-1018.
148. Zaidi, T. S.; Lyczak, J.; Preston, M.; Pier, G. B., Cystic Fibrosis Transmembrane Conductance Regulator-Mediated Corneal Epithelial Cell Ingestion of *Pseudomonas aeruginosa* Is a Key Component in the Pathogenesis of Experimental Murine Keratitis. *Infect. Immun.* **1999**, *67* (3), 1481-1492.

149. Guan, W.; Cai, L.; Wang, P. G., Highly Efficient Synthesis of UDP-GalNAc/GlcNAc Analogues with Promiscuous Recombinant Human UDP-GalNAc Pyrophosphorylase AGX1. *Chemistry – A European Journal* **2010**, *16* (45), 13343-13345.
150. al., J. L. J. e., *Comparative Genomics of Four Pseudomonas Species*. In: Ramos JL. (eds) *Pseudomonas*. Springer: Boston, MA, 2004.
151. Wiehlmann, L.; Wagner, G.; Cramer, N.; Siebert, B.; Gudowius, P.; Morales, G.; Köhler, T.; van Delden, C.; Weinel, C.; Slickers, P.; Tümmler, B., Population structure of *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* **2007**, *104* (19), 8101-8106.
152. He, J.; Baldini, R. L.; Déziel, E.; Saucier, M.; Zhang, Q.; Liberati, N. T.; Lee, D.; Urbach, J.; Goodman, H. M.; Rahme, L. G., The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (8), 2530-2535.
153. Scott, F. W.; Pitt, T. L., Identification and characterization of transmissible *Pseudomonas aeruginosa* strains in cystic fibrosis patients in England and Wales. *J. Med. Microbiol.* **2004**, *53* (7), 609-615.
154. Mandsberg, L. F.; Ciofu, O.; Kirkby, N.; Christiansen, L. E.; Poulsen, H. E.; Høiby, N., Antibiotic Resistance in *Pseudomonas aeruginosa* Strains with Increased Mutation Frequency Due to Inactivation of the DNA Oxidative Repair System. *Antimicrob. Agents Chemother.* **2009**, *53* (6), 2483-2491.
155. McCaslin, C. A.; Petrusca, D. N.; Poirier, C.; Serban, K. A.; Anderson, G. G.; Petrache, I., Impact of alginate-producing *Pseudomonas aeruginosa* on alveolar macrophage apoptotic cell clearance. *Journal of Cystic Fibrosis* **2015**, *14* (1), 70-77.
156. Lavoie, E. G.; Wangdi, T.; Kazmierczak, B. I., Innate immune responses to *Pseudomonas aeruginosa* infection. *Microb. Infect.* **2011**, *13* (14), 1133-1145.
157. Grandjean Lapierre, S.; Phelippeau, M.; Hakimi, C.; Didier, Q.; Reynaud-Gaubert, M.; Dubus, J.-C.; Drancourt, M., Cystic fibrosis respiratory tract salt concentration: An Exploratory Cohort Study. *Medicine* **2017**, *96* (47), e8423.
158. Conover, J. H.; Bonforte, R. J.; Hathaway, P.; Paciuc, S.; Conod, E. J.; Hirschhorn, K.; Kopel, F. B., Studies on Ciliary Dyskinesia Factor in Cystic Fibrosis. I. Bioassay and Heterozygote Detection in Serum. *Pediatr. Res.* **1973**, *7*, 220.
159. Lee, T. W. R.; Brownlee, K. G.; Conway, S. P.; Denton, M.; Littlewood, J. M., Evaluation of a new definition for chronic *Pseudomonas aeruginosa* infection in cystic fibrosis patients. *J. Cyst. Fibros.* **2003**, *2* (1), 29-34.
160. Rogers, G. B.; Skelton, S.; Serisier, D. J.; van der Gast, C. J.; Bruce, K. D., Determining Cystic Fibrosis-Affected Lung Microbiology: Comparison of Spontaneous and Serially Induced Sputum Samples by Use of Terminal Restriction Fragment Length Polymorphism Profiling. *J. Clin. Microbiol.* **2010**, *48* (1), 78-86.
161. Driscoll, J. A.; Brody, S. L.; Kollef, M. H., The Epidemiology, Pathogenesis and Treatment of *Pseudomonas aeruginosa* Infections. *Drugs* **2007**, *67* (3), 351-368.
162. Davis, B. D., Mechanism of bactericidal action of aminoglycosides. *Microbiol. Rev.* **1987**, *51* (3), 341-350.



163. Bumann, D.; Behre, C.; Behre, K.; Herz, S.; Gewecke, B.; Gessner, J. E.; von Specht, B. U.; Baumann, U., Systemic, nasal and oral live vaccines against *Pseudomonas aeruginosa*: a clinical trial of immunogenicity in lower airways of human volunteers. *Vaccine* **2010**, *28* (3), 707-713.
164. Döring, G.; Meisner, C.; Stern, M., A double-blind randomized placebo-controlled phase III study of a *Pseudomonas aeruginosa* flagella vaccine in cystic fibrosis patients. *Proceedings of the National Academy of Sciences* **2007**, *104* (26), 11020-11025.
165. Singh, P. K.; Schaefer, A. L.; Parsek, M. R.; Moninger, T. O.; Welsh, M. J.; Greenberg, E. P., Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **2000**, *407*, 762.
166. Walters, M. C.; Roe, F.; Bugnicourt, A.; Franklin, M. J.; Stewart, P. S., Contributions of Antibiotic Penetration, Oxygen Limitation, and Low Metabolic Activity to Tolerance of *Pseudomonas aeruginosa* Biofilms to Ciprofloxacin and Tobramycin. *Antimicrob. Agents Chemother.* **2003**, *47* (1), 317-323.
167. Winstanley, C.; O'Brien, S.; Brockhurst, M. A., *Pseudomonas aeruginosa* Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections. *Trends Microbiol.* **2016**, *24* (5), 327-337.
168. Singh, P. K.; Schaefer, A. L.; Parsek, M. R.; Moninger, T. O.; Welsh, M. J.; Greenberg, E. P., Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature (London)* **2000**, *407* (6805), 762-764.
169. Stewart, P. S., Theoretical aspects of antibiotic diffusion into microbial biofilms. *Antimicrob. Agents Chemother.* **1996**, *40* (11), 2517-2522.
170. Lanotte, P.; Watt, S.; Mereghetti, L.; Dartiguelongue, N.; Rastegar-Lari, A.; Goudeau, A.; Quentin, R., Genetic features of *Pseudomonas aeruginosa* isolates from cystic fibrosis patients compared with those of isolates from other origins. *J. Med. Microbiol.* **2004**, *53* (1), 73-81.
171. Soong, G.; Muir, A.; Gomez, M. I.; Waks, J.; Reddy, B.; Planet, P.; Singh, P. K.; Kanetko, Y.; Wolfgang, M. C.; Hsiao, Y.-S.; Tong, L.; Prince, A., Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. *The Journal of Clinical Investigation* **2006**, *116* (8), 2297-2305.
172. Wolska, K.; Kot, B.; Mioduszezewska, H.; Sempruch, C.; Borkowska, L.; Rymuza, K., Occurrence of the nan1 gene and adhesion of *Pseudomonas aeruginosa* isolates to human buccal epithelial cells. **2012**, *49* (1), 59.
173. Soong, G.; Muir, A.; Gomez, M. I.; Waks, J.; Reddy, B.; Planet, P.; Singh, P. K.; Kanetko, Y.; Wolfgang, M. C.; Hsiao, Y.-S.; Tong, L.; Prince, A., Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. *The Journal of Clinical Investigation* *116* (8), 2297-2305.
174. Taylor, G., Sialidases: structures, biological significance and therapeutic potential. *Curr. Opin. Struct. Biol.* **1996**, *6* (6), 830-837.
175. Wozniak, D. J.; Wyckoff, T. J. O.; Starkey, M.; Keyser, R.; Azadi, P.; O'Toole, G. A.; Parsek, M. R., Alginate is not a significant component of the extracellular polysaccharide matrix of PA14 and PAO1 *Pseudomonas aeruginosa* biofilms. *Proceedings of the National Academy of Sciences* **2003**, *100* (13), 7907-7912.
176. Krivan, H. C.; Roberts, D. D.; Ginsburg, V., Many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence GalNAc beta 1-4Gal found in some glycolipids. *Proceedings of the National Academy of Sciences* **1988**, *85* (16), 6157-6161.

177. Rajan, S.; Bryan, R.; Ratner, A. J.; Prince, A., *Pseudomonas aeruginosa* Interactions with Epithelial Cells: Adherence, Invasion and Apoptosis ◆ 896. *Pediatr. Res.* **1998**, *43*, 155.
178. Knirel, Y. A.; Vinogradov, E. V.; L'Vov, V. L.; Kocharova, N. A.; Shashkov, A. S.; Dmitriev, B. A.; Kochetkov, N. K., Sialic acids of a new type from the lipopolysaccharides of *Pseudomonas aeruginosa* and *Shigella boydii*. *Carbohydr. Res.* **1984**, *133* (2), C5-C8.
179. Schirm, M.; Schoenhofen, I. C.; Logan, S. M.; Waldron, K. C.; Thibault, P., Identification of Unusual Bacterial Glycosylation by Tandem Mass Spectrometry Analyses of Intact Proteins. *Anal. Chem.* **2005**, *77* (23), 7774-7782.
180. Schirm, M.; Soo, E. C.; Aubry, A. J.; Austin, J.; Thibault, P.; Logan, S. M., Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol. Microbiol.* **2003**, *48* (6), 1579-1592.
181. Hitchen, P.; Brzostek, J.; Panico, M.; Butler, J. A.; Morris, H. R.; Dell, A.; Linton, D., Modification of the *Campylobacter jejuni* flagellin glycan by the product of the Cj1295 homopolymeric-tract-containing gene. *Microbiology* **2010**, *156* (Pt 7), 1953-62.
182. Li, Z.; Hwang, S.; Ericson, J.; Bowler, K.; Bar-Peled, M., Pen and Pal are nucleotide-sugar dehydratases that convert UDP-GlcNAc to UDP-6-deoxy-D-GlcNAc-5,6-ene and then to UDP-4-keto-6-deoxy-L-AltNAc for CMP-pseudaminic acid synthesis in *Bacillus thuringiensis*. *J. Biol. Chem.* **2015**, *290* (2), 691-704.
183. Thibault, P.; Logan, S. M.; Kelly, J. F.; Brisson, J.-R.; Ewing, C. P.; Trust, T. J.; Guerry, P., Identification of the Carbohydrate Moieties and Glycosylation Motifs in *Campylobacter jejuni* Flagellin. *J. Biol. Chem.* **2001**, *276* (37), 34862-34870.
184. KNIREL, Y. A.; VINOGRADOV, E. V.; SHASHKOV, A. S.; DMITRIEV, B. A.; KOCHETKOV, N. K.; STANISLAVSKY, E. S.; MASHILOVA, G. M., Somatic antigens of *Pseudomonas aeruginosa*. *Eur. J. Biochem.* **1987**, *163* (3), 627-637.
185. Staaf, M.; Weintraub, A.; Widmalm, G., Structure determination of the O-antigenic polysaccharide from the enteroinvasive *Escherichia coli* O136. *Eur. J. Biochem.* **1999**, *263* (3), 656-661.
186. GIL-SERRANO, A. M.; RODRÍGUEZ-CARVAJAL, M. A.; TEJERO-MATEO, P.; ESPARTERO, J. L.; MENENDEZ, M.; CORZO, J.; RUIZ-SAINZ, J. E.; BUENDÍA-CLAVERÍA, A. M., Structural determination of a 5-acetamido-3,5,7,9-tetra-deoxy-7-(3-hydroxybutyramido)-L-glycero-L-manno-nonulosonic acid-containing homopolysaccharide isolated from *Sinorhizobium fredii* HH103. *Biochem. J* **1999**, *342* (3), 527-535.
187. Castric, P.; Cassels, F. J.; Carlson, R. W., Structural Characterization of the *Pseudomonas aeruginosa* 1244 Pilin Glycan. *J. Biol. Chem.* **2001**, *276* (28), 26479-26485.
188. Knirel, Y. A.; Kocharova, N. A.; Shashkov, A. S.; Dmitriev, B. A.; Kochetkov, N. K.; Stanislavskii, E. S.; Mashilova, G. M., Somatic antigens of *Pseudomonas aeruginosa*. The structure of O-specific polysaccharide chains of the lipopolysaccharides from *P. aeruginosa* O5 (Lanyi) and immunotype 6 (Fisher). *Eur. J. Biochem.* **1987**, *163* (3), 639-52.
189. Hopf, P. S.; Ford, R. S.; Zebian, N.; Merckx-Jacques, A.; Vijayakumar, S.; Ratnayake, D.; Hayworth, J.; Creuzenet, C., Protein glycosylation in *Helicobacter pylori*: beyond the flagellins? *PLoS One* **2011**, *6* (9), e25722.
190. Ewing, C. P.; Andreishcheva, E.; Guerry, P., Functional Characterization of Flagellin Glycosylation in *Campylobacter jejuni* 81-176. *J. Bacteriol.* **2009**, *191* (22), 7086-7093.

191. Ottemann, K. M.; Lowenthal, A. C., *Helicobacter pylori* Uses Motility for Initial Colonization and To Attain Robust Infection. *Infect. Immun.* **2002**, *70* (4), 1984-1990.
192. Baldvinsson, S. B.; Sørensen, M. C. H.; Vegge, C. S.; Clokie, M. R.; Brøndsted, L., *Campylobacter jejuni* motility is required for infection by the flagellotropic bacteriophage F341. *Appl. Environ. Microbiol.* **2014**, AEM. 02057-14.
193. Andersen-Nissen, E.; Smith, K. D.; Strobe, K. L.; Barrett, S. L. R.; Cookson, B. T.; Logan, S. M.; Aderem, A., Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (26), 9247-9252.
194. Stephenson, H. N.; Jones, H.; Milioris, E.; Copland, A.; Bajaj-Elliott, M.; Mills, D. C.; Dorrell, N.; Wren, B. W.; Crocker, P. R.; Escors, D., Pseudaminic acid on *Campylobacter jejuni* flagella modulates dendritic cell IL-10 expression via Siglec-10 receptor: a novel flagellin-host interaction. *J. Infect. Dis.* **2014**, *210* (9), 1487-98.
195. Tsvetkov, Y. E.; Shashkov, A. S.; Knirel, Y. A.; Zähringer, U., Synthesis and identification in bacterial lipopolysaccharides of 5,7-diacetamido-3,5,7,9-tetradeoxy-d-glycero-d-galacto- and -d-glycero-d-talo-non-2-ulosonic acids. *Carbohydr. Res.* **2001**, *331* (3), 233-237.
196. Lee, Y. J.; Kubota, A.; Ishiwata, A.; Ito, Y., Synthesis of pseudaminic acid, a unique nonulopyranoside derived from pathogenic bacteria through 6-deoxy-AltdiNAc. *Tetrahedron Lett.* **2011**, *52* (3), 418-421.
197. Zunk, M.; Williams, J.; Carter, J.; Kiefel, M. J., A new approach towards the synthesis of pseudaminic acid analogues. *Org. Biomol. Chem.* **2014**, *12* (18), 2918-2925.
198. Williams, J. T.; Corcilius, L.; Kiefel, M. J.; Payne, R. J., Total Synthesis of Native 5,7-Diacetyl-pseudaminic Acid from N-Acetylneuraminic Acid. *The Journal of Organic Chemistry* **2016**, *81* (6), 2607-2611.
199. Liu, H.; Zhang, Y.; Wei, R.; Andolina, G.; Li, X., Total Synthesis of *Pseudomonas aeruginosa* 1244 Pilin Glycan via de Novo Synthesis of Pseudaminic Acid. *J. Am. Chem. Soc.* **2017**, *139* (38), 13420-13428.
200. Rotstein, B. H.; Winternheimer, D. J.; Yin, L. M.; Deber, C. M.; Yudin, A. K., Thioester-isocyanides: versatile reagents for the synthesis of cycle-tail peptides. *Chem. Commun.* **2012**, *48* (31), 3775-3777.
201. Chan, T.-H.; Li, C.-J., A concise chemical synthesis of (+)-3-deoxy-D-glycero-D-galactononulosonic acid (KDN). *J. Chem. Soc., Chem. Commun.* **1992**, (10), 747-748.
202. Szymanski, C. M.; Yao, R.; Ewing, C. P.; Trust, T. J.; Guerry, P., Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **1999**, *32* (5), 1022-1030.
203. Schoenhofen, I. C.; McNally, D. J.; Brisson, J.-R.; Logan, S. M., Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* **2006**, *16* (9), 8C-14C.
204. Davies, G.; Henrissat, B., Structures and mechanisms of glycosyl hydrolases. *Structure* **1995**, *3* (9), 853-859.
205. Rui Shen, S. W., Xiaofeng Ma, Junyang Xian, Jing Li, Lianwen Zhang, and Peng Wang, An Easy Colorimetric Assay for Glycosyltransferases. *Biochemistry (Moscow)* **2010**, *75* (7), 944-950.
206. Lalégerie, P.; Legler, G.; Yon, J. M., The use of inhibitors in the study of glycosidases. *Biochimie* **1982**, *64* (11), 977-1000.

207. King, S. J.; Hippe, K. R.; Weiser, J. N., Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol. Microbiol.* **2006**, *59* (3), 961-974.
208. Varki, A.; Gagneux, P., Multifarious roles of sialic acids in immunity. *Ann. N.Y. Acad. Sci.* **2012**, *1253* (1), 16-36.
209. Varki, A., Sialic acids in human health and disease. *Trends Mol. Med.* **2008**, *14* (8), 351-360.
210. Blessia, T. F.; Rapheal, V. S.; Sharmila, D. J. S., Molecular Dynamics of Sialic Acid Analogues and their Interaction with Influenza Hemagglutinin. *Indian J. Pharm. Sci.* **2010**, *72* (4), 449-457.
211. Towler, D. A.; Eubanks, S. R.; Towery, D. S.; Adams, S. P.; Glaser, L., Amino-terminal processing of proteins by N-myristoylation. Substrate specificity of N-myristoyl transferase. *J. Biol. Chem.* **1987**, *262* (3), 1030-6.
212. Vimr, E. R.; Kalivoda, K. A.; Deszo, E. L.; Steenbergen, S. M., Diversity of Microbial Sialic Acid Metabolism. *Microbiol. Mol. Biol. Rev.* **2004**, *68* (1), 132-153.
213. Spiwok, V.; Tvaroška, I., Conformational Free Energy Surface of  $\alpha$ -N-Acetylneuraminic Acid: An Interplay Between Hydrogen Bonding and Solvation. *The Journal of Physical Chemistry B* **2009**, *113* (28), 9589-9594.
214. Warren, L., [67] Thiobarbituric acid assay of sialic acids. In *Methods Enzymol.*, Academic Press: 1963; Vol. 6, pp 463-465.
215. Myers, R. W.; Lee, R. T.; Lee, Y. C.; Thomas, G. H.; Reynolds, L. W.; Uchida, Y., The synthesis of 4-methylumbelliferyl  $\alpha$ -ketoside of N-acetylneuraminic acid and its use in a fluorometric assay for neuraminidase. *Anal. Biochem.* **1980**, *101* (1), 166-174.
216. Salah Ud-Din, A. I. M.; Roujeinikova, A., Flagellin glycosylation with pseudaminic acid in *Campylobacter* and *Helicobacter*: prospects for development of novel therapeutics. *Cell. Mol. Life Sci.* **2018**, *75* (7), 1163-1178.
217. Parker, J. L.; Day-Williams, M. J.; Tomas, J. M.; Stafford, G. P.; Shaw, J. G., Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *MicrobiologyOpen* **2012**, *1* (2), 149-160.
218. Kenyon, J. J.; Marzaioli, A. M.; Hall, R. M.; De Castro, C., Structure of the K2 capsule associated with the KL2 gene cluster of *Acinetobacter baumannii*. *Glycobiology* **2014**, *24* (6), 554-563.
219. Cacalano, G.; Kays, M.; Saiman, L.; Prince, A., Production of the *Pseudomonas aeruginosa* neuraminidase is increased under hyperosmolar conditions and is regulated by genes involved in alginate expression. *J. Clin. Invest.* **1992**, *89* (6), 1866-1874.
220. Horzempa, J.; Held, T. K.; Cross, A. S.; Furst, D.; Qutyan, M.; Neely, A. N.; Castric, P., Immunization with a *Pseudomonas aeruginosa* 1244 Pilin Provides O-Antigen-Specific Protection. *Clin. Vaccine Immunol.* **2008**, *15* (4), 590-597.
221. Vann, W. F.; Daines, D. A.; Murkin, A. S.; Tanner, M. E.; Chaffin, D. O.; Rubens, C. E.; Vionnet, J.; Silver, R. P., The NeuC Protein of *Escherichia coli* K1 Is a UDP-N-Acetylglucosamine 2-Epimerase. *J. Bacteriol.* **2004**, *186* (3), 706-712.

222. Vann, W. F.; Tavarez, J. J.; Crowley, J.; Vimr, E.; Silver, R. P., Purification and characterization of the Escherichia coli Kl neuB gene product N-acetylneuraminic acid synthetase. *Glycobiology* **1997**, *7* (5), 697-701.
223. Chen, H.; Blume, A.; Zimmermann-Kordmann, M.; Reutter, W.; Hinderlich, S., Purification and characterization of N-acetylneuraminic acid-9-phosphate synthase from rat liver. *Glycobiology* **2002**, *12* (2), 65-71.
224. Hinderlich, S.; Stäsche, R.; Zeitler, R.; Reutter, W., A Bifunctional Enzyme Catalyzes the First Two Steps in N-Acetylneuraminic Acid Biosynthesis of Rat Liver: PURIFICATION AND CHARACTERIZATION OF UDP-N-ACETYLGLUCOSAMINE 2-EPIMERASE/N-ACETYLMANNOSAMINE KINASE. *J. Biol. Chem.* **1997**, *272* (39), 24313-24318.
225. Creuzenet, C.; Schur, M. J.; Li, J.; Wakarchuk, W. W.; Lam, J. S., FlaA1, a new bifunctional UDP-GlcNAc C6 Dehydratase/ C4 reductase from Helicobacter pylori. *J. Biol. Chem.* **2000**, *275* (45), 34873-80.
226. Schoenhofen, I. C.; McNally, D. J.; Vinogradov, E.; Whitfield, D.; Young, N. M.; Dick, S.; Wakarchuk, W. W.; Brisson, J.-R.; Logan, S. M., Functional Characterization of Dehydratase/Aminotransferase Pairs from Helicobacter and Campylobacter: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.* **2006**, *281* (2), 723-732.
227. Guerry, P.; Ewing, C. P.; Schirm, M.; Lorenzo, M.; Kelly, J.; Pattarini, D.; Majam, G.; Thibault, P.; Logan, S., Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence. *Mol. Microbiol.* **2006**, *60* (2), 299-311.
228. Liu, F.; Tanner, M. E., PseG of Pseudaminic Acid Biosynthesis: a UDP-sugar hydrolase as a masked glycosyltransferase. *J. Biol. Chem.* **2006**, *281* (30), 20902-20909.
229. Tabei, S. M. B.; Hitchen, P. G.; Day-Williams, M. J.; Merino, S.; Vart, R.; Pang, P.-C.; Horsburgh, G. J.; Viches, S.; Wilhelms, M.; Tomas, J. M.; Dell, A.; Shaw, J. G., An Aeromonas caviae genomic island is required for both O-antigen lipopolysaccharide biosynthesis and flagellin glycosylation. *J. Bacteriol.* **2009**, *191* (8), 2851-2863.
230. Ishiyama, N.; Creuzenet, C.; Miller, W. L.; Demendi, M.; Anderson, E. M.; Harauz, G.; Lam, J. S.; Berghuis, A. M., Structural Studies of FlaA1 from Helicobacter pylori Reveal the Mechanism for Inverting 4,6-Dehydratase Activity. *J. Biol. Chem.* **2006**, *281* (34), 24489-24495.
231. Jansonius, J. N., Structure, evolution and action of vitamin B6-dependent enzymes. *Curr. Opin. Struct. Biol.* **1998**, *8* (6), 759-769.
232. Song, W. S.; Nam, M. S.; Namgung, B.; Yoon, S.-i., Structural analysis of PseH, the Campylobacter jejuni N-acetyltransferase involved in bacterial O-linked glycosylation. *Biochem. Biophys. Res. Commun.* **2015**, *458* (4), 843-848.
233. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B., An Evolving Hierarchical Family Classification for Glycosyltransferases. *J. Mol. Biol.* **2003**, *328* (2), 307-317.
234. Rangarajan, E. S.; Proteau, A.; Cui, Q.; Logan, S. M.; Potetinova, Z.; Whitfield, D.; Purisima, E. O.; Cygler, M.; Matte, A.; Sulea, T.; Schoenhofen, I. C., Structural and Functional Analysis of Campylobacter jejuni PseG: A UDP-sugar hydrolase from the pseudaminic acid biosynthetic pathway. *J. Biol. Chem.* **2009**, *284* (31), 20989-21000.
235. Lv, X.; Cao, H.; Lin, B.; Wang, W.; Zhang, W.; Duan, Q.; Tao, Y.; Liu, X.-W.; Li, X., Synthesis of Sialic Acids, Their Derivatives, and Analogs by Using a Whole-Cell Catalyst. *Chemistry – A European Journal* **2017**, *23* (60), 15143-15149.

236. Chou, W. K.; Dick, S.; Wakarchuk, W. W.; Tanner, M. E., Identification and characterization of NeuB3 from *Campylobacter jejuni* as a pseudaminic acid synthase. *J. Biol. Chem.* **2005**, *280* (43), 35922-35928.
237. Abnave, P.; Muracciole, X.; Ghigo, E., Coxiella burnetii Lipopolysaccharide: What Do We Know? *Int. J. Mol. Sci.* **2017**, *18* (12), 2509.
238. Gaugler, R. W.; Gabriel, O., Biological Mechanisms Involved in the Formation of Deoxy Sugars: VII BIOSYNTHESIS OF 6-DEOXY-L-TALOSE. *J. Biol. Chem.* **1973**, *248* (17), 6041-6049.
239. Creuzenet, C.; Belanger, M.; Wakarchuk, W. W.; Lam, J. S., Expression, Purification, and Biochemical Characterization of WbpP, a New UDP-GlcNAc C4 Epimerase from *Pseudomonas aeruginosa* Serotype O6. *J. Biol. Chem.* **2000**, *275* (25), 19060-19067.
240. Jörnvall, H.; Persson, B.; Krook, M.; Atrian, S.; Gonzalez-Duarte, R.; Jeffery, J.; Ghosh, D., Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **1995**, *34* (18), 6003-6013.
241. McNally, D. J.; Hui, J. P. M.; Aubry, A. J.; Mui, K. K. K.; Guerry, P.; Brisson, J.-R.; Logan, S. M.; Soo, E. C., Functional Characterization of the Flagellar Glycosylation Locus in *Campylobacter jejuni* 81-176 Using a Focused Metabolomics Approach. *J. Biol. Chem.* **2006**, *281* (27), 18489-18498.
242. Morrison, J. P.; Schoenhofen, I. C.; Tanner, M. E., Mechanistic studies on PseB of pseudaminic acid biosynthesis: A UDP-N-acetylglucosamine 5-inverting 4,6-dehydratase. *Bioorg. Chem.* **2008**, *36* (6), 312-320.
243. Reid, C. W.; Stupak, J.; Chen, M. M.; Imperiali, B.; Li, J.; Szymanski, C. M., Affinity-Capture Tandem Mass Spectrometric Characterization of Polyprenyl-Linked Oligosaccharides: Tool to Study Protein N-Glycosylation Pathways. *Anal. Chem.* **2008**, *80* (14), 5468-5475.
244. MEHTA, P. K.; HALE, T. I.; CHRISTEN, P., Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur. J. Biochem.* **1993**, *214* (2), 549-561.
245. Catazaro, J.; Caprez, A.; Guru, A.; Swanson, D.; Powers, R., Functional Evolution of PLP-dependent Enzymes based on Active-Site Structural Similarities. *Proteins* **2014**, *82* (10), 2597-2608.
246. Noland, B. W.; Newman, J. M.; Hendle, J.; Badger, J.; Christopher, J. A.; Tresser, J.; Buchanan, M. D.; Wright, T. A.; Rutter, M. E.; Sanderson, W. E.; Müller-Dieckmann, H.-J.; Gajiwala, K. S.; Buchanan, S. G., Structural Studies of *Salmonella typhimurium* ArnB (PmrH) Aminotransferase: A 4-Amino-4-Deoxy-L-Arabinose Lipopolysaccharide-Modifying Enzyme. *Structure* **2002**, *10* (11), 1569-1580.
247. Obhi, R. K.; Creuzenet, C., Biochemical Characterization of the *Campylobacter jejuni* Cj1294, a Novel UDP-4-keto-6-deoxy-GlcNAc Aminotransferase That Generates UDP-4-amino-4,6-dideoxy-GalNAc. *J. Biol. Chem.* **2005**, *280* (21), 20902-20908.
248. Liu, Y. C.; Ud-Din, A. I.; Roujeinikova, A., Cloning, purification and preliminary crystallographic analysis of the *Helicobacter pylori* pseudaminic acid biosynthesis N-acetyltransferase PseH. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2014**, *70* (9), 1276-1279.
249. Dyda, F.; Klein, D. C.; Hickman, A. B., GCN5-Related N-Acetyltransferases: A Structural Overview. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29* (1), 81-103.
250. Vetting, M. W.; S. de Carvalho, L. P.; Yu, M.; Hegde, S. S.; Magnet, S.; Roderick, S. L.; Blanchard, J. S., Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **2005**, *433* (1), 212-226.

251. Ud-Din, A. I.; Liu, Y. C.; Roujeinikova, A., Crystal structure of Helicobacter pylori pseudaminic acid biosynthesis N-acetyltransferase PseH: implications for substrate specificity and catalysis. *PLoS One* **2015**, *10* (3), e0115634/1-e0115634/14.
252. Simplicio, A.; Clancy, J.; Gilmer, J., Prodrugs for Amines. *Molecules* **2008**, *13* (3), 519.
253. Sagandira, C. R.; Watts, P., Synthesis of Amines, Carbamates and Amides by Multi-Step Continuous Flow Synthesis. *Eur. J. Org. Chem.* **2017**, *2017* (44), 6554-6565.
254. Majorek, K. A.; Kuhn, M. L.; Chruszcz, M.; Anderson, W. F.; Minor, W., Structural, functional and inhibition studies of a GNAT superfamily protein PA4794: a new C-terminal lysine protein acetyltransferase from Pseudomonas aeruginosa. *J. Biol. Chem.* **2013**.
255. Favrot, L.; Blanchard, J. S.; Vergnolle, O., Bacterial GCN5-Related N-Acetyltransferases: From Resistance to Regulation. *Biochemistry* **2016**, *55* (7), 989-1002.
256. Okamoto, S.; Suzuki, Y., Chloramphenicol-, Dihydrostreptomycin-, and Kanamycin-Inactivating Enzymes from Multiple Drug-Resistant Escherichia coli Carrying Episome 'R'. *Nature* **1965**, *208*, 1301.
257. Brownell, J. E.; Zhou, J.; Ranalli, T.; Kobayashi, R.; Edmondson, D. G.; Roth, S. Y.; Allis, C. D., Tetrahymena Histone Acetyltransferase A: A Homolog to Yeast Gcn5p Linking Histone Acetylation to Gene Activation. *Cell* **1996**, *84* (6), 843-851.
258. Zheng, W.; Scheibner, K. A.; Ho, A. K.; Cole, P. A., Mechanistic studies on the alkyltransferase activity of serotonin N-acetyltransferase. *Chem. Biol.* **2001**, *8* (4), 379-389.
259. Magnet, S.; Blanchard, J. S., Molecular Insights into Aminoglycoside Action and Resistance. *Chem. Rev.* **2005**, *105* (2), 477-498.
260. Starai, V. J.; Escalante-Semerena, J. C., Identification of the Protein Acetyltransferase (Pat) Enzyme that Acetylates Acetyl-CoA Synthetase in Salmonella enterica. *J. Mol. Biol.* **2004**, *340* (5), 1005-1012.
261. Mizzen, C. A.; Allis, C. D., Linking histone acetylation to transcriptional regulation. *Cellular and Molecular Life Sciences CMLS* **1998**, *54* (1), 6-20.
262. Salah Ud-Din, A. I. M.; Tikhomirova, A.; Roujeinikova, A., Structure and Functional Diversity of GCN5-Related N-Acetyltransferases (GNAT). *Int. J. Mol. Sci.* **2016**, *17* (7), 1018.
263. Pietrocola, F.; Galluzzi, L.; Bravo-San Pedro, José M.; Madeo, F.; Kroemer, G., Acetyl Coenzyme A: A Central Metabolite and Second Messenger. *Cell Metab.* **2015**, *21* (6), 805-821.
264. Mishra, P. K.; Drueckhammer, D. G., Coenzyme A Analogues and Derivatives: Synthesis and Applications as Mechanistic Probes of Coenzyme A Ester-Utilizing Enzymes. *Chem. Rev.* **2000**, *100* (9), 3283-3310.
265. Taylor, J. E.; Bull, S. D., 6.11 N-Acylation Reactions of Amines. In *Comprehensive Organic Synthesis II (Second Edition)*, Knochel, P., Ed. Elsevier: Amsterdam, 2014; pp 427-478.
266. Bülter, T.; Elling, L., Enzymatic synthesis of nucleotide sugars. *Glycoconjugate J.* **1999**, *16* (2), 147-159.
267. Krusemark, C. J.; Frey, B. L.; Smith, L. M.; Belshaw, P. J., Complete Chemical Modification of Amine and Acid Functional Groups of Peptides and Small Proteins. In *Gel-Free Proteomics: Methods and Protocols*, Gevaert, K.; Vandekerckhove, J., Eds. Humana Press: Totowa, NJ, 2011; pp 77-91.

268. Aerry, S.; Kumar, A.; Saxena, A.; De, A.; Mozumdar, S., Chemoselective acetylation of amines and thiols using monodispersed Ni-nanoparticles. *Green Chemistry Letters and Reviews* **2013**, *6* (2), 183-188.
269. BASU, K.; CHAKRABORTY, S.; SARKAR, A. K.; SAHA, C., Efficient acetylation of primary amines and amino acids in environmentally benign brine solution using acetyl chloride. *Journal of Chemical Sciences* **2013**, *125* (3), 607-613.
270. Sanz Sharley, D. D.; Williams, J. M. J., Acetic acid as a catalyst for the N-acylation of amines using esters as the acyl source. *Chem. Commun.* **2017**, *53* (12), 2020-2023.
271. Phukan, K.; Ganguly, M.; Devi, N., Mild and Useful Method for N-Acylation of Amines. *Synth. Commun.* **2009**, *39* (15), 2694-2701.
272. Roseman, S.; Ludowieg, J., N-Acetylation of the Hexosamines. *J. Am. Chem. Soc.* **1954**, *76* (1), 301-302.
273. Patel, M. S.; Korotchkina, L. G., The biochemistry of the pyruvate dehydrogenase complex\*. *Biochem. Mol. Biol. Educ.* **2003**, *31* (1), 5-15.
274. Chenault, H. K.; Simon, E. S.; Whitesides, G. M., Cofactor Regeneration for Enzyme-Catalysed Synthesis. *Biotechnol. Genet. Eng. Rev.* **1988**, *6* (1), 221-270.
275. Zhao, H.; van der Donk, W. A., Regeneration of cofactors for use in biocatalysis. *Curr. Opin. Biotechnol.* **2003**, *14* (6), 583-589.
276. Luong, A.; Hannah, V. C.; Brown, M. S.; Goldstein, J. L., Molecular Characterization of Human Acetyl-CoA Synthetase, an Enzyme Regulated by Sterol Regulatory Element-binding Proteins. *J. Biol. Chem.* **2000**, *275* (34), 26458-26466.
277. Hunter, G. A.; Ferreira, G. C., A Continuous Spectrophotometric Assay for 5-Aminolevulinate Synthase That Utilizes Substrate Cycling. *Anal. Biochem.* **1995**, *226* (2), 221-224.
278. Patel, S. S.; Conlon, H. D.; Walt, D. R., Enzymic synthesis of L-acetylcarnitine and citric acid using acetyl coenzyme A recycling. *The Journal of Organic Chemistry* **1986**, *51* (14), 2842-2844.
279. Simon, E. J.; Shemin, D., The Preparation of S-Succinyl Coenzyme A. *J. Am. Chem. Soc.* **1953**, *75* (10), 2520-2520.
280. Lowe, D. M.; Tubbs, P. K., Preparation of bromo[1-14C]acetyl-coenzyme A as an affinity label for acetyl-coenzyme A binding sites. *Anal. Biochem.* **1983**, *132* (2), 276-284.
281. Ouyang, T.; Walt, D. R., A new chemical method for synthesizing and recycling acyl coenzyme A thioesters. *The Journal of Organic Chemistry* **1991**, *56* (11), 3752-3755.
282. Modis, Y.; Wierenga, R., Two crystal structures of N-acetyltransferases reveal a new fold for CoA-dependent enzymes. *Structure* **1998**, *6* (11), 1345-1350.
283. Mizanur, R. M.; Jaipuri, F. A.; Pohl, N. L., One-Step Synthesis of Labeled Sugar Nucleotides for Protein O-GlcNAc Modification Studies by Chemical Function Analysis of an Archaeal Protein. *J. Am. Chem. Soc.* **2005**, *127* (3), 836-837.
284. Moynihan, P. J.; Clarke, A. J., O-Acetylation of Peptidoglycan in Gram-negative Bacteria: IDENTIFICATION AND CHARACTERIZATION OF PEPTIDOGLYCAN O-ACETYLTRANSFERASE IN NEISSERIA GONORRHOEAE. *J. Biol. Chem.* **2010**, *285* (17), 13264-13273.
285. Moynihan, P. J.; Clarke, A. J., Assay for peptidoglycan O-acetyltransferase: A potential new antibacterial target. *Anal. Biochem.* **2013**, *439* (2), 73-79.



286. Moynihan, P. J.; Clarke, A. J., Substrate Specificity and Kinetic Characterization of Peptidoglycan O-Acetyltransferase B from *Neisseria gonorrhoeae*. *J. Biol. Chem.* **2014**, *289* (24), 16748-16760.
287. Bae, B.; Cobb, R. E.; DeSieno, M. A.; Zhao, H.; Nair, S. K., New N-Acetyltransferase Fold in the Structure and Mechanism of the Phosphonate Biosynthetic Enzyme FrbF. *J. Biol. Chem.* **2011**, *286* (41), 36132-36141.
288. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R., Engineering Chemical Reactivity on Cell Surfaces Through Oligosaccharide Biosynthesis. *Science* **1997**, *276* (5315), 1125-1128.
289. Saxon, E.; Bertozzi, C. R., Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000**, *287* (5460), 2007-2010.
290. Winans, K. A.; Bertozzi, C. R., Inner space exploration: the chemical biologist's guide to the cell. *Chem. Biol.* **1998**, *5* (12), R313-R315.
291. Laughlin, S. T.; Bertozzi, C. R., Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat. Protoc.* **2007**, *2*, 2930.
292. Nwe, K.; Brechbiel, M. W., Growing Applications of "Click Chemistry" for Bioconjugation in Contemporary Biomedical Research. *Cancer Biother. Radiopharm.* **2009**, *24* (3), 289-302.
293. Yang, C.; Mi, J.; Feng, Y.; Ngo, L.; Gao, T.; Yan, L.; Zheng, Y. G., Labeling Lysine Acetyltransferase Substrates with Engineered Enzymes and Functionalized Cofactor Surrogates. *J. Am. Chem. Soc.* **2013**, *135* (21), 7791-7794.
294. Wang, Y.; Lazor, K. M.; DeMeester, K. E.; Liang, H.; Heiss, T. K.; Grimes, C. L., Postsynthetic Modification of Bacterial Peptidoglycan Using Bioorthogonal N-Acetylcysteamine Analogs and Peptidoglycan O-Acetyltransferase B. *J. Am. Chem. Soc.* **2017**, *139* (39), 13596-13599.
295. Liu, F.; Aubry, A. J.; Schoenhofen, I. C.; Logan, S. M.; Tanner, M. E., The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella. *ChemBioChem* **2009**, *10* (8), 1317-1320.
296. Liav, A.; Sharon, N., Synthesis of 2,4-diacetamido-2,4,6-trideoxy-L-altrose,-L-idose, and -L-talose from benzyl 6-deoxy- 3,4-O-isopropylidene- $\beta$ -L-galactopyranoside. *Carbohydr. Res.* **1973**, *30* (1), 109-126.
297. Schwab, J. M.; Klassen, J. B., Steric course of the allylic rearrangement catalyzed by .beta.-hydroxydecanoylthioester dehydrase. Mechanistic implications. *J. Am. Chem. Soc.* **1984**, *106* (23), 7217-7227.
298. Wang, H.; Vath, G. M.; Gleason, K. J.; Hanna, P. E.; Wagner, C. R., Probing the Mechanism of Hamster Arylamine N-Acetyltransferase 2 Acetylation by Active Site Modification, Site-Directed Mutagenesis, and Pre-Steady State and Steady State Kinetic Studies. *Biochemistry* **2004**, *43* (25), 8234-8246.
299. Means, G. E.; Bender, M. L., Acetylation of human serum albumin by p-nitrophenyl acetate. *Biochemistry* **1975**, *14* (22), 4989-4994.
300. Westwood, I. M.; Sim, E., Kinetic characterisation of arylamine N-acetyltransferase from *Pseudomonas aeruginosa*. *BMC Biochem.* **2007**, *8*, 3-3.
301. Minchin, R. F.; Butcher, N. J., The role of lysine100 in the binding of acetylcoenzyme A to human arylamine N-acetyltransferase 1: Implications for other acetyltransferases. *Biochem. Pharmacol.* **2015**, *94* (3), 195-202.

302. Yang, W.; Drueckhammer, D. G., Understanding the Relative Acyl-Transfer Reactivity of Oxoesters and Thioesters: Computational Analysis of Transition State Delocalization Effects. *J. Am. Chem. Soc.* **2001**, *123* (44), 11004-11009.