
Article:
Chidwick, Harriet and Fascione, Martin Anthony orcid.org/0000-0002-0066-4419 (2019) Mechanistic and structural studies into the biosynthesis of the bacterial sugar pseudaminic acid (Pse5Ac7Ac). Organic and Biomolecular Chemistry. pp. 799-809. ISSN 1477-0520

https://doi.org/10.1039/C9OB02433F

Organic & Biomolecular Chemistry

REVIEW

Mechanistic and structural studies into the biosynthesis of the bacterial sugar pseudaminic acid (Pse5Ac7Ac)

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Harriet S. Chidwick, Martin A. Fascione*

The non-mammalian nonulosonic acid sugar pseudaminic acid (Pse) is present on the surface of a number of human pathogens including *Campylobacter jejuni* and *Helicobacter pylori* and other bacteria such as multidrug resistant *Acinetobacter baumannii*. It is likely important for evasion of the host immune sysyem, and also plays a role in bacterial motility through flagellin glycosylation. Herein we review the mechanistic and structural characterisation of the enzymes responsible for the biosynthesis of the Pse parent structure, Pse5Ac7Ac in bacteria.

Chemical abbrieviations; *Pse* pseudaminic acid, *Neu5Ac* neuraminic acid, *Pse5Ac7Ac* 5,7-diamino-3,5,7,9-tetradeoxy-L-glycero-L-manno-non-2-ulopyranosonic acid *UDP-* uridine diphosphate-, *CMP-*cytidine monophosphate, *PEP* phosphoenol pyruvate, *PLP* pyridoxal phosphate, *PMP* pyridoxamine phosphate, *Ac-CoA* acetyl coenzyme A, *NAD(P)H* nicotine adenine dinucleotide (phosphate), L-*Glu* L-glutamate, *KDO* 3-deoxy-D-manno-octulosonic acid

Introduction

Nonulosonic acids (NulOs) are nine-carbon α-keto-acid sugars that are ubiquitous in Nature, occuring in several cell surface glycoconjugates where they play a crucial role in cell-cell interactions. Pseudaminic acids (Pse) are rare non-mammalian nonulosonic acids that belong to this class of sugars, and are epimeric at the C5, C7 and C8 positions when compared to the stereochemistry of the ubiquitous human sialic acid; neuraminic acid (Neu5Ac) **1** (**Figure 1**). The general term pseudaminic acid is commonly used to describe the parent Pse5Ac7Ac structure **2**, which differs in functionality from Neu5Ac by replacement of the C7 hydroxyl with an acetamido group and loss of a C9 hydroxyl group (**Figure 1**). Pse5Ac7Ac **2** was first discovered as a component of the *Pseudonomas aeruginosa* and *Shigella Boydii* lipopolysaccharides (LPS), with NMR used to assign the configuration of this previously undisclosed sugar as L-glycero-L-manno.[1] Pse5Ac7Ac **2** has since been identified as biosynthesised by a number of gram negative bacteria including pathogens *Aeromonas caviae*,[2] *Helicobacter pylori*[3] and *Campylobacter jejuni*[4] and more recently it has also been discovered in the gram positive bacteria *Bacillus thuringiensis*.[5] During the initial identification of Pse5Ac7Ac **2**, a derivative was also tentatively assigned with NMR peaks observed consistent with an *N*-(3-hydroxybutyryl) group at the C5 position (Pse5Hb7Ac).[6] Subsequent research on other Pse structures has highlighted that derivatisation of the C5 and C7 acetamido groups is common (**Figure 2**).[7] For example, *C. jejuni* flagellin protein is extensively glycosylated with Pse in the form of Pse5Ac7Ac **2**, Pse5Am7Ac (C5 acetaminido) and minor amounts of other derivatives.[3]

Pse5Ac7Ac **2** and derivatives are most commonly found on bacterial cell surfaces with reports of LPS and capsular polysaccharide (CPS) containing Pse, in addition to Pse glycosylated flagella and pili. Unusually for a nulO, there is evidence for Pse5Ac7Ac **2** linked in both the α- and β-anomeric configuration. For example *P. aeruginosa* O13 LPS incorporates a Pse5Ac7Ac(α2-3)L-FucAm bond **3**[8] whereas *Escherichia coli* O136 LPS has a Pse5Ac7Ac(β2-4)D-Gal **4** (**Figure 3**).[9] Notably, Pse5Ac7Ac **2** is more commonly found linked within the glycan in LPS and CPS structures rather than at the terminal position which is more typical of mamallian human sialic acids, and can act as a glycosyl acceptor via its derivatives at the C5, or C7 positions. Indeed the CPS of symbiont *Sinorhizobium fredii* is a Pse5Ac7Hb (C7 *N*-(3-hydroxybutyryl)) homo-polysaccharide **5** with the α glycosidic linkage between the anomeric position and the hydroxyl group on the C5 *N*-hydroxybutyryl (**Figure 3**).[10] Interestingly *P. aeruginosa* PA1244 modifies its pili with an α-Pse5Am7Hb (C5 acetaminido, C7 *N*-(3-hydroxybutyryl)) terminating trisaccharide[11] which also forms the repeating unit in the LPS of *P. aeruginosa* belonging to the O7 serotype.[12] This consistent positioning of Pse on bacterial cell surface structures such as flagella, pili, CPS and LPS, has been shown to provide a selective advantage to the bacteria and can have a direct role as a virulence factor.[7] For example, *C. jejuni* flagellin proteins (such as FlaA1) were shown to be glycosylated with Pse5Ac7Ac **2** monosaccharides (or a derivative) at 19 sites.[13] Subsequent mutations to C. *jejuni* Pse5Ac7Ac **2** biosynthesis genes resulted in incorrect assembly of flagellar and therefore reduced

[a] Department of Chemistry, University of York, York, YO10 5DD, UK. Email: martin.fascione@york.ac.uk

† Footnotes relating to the title and/or authors should appear here.

motility of the bacteria.[14] Similarly *H. pylori* FlaA and FlaB flagellin proteins have been found to be exclusively glycosylated with Pse5Ac7Ac **2** and prevention of flagellin glycosylation resulted in no detectable flagella and non-motile bacterium.[3] As motility has previously been shown to be a key factor in the ability of these two bacteria to establish an infection,[15, 16] it is clear that therapeutics targeting the Pse5Ac7Ac **2** biosynthesis could reduce the virulence of such bacteria. Furthermore it has been proposed that Pse5Ac7Ac **2** may play a role in bacterial evasion of the host immune system due to its structural similarities to the human sugar Neu5Ac **1**. Indeed, many pathogens manipulate host immune system by binding to receptors promoting interleukin-10 (IL-10) induction which can supress immune responses.[17] Importantly *C. jejuni* FlaA1 was found to be important for binding to siglec-10 (a host immune system glycan receptor) with FlaA1 lacking Pse5Ac7Ac glycosylation resulting in a significant decrease in IL-10 induction, suggesting that Pse5Ac7Ac **2** can be recognised by siglec-10 and dampen the immune response.[18]

A detailed understanding of the biosynthesis of this bacterial sugar is therefore essential as inhibition of this pathway could form the basis of a novel antimicrobial strategy. Herein we provide an overview of the biosynthesis of Pse5Ac7Ac **2**, specifically highlighting mechanistic and structural studies of the five essential enzymes PseB, PseC, PseH, PseG and PseI conserved in *C. jejuni* and *H. pylori*.

Pseudaminic acid biosynthesis

Neu5Ac biosynthesis

Commonly nulOs are biosynthesised from the activated sugar uridine diphosphate-*N*-acetyl-α-D-glucosamine (UDP-GlcNAc) **7**,[19] before conversion to their activated glycosyl donor analogues by enzyme catalysed transfer of a cytidine monophosphate (CMP) group. In bacteria, biosynthesis of Neu5Ac **1** from UDP-GlcNAc **7** requires only two enzymes; a UDP-GlcNAc 2-epimerase to hydrolyse the UDP- group and epimerise the stereochemistry at C2,[20] and a Neu5Ac synthase that catalyses the condensation reaction of ManNAc **8** and phospho-enol-pyruvate (PEP) to afford the nonulosonic acid backbone (**Scheme 1**).[21] In mammals, the UDP-GlcNAc 2-epimerase is bifunctional and after conversion to ManNAc **8** it subsequently catalyses product phosphorylation to yeild ManNAc6P **9**.[22] A Neu5Ac-9-P synthase is then employed to catalyse the condensation reaction with PEP to afford the nulO **10** which can then de-phosphorylated yielding the desired Neu5Ac **1** (**Scheme 1**). It was hypothesised that the Pse5Ac7Ac **2** biosynthetic pathway may resemble the biosynthesis of Neu5Ac **1**; with enzymes catalysing the conversion of the UDP-GlcNAc starting material **7** into a precursor that could act as a substrate for a Pse5Ac7Ac synthase catalysed reaction to afford Pse5Ac7Ac **2**.

Identification of the pseudaminic acid biosynthetic enzymes

H. pylori and *C. jejuni* flagella are predominantly glycosylated with Pse5Ac7Ac **2**, and derivatives, therefore the flagellin glycosylation gene clusters in these bacteria were initially inspected for potential Pse biosynthesis genes. Schirm *et al.* initially identified four genes (*HP0326A*, *HP0326B*, *HP0840* and *HP0178*) in *H. pylori* that were found to be essential for the production of Pse5Ac7Ac **2**.[3] HP0840 was identified as the first enzyme in the Pse5Ac7Ac **2** biosynthetic pathway and upon inactivation of this gene there was an accumulation of UDP-GlcNAc **7** providing evidence for this as the Pse5Ac7Ac **2** precursor.[23] Sequence homology studies enabled assignment of the HP0178 enzyme as a Neu5Ac synthase and identified the HP0326A protein to have the highest homology with CMP-Neu5Ac synthetases.[3] Thus it was proposed that the final stages of Pse5Ac7Ac biosynthesis did indeed mimic that of Neu5Ac **1** and that other biosynthetic enzymes would be required to process UDP-GlcNAc **7** into the Pse5Ac7Ac synthase substrate **11**. Putative UDP-GlcNAc dehydratase and aminotransferase enzymes were purified and their reaction products analysed by NMR. *H. pylori* (HP0840/HP0366) and *C. jejuni* (Cj1293/Cj1294) products **12** and **13** were assigned as the first two intermediates in the Pse5Ac7Ac **2** biosynthetic pathway based on stereochemical assignments (**Scheme 2**).[24] *HP0326B* and *HP0327* were identified as potential Pse5Ac7Ac **2** biosynthetic genes by comparing the genomes of other bacteria also known to display surface Pse5Ac7Ac **2**. Sequence alignment of HP0327 suggested the enzyme was the acetyl-transferase required to afford the desired diacetamido product **14**.[25] HP0326B displayed sequence homology with a glycosyltransferase enzyme suggesting it may be required for the transfer of Pse onto other sugars or proteins rather than involved in Pse5Ac7Ac **2** biosynthesis. However upon mutation of this gene there was an accumulation of Pse5Ac7Ac **2** biosynthetic intermediates (**12**, **13** and **14**) rather than the expected CMP-Pse5Ac7Ac that would accumulate if the gene encoded for a glycosyltransferase suggesting that it may actually be the required UDP-hydrolase in the Pse5Ac7Ac **2** biosynthetic pathway.[3] Thorough structural and biochemical analyses of each enzyme has led to confirmation of all of the biosynthetic intermediates in the pathway which are analogous in *H. pylori* and *C. jejuni*. The five biosynthetic enzymes required for the production of Pse5Ac7Ac **2** are now known as PseB, PseC, PseH, PseG, and PseI (**Scheme 2**).[26]

A slightly different Pse5Ac7Ac **2** biosynthetic pathway has been assigned in *A. caviae*; another pathogen that displays Pse5Ac7Ac **2** on its flagellin as well as in the LPS. Mutations to the flm gene locus caused a reduction in motility, through loss of flagella, and also a loss of LPS, suggesting this locus has a role in both flagellar assembly and LPS biosynthesis.[27] A cluster of genes in this locus displayed homology to the biosynthetic genes found in *C. jejuni* and *H. pylori* and comparison highlighted conserved domains for FlmA with PseB, FlmB with PseC and NeuB with PseI. However in *A. cavaie* 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.[27] The *A. cavaie* 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 has yet to be carried out. Further deviations from the original Pse5Ac7Ac **2** 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 Pse5Ac7Ac **2** biosynthetic pathway.[5] Sequence alignment allocated homologues for the PseC, PseH, PseG

and PseI 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 **7** into UDP-6-deoxy-D-GlcNAc-5,6-ene **15** and a second enzyme (Pal) acts as a C4 oxidase and C5,6 reductase resulting in epimerisation at C5 compared to the starting material **7**.[5]

PseB; a UDP-GlcNAc 5-inverting-4,6-dehydratase

PseB is the first enzyme in the pathway and converts UDP-GlcNAc **7** to UDP-4-keto-6-deoxy-β-L-IdoNAc **12**. 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.[28] Characterisation of this enzyme was initially ambiguous, with confusion over its catalytic activity, co-factor requirement and products released. Although PseB (*FlaA1/HP0840* in *H. pylori Cj1293* in *C. jejuni*) displays conserved domains with 4,6-dehydratases it also displays remarkable sequence similarities with UDP-GlcNAc **7** C4 epimerases,[23] such as WbpP from *P. aeruginosa*.[29] However, 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[30] and part of a sub-group that exhibits 4,6-dehydratase activity on nucleotide activated sugars to form deoxy-hexoses.[31] There was also initial uncertainty regarding the observed 4-keto product with reports of formation of both of the C4 epimers; UDP-4-keto-6-deoxy-β-L-IdoNAc **12**[28] and UDP-4-keto-6-deoxy-α-D-GlcNAc.[23] NMR data collected *in situ* confirmed the *C. jeuni* PseB initial product as UDP-4-keto-6-deoxy-β-L-IdoNAc **12**,[28] observed as the hydrated form in aqueous solution **16**.[32] Upon *in vitro* incubation of UDP-GlcNAc **7** with the five Pse5Ac7Ac **2** biosynthetic enzymes and the CMP-Pse5Ac7Ac synthetase enzyme, there was a drastic reduction in PseB activity as CMP-Pse5Ac7Ac accumulated.[26] Therefore PseB inhibition with CMP-Pse5Ac7Ac was investigated and found to be potent at very low concentrations (100 μM) suggesting that the biosynthetic pathway is controlled by negative feedback.[33]

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 nicotine adenine dinucleotide (phosphate) (NAD(P)H) co-factor binding site.[23] UDP-GlcNAc **7** was initially proposed to be the substrate in activity assays and was later confirmed by observation of binding in crystal structure substrate complexes. "Apo" crystal structures of PseB revealed electron density concurrent with a bound NADPH molecule even without exogenous addition, suggesting this molecule tightly binds to PseB and is the likely co-factor (**Figure 4**).[28] The deeply buried co-factor is surrounded by residues orientated to have favourable interactions with the molecule including multiple H-bonding residues (**Figure 4**).[28] 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. The combination of biochemical analysis and structural studies of the PseB enzyme facilitated the proposal of an enzyme mechanism and the predcited sugar conformations as the reaction progresses in *H. pylori* (**Scheme 3**).[28] Firstly employing the (S/T)YK triad and $NADP^+$ to oxidise the C4 hydroxyl producing the ketone intermediate **17** 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 **18**. 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 **12**.[28]

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

A number of biomacromolecules employ amino sugars within their structures and biosynthesis often requires an aminotransferase catalysed transamination. Pyridoxal phosphate (PLP)-dependent aminotransferases are classified into four subgroups *via* comparison of amino acid sequences aligned based on the predicted secondary structure.[34] 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-β-L-IdoNAc **12**.[35] *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 5**). Crystal structures of *H. pylori* PseC additionally revealed the characteristic Type 1 aminotransferase PLP-binding site adjacent to the active site.[36] In particular the highly conserved aspartic acid and phenylalanine residue were identified as Asp154 and Phe84 respectively in *H. pylori* PseC (**Figure 6.a**). 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 π-stack with the co-factor pyridine ring, stabilising binding of the co-factor (**Figure 6.a**). Crystal structures of *H. pylori* PseC in complex with PLP showed it to form an internal aldimine with the Lys183 residue providing evidence for one of the intermediates in the proposed mechanism of the first half transamination reaction (**Figure 6.b**).[35] The identification of the natural amino donor of PLP-dependent aminotransferases has often 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.[37] In *C. jejuni* PseC a screen of all twenty amino acids as the amino donor revealed that maximum conversion was achieved using 10 mM L-Glu (20 molar equivalents of the substrate).[38]

Co-crystallisation of *H. pylori* PseC, PLP and the proposed product UDP-4-amino-4,6-dideoxy-β-L-AltNAc **13** also provided insight into the second half-transamination reaction as complexes formed showing the enzyme acting in reverse. Electron density in the active site could be attributed to a pyridoxamine phosphate (PMP)-sugar aldimine **19** suggesting that direct aminotransfer from the PMP **20** to the keto-sugar **12** occurs and confirming the proximity of the co-factor binding site and active sites (**Figure 7.a**). After consideration of the orientation of the active site residues, Lys183 was also suggested as the catalytic residue for the second half-reaction (**Figure 7.b**).[35] The importance of this residue was confirmed upon observation of drastically reduced

aminotransferase activity upon introduction of a Lys183Arg mutation.[35] These active site structural features led to the proposal of a mechanism mimicking that of ArnB (another UDP-4-keto aminotransferase)[37] whereby the PMP amine attacks the substrate C4 and Lys183 is utilised in a transaldimination reaction, triggering product release (**Scheme 4**). During the first half transamination, initially an enzyme-PLP Schiff base forms then a free amino donor releases the enzyme from the internal aldimine forming an external aldimine *via* a transaldimination reaction. Finally hydrolysis occurs resulting in release of a glyoxylic acid to produce free PMP **20** in the active site.[39] The second amino transfer occurs *via* attack of the PMP amine to the sugar C4 keto producing water and a ketimine intermediate **21** (**Scheme 4**). 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 **22**. Re-protonation at the sugar C4 position allows for formation of the sugar-co-factor aldimine **19** which is then released in a transaldimination reaction to give the bound co-factor enzyme complex **23** and the free aminated sugar **13** (**Scheme 4**).[35]

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% 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. For example the *H. pylori* catalytic residue Lys183 is aligned with the predicted *C. jejuni* catalytic residue; Lys181 surrounded by homologous residues. 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-β-L-AltNAc **13** and suggested its role as the second enzyme in the Pse5Ac7Ac **2** biosynthetic pathway.[24] Coupled reactions also revealed that aminotransferase activity occurred without the addition of exogenous PLP, suggesting that similarly to PseB, there is tight binding of the co-factor throughout purification.[38, 40] However in some cases it was found that introduction of PLP into the reaction mixture could increase activity suggesting that the intracellular level of PLP was not sufficient for saturation of PseC with its co-factor.[38]

PseH; a UDP-4-amino-4,6-dideoxy -β-L-IdoNAc acetyl transferase

The third step in the Pse5Ac7Ac **2** biosynthetic pathway is the transfer of an acetyl group to C4 of UDP-4-amino-4,6-dideoxy-β-L-AltNAc **13**.[41] This step is catalysed by PseH, an aminoglycoside *N*-acetyltransferase from the GCN5-related *N*-acetyltransferase (GNAT) superfamily, and as such, utilises acetyl coenzyme A (Ac-CoA) **24** as a co-factor.[42] The GNAT superfamily consists of more than 100,000 members found in all kingdoms of life,[43] and although the GNAT family has low sequence homology there is conservation of the core fold in all structures to date.[44] Structural characterisation of over twenty GNAT enzymes has revealed conserved secondary elements consisting largely of antiparallel β sheets (six or seven strands) connected by loops or one of the four conserved α helices.[42] This overall structure is evident in the highly similar *C. jejuni*[45] and *H. pylori*[46] PseH crystal structures (RMSD 1.04 Å for 134 C_α atoms) with the major difference being the extra α helix observed in *H. pylori* PseH (**Figure 8**). Both of the PseH crystal structures were identified as displaying structural similarities to WecD,[45, 46] another acetyl-transferase with a nucleotide-linked sugar substrate.[47] However the GNAT family member with the highest structural similarity to *C. jejun*i PseH, *Vibrio cholerae* spermidine acetyltransferase, utilises a structurally unrelated substrate.[45] *H. pylori* PseH was found to display highest structural similarity with a *E. coli* microcin C7 acetyltransferase,[46] a nucleotide containing molecule.[48] Another notable structural feature present in both structures is the "β-4 bulge" forcing the β-4 and β-5 strands apart at one end to form a V-shape cleft, commonly forming the Ac-CoA **24** binding site.[44] *C. jejuni* PseH Ac-CoA **24** is bound shaped like the letter 'L' and interactions with the enzyme occur with residues in the proximal α3, α4, β4 and β5 secondary structures (**Figure 9.a**).[45] The co-factor in *H. pylori* PseH is held slightly differently within the enzyme but with the thio-ester still buried deep within the cleft between the β-4 and β-5 sheets (**Figure 9.b**).[45]

Based on kinetic and structural data it has been proposed that this family of acetyltransferase enzymes catalyse the direct acetyl transfer from Ac-CoA **24** to substrates,[49-54] however there is an exception with evidence suggesting a yeast histone acetyltransferase progresses *via* an acetylated enzyme intermediate.[55] The crystal structures of the two PseH enzymes suggest that they follow the common direct acetyl transfer mechanism, requiring a basic residue proximal to the sugar **13** amino group and an acidic residue close to the Ac-CoA **24** sulfur. A suitable catalytic base proximal to the proposed substrate C4 amino group could not be easily identified so the substrate **13** was docked into the *H. pylori* PseH crystal structure.[45] A well-ordered water molecule was identified with hydrogen bonds to acidic residues and it was suggested that this molecule could mediate the deprotonation of substrate **13** to promote nucleophilic attack on the co-factor carbonyl **24** forming the tetrahedral intermediate **25**.[46] A conserved tyrosine was identified in both PseH enzymes (Tyr128 in *C. jejuni* PseH and Tyr138 in *H. pylori* PseH) as the potenital acid required to stablise the thiolate ion **26** generated during the collapse of the tetrahedral intermediate **25** (**Scheme 5**).[46]

PseG; a UDP-4-acetamido-4,6-dideoxy-β-L-AltNAc hydrolase

The *PseG* gene loci *Cj1312* and *HP0326B* were originally assigned to code for glycosyltransferases as they displayed significant alignment with conserved sequences of some UDP-sugar transferases.[3] PSI-BLAST homology sequencing of *Cj1312* classified it as belonging to the metal-independent GT-B superfamily,[56] however the amino acid sequence displays only modest overall identity with this class of protein. Therefore it was postulated that these enzymes could be the glycosyl hydrolase required to catalyse the "transfer" of the sugar substrate **14** onto a water molecule, resulting in the desired UDP- hydrolysis to afford the Pse5Ac7Ac synthase substrate **11**.[57] *In vitro* biochemical analysis of the Cj1312 enzyme confirmed it to catalyse hydrolysis of UDP- from UDP-4-acetamido-4,6-dideoxy-β-L-AltNAc **14** when no activity was observed with either the sterochemically similar UDP-4-amino-4,6-dideoxy-β-L-AltNAc **13** or

the functionally similar UDP-2,4-diacetamido-bacillosamine.[58] During *in vitro* characterisation of the *H. pylori* PseG enzyme, it was also found to efficiently catalyse UDP-hydrolysis of UDP-4-acetamido-4,6-dideoxy-β-L-AltNAc **14**.[26] Calculation of the kinetic constants using a coupled assay for monitoring UDP- release confirmed that this was the biologicallly relevant role of PseG as there is a relatively large specificity constant ($\boldsymbol{k_{cat}/K_m = 1.5\times 10^5\ M^{-1}\ s^{-1}}$).[57] Mutations to *Cj1312* resulted in non-motile *C. jejuni* phenotypes lacking in flagella filaments and hook structures suggesting this enzyme to be essential for the correct assembly of flagella.[25] PseG candidates have been identified in a number of Pse biosynthesis gene clusters in other bacteria and have been found to only have weak sequence similarity, however a minimum consensus sequence of DX_5GXGHX_2R was identified across eight putative PseG sequences.[58]

Glycosyl hydrolases have been shown to catalyse UDP-hydrolysis *via* attack of a water molecule onto the anomeric carbon[59, 60] or onto one of the phosphorus atoms in the UDP-group.[61, 62] The second mechanism was discounted for PseG following an experiment utilising $H_2{}^{18}O$, whereby mass spectrometry only identifed incorporation of the solvent ^{18}O into the hydrolysed sugar and not the released nucleotide moiety.[57] Further mechanistic detail was gathered by tracking the stereochemistry of the anomeric centre with ^1H NMR during PseG catalysed reactions in the presence of D_2O. It was demonstrated that hydrolysis of UDP-4-acetamido-4,6-dideoxy-β-L-AltNAc **14** occurs with an inversion of stereochemistry at C1 to afford 4-acetamido-4,6-dideoxy-α-L-AltNAc **11** which then undergoes non-enzymatic mutorotation in solution to exist predominantly as the β-anomer (**Scheme 6**).[57] This experiment also provided evidence against the mechanism proceeding through a glycal intermediate as there was no incorporation of solvent deuterium into product **11** during catalysis. Therefore the mechanism was predicted to involve a single displacement step *via* direct attack of a water molecule to the anomeric carbon (**Scheme 6**).

The overall crystal structure of *C. jejuni* PseG (PDB 3HBN)[58] showed high homology to that of an *E. coli* UDP-GlcNAc glycosyltransferase MurG (PDB 1NLM),[63] especially when ligand bound (**Figure 10**). MurG His19 has been suggested as the catalytic base[64] as incorporation of a H19A mutation drastically reduces the enzyme $k_{cat.}$[63] The same reduction in turnover is observed upon mutation of the structurally conserved PseG His17 to Phe or Leu suggesting that the enzymes share a mechanism; a catalytic His residue activating a water molecule for nucelophilic attack of the anomeric carbon to release UDP- in a concerted mechanism .[58] The crystal structure elucidated a well-ordered active site water, that appears anchored *via* hydrogen bonds to His17 and the main chain carbonyl of Ile13, as the proposed nucleophile for hydrolysis (**Scheme 6**). Docking the substrate into the active site in three energetically favorable free-substrate conformations led to convergence to a twist-boat as the lowest energy conformation in the ligand-enzyme complex.[58]

PseI; a Pse5Ac7Ac synthase

The final enzyme in the biosynthetic pathway, PseI was the first to be identified as being involved in the biosynthesis of Pse5Ac7Ac **2** as it mimics the final biosynthetic step for the production of sialic acids and hence displays homology to well-characterised Neu5Ac synthase sequences.[3] Genes displaying homology to Neu5Ac synthases were identified in *C. jejuni* and *NeuB3* was proposed as the most likely to be a Pse synthase as mutations to this gene prevented the correct formation of flagella.[65] *In vitro* assays confirmed that the *C. jejuni* NeuB3 enzyme displayed activity for conversion of 4-acetamido-4,6-dideoxy-α-L-AltNAc**11** into Pse5Ac7Ac **2** when incubated with PEP and was metal dependent.[66] The *H. pylori* homolog *HP0178* has also been confirmed to encode the PseI enzyme and utilised *in vitro* for the enzymatic synthesis of Pse5Ac7Ac **2**.[26]

Enzyme catalysed condensation reactions with PEP have been cited to proceed along one of two mechanistic pathways, for example, as in pyruvate kinases the PEP phosphorus could be released to afford a reactive pyruvate enolate ion which then could attack the substrate,[67, 68] or as in other nulO synthases such as in the biosynthesis of 3-deoxy-D-*manno-octulosonic acid* (KDO), PseI could catalyse attack of the PEP C3 onto the sugar open chain formyl and progess through an oxocarbenium ino intermediate and tetrahedral intermediate.[69-71] To establish which mechansim PseI catalysed, C2 ^{18}O labelled PEP was utilised and the reaction with 4-acetamido-4,6-dideoxy-α-L-AltNAc **11** monitored by ^{31}P NMR to track whether the O-P bond is cleaved during the mechansim. The results showed that the labelled oxygen remained bound to the released inorganic phosphate suggesting that PseI follows that of the KDO synthase.[66] Specifically, the ring open form of the substrate **27** is activated by a metal ion to aid nucleophilic attack of the PEP C3 to generate an oxocarbenium ion **28**. A free hydroxyl then readily attacks the carbonyl to afford the tetrahedral intermediate **29** which releases an inorganic phopshate, thus resulting in the formation of a ring opened nonulosonic acid **30** which can then ring close to afford Pse5Ac7Ac **2** (**Scheme 7**).[66]

Conclusions

In this review we have detailed the mechanistic and structural investigations into the enzymes integral for the biosynthesis of Pse5Ac7Ac **2** in bacteria, including pathogens *C. jejuni* and *H. pylori*. Using insights gleamed from enzymology studies these well characterised enzymes have previously been screened with small molecules for inhibition, with three validated PseB inhibitors displaying inhibition activity in cell-based assays.[72] It is our hope that this summary will serve to catalyse further explorations and design of potential therapeutics in this area.

Conflicts of Interest

None to declare

Acknowledgements

We thank the University of York and The Rosetrees Trust for funding (PhD award to HSC).

Notes and references

1. Y. A. Knirel, E. V. Vinogradov, V. L. L'Vov, N. A. Kocharova, A. S. Shashkov, B. A. Dmitriev and N. K. Kochetkov, *Carbohydr. Res.*, 1984, **133**, C5-C8.
2. M. Schirm, I. C. Schoenhofen, S. M. Logan, K. C. Waldron and P. Thibault, *Anal. Chem.*, 2005, **77**, 7774-7782.
3. M. Schirm, E. C. Soo, A. J. Aubry, J. Austin, P. Thibault and S. M. Logan, *Mol. Microbiol.*, 2003, **48**, 1579-1592.
4. P. Hitchen, J. Brzostek, M. Panico, J. A. Butler, H. R. Morris, A. Dell and D. Linton, *Microbiology (Reading, U. K.)*, 2010, **156**, 1953-1962.
5. Z. Li, S. Hwang, J. Ericson, K. Bowler and M. Bar-Peled, *J. Biol. Chem.*, 2015, **290**, 691-704.
6. Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, N. K. Kochetkov, V. L'Vov and B. A. Dmitriev, *Carbohydr. Res.*, 1985, **141**, C1-C5.
7. M. Zunk and M. J. Kiefel, *RSC Adv.*, 2014, **4**, 3413-3421.
8. Y. A. Knirel, E. V. Vinogradov, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavsky and G. M. Mashilova, *Eur. J. Biochem.*, 1986, **157**, 129-138.
9. M. Staaf, A. Weintraub and G. Widmalm, *Eur. J. Biochem.*, 1999, **263**, 656-661.
10. A. M. GIL-SERRANO, M. A. RODRÍGUEZ-CARVAJAL, P. TEJERO-MATEO, J. L. ESPARTERO, M. MENENDEZ, J. CORZO, J. E. RUIZ-SAINZ and A. M. BUENDÍA-CLAVERÍA, *Biochem. J*, 1999, **342**, 527-535.
11. P. Castric, F. J. Cassels and R. W. Carlson, *J. Biol. Chem.*, 2001, **276**, 26479-26485.
12. Y. A. Knirel, N. A. Kocharova, A. S. Shashkov, B. A. Dmitriev, N. K. Kochetkov, E. S. Stanislavskii and G. M. Mashilova, *Eur. J. Biochem.*, 1987, **163**, 639-652.
13. P. Thibault, S. M. Logan, J. F. Kelly, J.-R. Brisson, C. P. Ewing, T. J. Trust and P. Guerry, *J. Biol. Chem.*, 2001, **276**, 34862-34870.
14. C. P. Ewing, E. Andreishcheva and P. Guerry, *J. Bacteriol.*, 2009, **191**, 7086-7093.
15. K. M. Ottemann and A. C. Lowenthal, *Infect. Immun.*, 2002, **70**, 1984-1990.
16. S. B. Baldvinsson, M. C. H. Sørensen, C. S. Vegge, M. R. Clokie and L. Brøndsted, *Appl. Environ. Microbiol.*, 2014, AEM. 02057-02014.
17. E. Andersen-Nissen, K. D. Smith, K. L. Strobe, S. L. R. Barrett, B. T. Cookson, S. M. Logan and A. Aderem, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 9247-9252.
18. H. N. Stephenson, H. Jones, E. Milioris, A. Copland, M. Bajaj-Elliott, D. C. Mills, N. Dorrell, B. W. Wren, P. R. Crocker and D. Escors, *J. Infect. Dis.*, 2014, **210**, 1487-1498.
19. T. Angata and A. Varki, *Chem. Rev.*, 2002, **102**, 439-470.
20. W. F. Vann, D. A. Daines, A. S. Murkin, M. E. Tanner, D. O. Chaffin, C. E. Rubens, J. Vionnet and R. P. Silver, *J. Bacteriol.*, 2004, **186**, 706-712.
21. W. F. Vann, J. J. Tavarez, J. Crowley, E. Vimr and R. P. Silver, *Glycobiology*, 1997, **7**, 697-701.
22. S. Hinderlich, R. Stäsche, R. Zeitler and W. Reutter, *J. Biol. Chem.*, 1997, **272**, 24313-24318.
23. C. Creuzenet, M. J. Schur, J. Li, W. W. Wakarchuk and J. S. Lam, *J. Biol. Chem.*, 2000, **275**, 34873-34880.
24. I. C. Schoenhofen, D. J. McNally, E. Vinogradov, D. Whitfield, N. M. Young, S. Dick, W. W. Wakarchuk, J.-R. Brisson and S. M. Logan, *J. Biol. Chem.*, 2006, **281**, 723-732.
25. P. Guerry, C. P. Ewing, M. Schirm, M. Lorenzo, J. Kelly, D. Pattarini, G. Majam, P. Thibault and S. Logan, *Mol. Microbiol.*, 2006, **60**, 299-311.
26. I. C. Schoenhofen, D. J. McNally, J.-R. Brisson and S. M. Logan, *Glycobiology*, 2006, **16**, 8C-14C.
27. S. M. B. Tabei, P. G. Hitchen, M. J. Day-Williams, S. Merino, R. Vart, P.-C. Pang, G. J. Horsburgh, S. Viches, M. Wilhelms, J. M. Tomas, A. Dell and J. G. Shaw, *J. Bacteriol.*, 2009, **191**, 2851-2863.
28. N. Ishiyama, C. Creuzenet, W. L. Miller, M. Demendi, E. M. Anderson, G. Harauz, J. S. Lam and A. M. Berghuis, *J. Biol. Chem.*, 2006, **281**, 24489-24495.
29. C. Creuzenet, M. Belanger, W. W. Wakarchuk and J. S. Lam, *J. Biol. Chem.*, 2000, **275**, 19060-19067.
30. H. Jörnvall, B. Persson, M. Krook, S. Atrian, R. Gonzalez-Duarte, J. Jeffery and D. Ghosh, *Biochemistry*, 1995, **34**, 6003-6013.
31. D. J. McNally, J. P. M. Hui, A. J. Aubry, K. K. K. Mui, P. Guerry, J.-R. Brisson, S. M. Logan and E. C. Soo, *J. Biol. Chem.*, 2006, **281**, 18489-18498.
32. J. P. Morrison, I. C. Schoenhofen and M. E. Tanner, *Bioorg. Chem.*, 2008, **36**, 312-320.
33. D. J. McNally, I. C. Schoenhofen, R. S. Houliston, N. H. Khieu, D. M. Whitfield, S. M. Logan, H. C. Jarrell and J.-R. Brisson, *ChemMedChem*, 2008, **3**, 55-59.
34. P. K. MEHTA, T. I. HALE and P. CHRISTEN, *Eur. J. Biochem.*, 1993, **214**, 549-561.
35. I. C. Schoenhofen, V. V. Lunin, J.-P. Julien, Y. Li, E. Ajamian, A. Matte, M. Cygler, J.-R. Brisson, A. Aubry, S. M. Logan, S. Bhatia, W. W. Wakarchuk and N. M. Young, *J. Biol. Chem.*, 2006, **281**, 8907-8916.
36. J. Catazaro, A. Caprez, A. Guru, D. Swanson and R. Powers, *Proteins*, 2014, **82**, 2597-2608.
37. B. W. Noland, J. M. Newman, J. Hendle, J. Badger, J. A. Christopher, J. Tresser, M. D. Buchanan, T. A. Wright, M. E. Rutter, W. E. Sanderson, H.-J. Müller-Dieckmann, K. S. Gajiwala and S. G. Buchanan, *Structure*, 2002, **10**, 1569-1580.
38. R. K. Obhi and C. Creuzenet, *J. Biol. Chem.*, 2005, **280**, 20902-20908.
39. J. N. Jansonius, *Curr. Opin. Struct. Biol.*, 1998, **8**, 759-769.
40. J. M. Reimer, I. Harb, O. G. Ovchinnikova, J. Jiang, C. Whitfield and T. M. Schmeing, *ACS Chem. Biol.*, 2018, **13**, 3161-3172.
41. Y. C. Liu, A. I. Ud-Din and A. Roujeinikova, *Acta Crystallogr., Sect. F: Struct. Biol. Commun.*, 2014, **70**, 1276-1279.
42. F. Dyda, D. C. Klein and A. B. Hickman, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 81-103.

43. L. Favrot, J. S. Blanchard and O. Vergnolle, *Biochemistry*, 2016, **55**, 989-1002.
44. M. W. Vetting, L. P. S. de Carvalho, M. Yu, S. S. Hegde, S. Magnet, S. L. Roderick and J. S. Blanchard, *Arch. Biochem. Biophys.*, 2005, **433**, 212-226.
45. W. S. Song, M. S. Nam, B. Namgung and S.-i. Yoon, *Biochem. Biophys. Res. Commun.*, 2015, **458**, 843-848.
46. A. I. Ud-Din, Y. C. Liu and A. Roujeinikova, *PLoS One*, 2015, **10**, e0115634/0115631-e0115634/0115614.
47. M.-N. Hung, E. Rangarajan, C. Munger, G. Nadeau, T. Sulea and A. Matte, *J. Bacteriol.*, 2006, **188**, 5606-5617.
48. V. Agarwal, A. Metlitskaya, K. Severinov and S. K. Nair, *The Journal of biological chemistry*, 2011, **286**, 21295-21303.
49. A. B. Hickman, M. A. A. Namboodiri, D. C. Klein and F. Dyda, *Cell*, 1999, **97**, 361-369.
50. J. R. Rojas, R. C. Trievel, J. Zhou, Y. Mo, X. Li, S. L. Berger, C. D. Allis and R. Marmorstein, *Nature*, 1999, **401**, 93-98.
51. J. De Angelis, J. Gastel, D. C. Klein and P. A. Cole, *J. Biol. Chem.*, 1998, **273**, 3045-3050.
52. K. G. Tanner, R. C. Trievel, M.-H. Kuo, R. M. Howard, S. L. Berger, C. D. Allis, R. Marmorstein and J. M. Denu, *J. Biol. Chem.*, 1999, **274**, 18157-18160.
53. A. Schuetz, G. Bernstein, A. Dong, T. Antoshenko, H. Wu, P. Loppnau, A. Bochkarev and A. N. Plotnikov, *Proteins: Structure, Function, and Bioinformatics*, 2007, **68**, 403-407.
54. K. G. Tanner, M. R. Langer, Y. Kim and J. M. Denu, *J. Biol. Chem.*, 2000, **275**, 22048-22055.
55. Y. Yan, S. Harper, D. W. Speicher and R. Marmorstein, *Nat. Struct. Biol.*, 2002, **9**, 862-869.
56. P. M. Coutinho, E. Deleury, G. J. Davies and B. Henrissat, *J. Mol. Biol.*, 2003, **328**, 307-317.
57. F. Liu and M. E. Tanner, *J. Biol. Chem.*, 2006, **281**, 20902-20909.
58. E. S. Rangarajan, A. Proteau, Q. Cui, S. M. Logan, Z. Potetinova, D. Whitfield, E. O. Purisima, M. Cygler, A. Matte, T. Sulea and I. C. Schoenhofen, *J. Biol. Chem.*, 2009, **284**, 20989-21000.
59. A. S. Murkin, W. K. Chou, W. W. Wakarchuk and M. E. Tanner, *Biochemistry*, 2004, **43**, 14290-14298.
60. L. L. Lairson and S. G. Withers, *Chem. Commun.*, 2004, DOI: 10.1039/B406490A, 2243-2248.
61. L. E. t. Metzger and C. R. H. Raetz, *Biochemistry*, 2010, **49**, 6715-6726.
62. K. J. Babinski, A. A. Ribeiro and C. R. H. Raetz, *J. Biol. Chem.*, 2002, **277**, 25937-25946.
63. Y. Hu, L. Chen, S. Ha, B. Gross, B. Falcone, D. Walker, M. Mokhtarzadeh and S. Walker, *Proceedings of the National Academy of Sciences*, 2003, **100**, 845-849.
64. A. Bouhss, A. E. Trunkfield, T. D. H. Bugg and D. Mengin-Lecreulx, *FEMS Microbiol. Rev.*, 2007, **32**, 208-233.
65. D. Linton, A. V. Karlyshev, P. G. Hitchen, H. R. Morris, A. Dell, N. A. Gregson and B. W. Wren, *Mol. Microbiol.*, 2000, **35**, 1120-1134.
66. W. K. Chou, S. Dick, W. W. Wakarchuk and M. E. Tanner, *J. Biol. Chem.*, 2005, **280**, 35922-35928.
67. A. Hassett, W. Blaettler and J. R. Knowles, *Biochemistry*, 1982, **21**, 6335-6340.
68. A. M. Reynard, L. F. Hass, D. D. Jacobsen and P. D. Boyer, *J. Biol. Chem.*, 1961, **236**, 2277-2283.
69. P. H. Liang, J. Lewis, K. S. Anderson, A. Kohen, F. W. D'Souza, Y. Benenson and T. Baasov, *Biochemistry*, 1998, **37**, 16390-16399.
70. J. Wang, H. S. Duewel, R. W. Woodard and D. L. Gatti, *Biochemistry*, 2001, **40**, 15676-15683.
71. D. L. Howe, A. K. Sundaram, J. Wu, D. L. Gatti and R. W. Woodard, *Biochemistry*, 2003, **42**, 4843-4854.
72. R. Menard, I. C. Schoenhofen, L. Tao, A. Aubry, P. Bouchard, C. W. Reid, P. Lachance, S. M. Twine, K. M. Fulton, Q. Cui, H. Hogues, E. O. Purisima, T. Sulea and S. M. Logan, *Antimicrob. Agents Chemother.*, 2014, **58**, 7430-7440, 7412 pp.

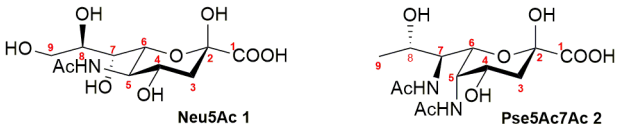


Figure 1. The structurally related nonulosonic acids; the common sialic acid Neu5Ac **1** and rare bacterial sugar Pse5Ac7Ac **2**.

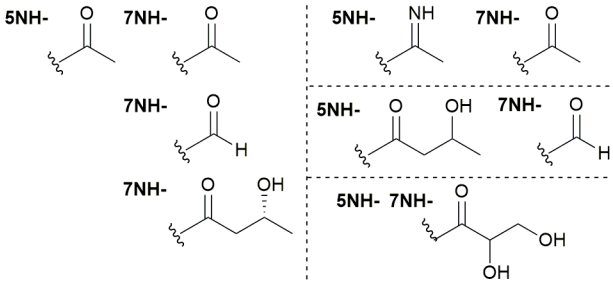


Figure 2. Examples of different acetaminido functionality observed at the C5 and C7 of pseudaminic acid structures identified in bacteria.

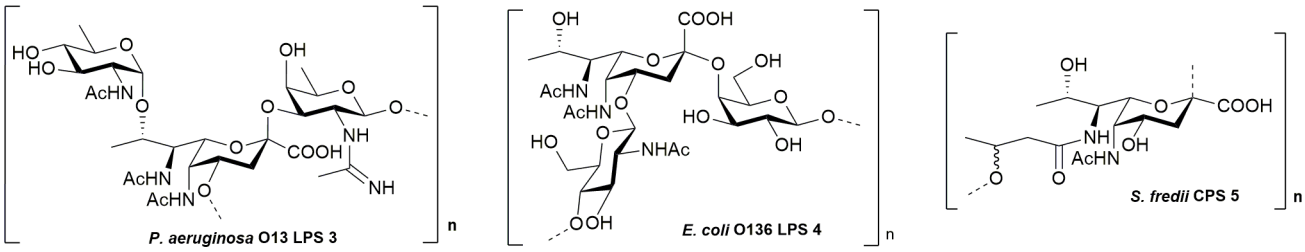
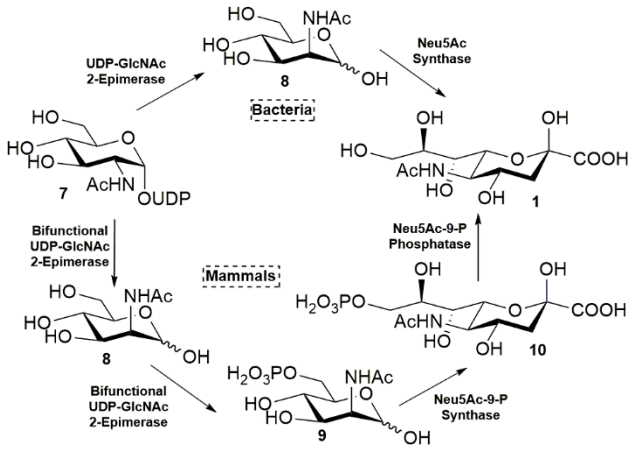
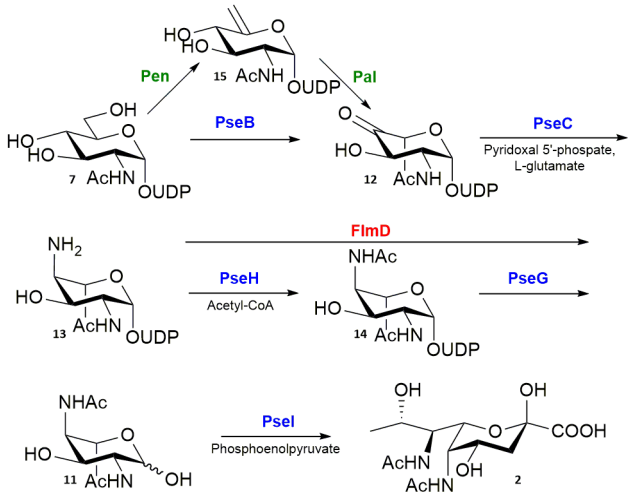


Figure 3. Examples of pseudaminic acid cell surface structures, showing a selection of the different glycosidic linkages utilised in nature.



Scheme 1. The biosynthetic pathway from UDP-GlcNAc **7** to Neu5Ac **1** in bacteria and mammals.



Scheme 2. Pse5Ac7Ac **2** biosynthesis; detailing the intermediates identified *in C. jejuni* and *H. pylori* (blue) and enzymatic deviations occurring in *A. caviae* (red) and *B. thuringiensis* (green) highlighted.

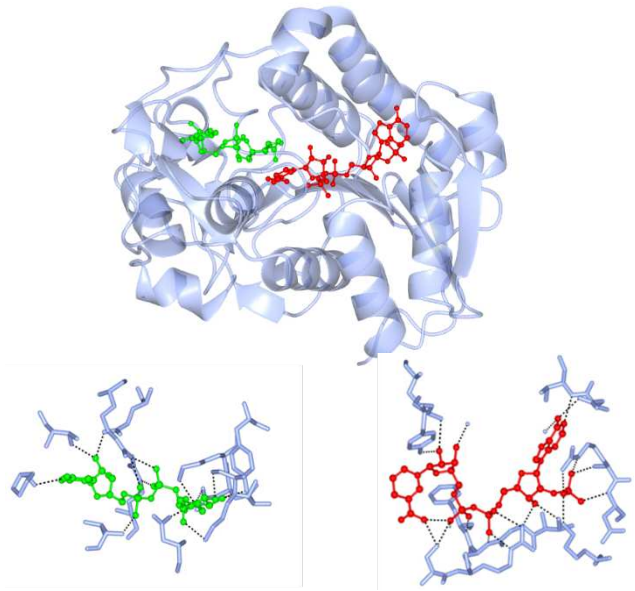
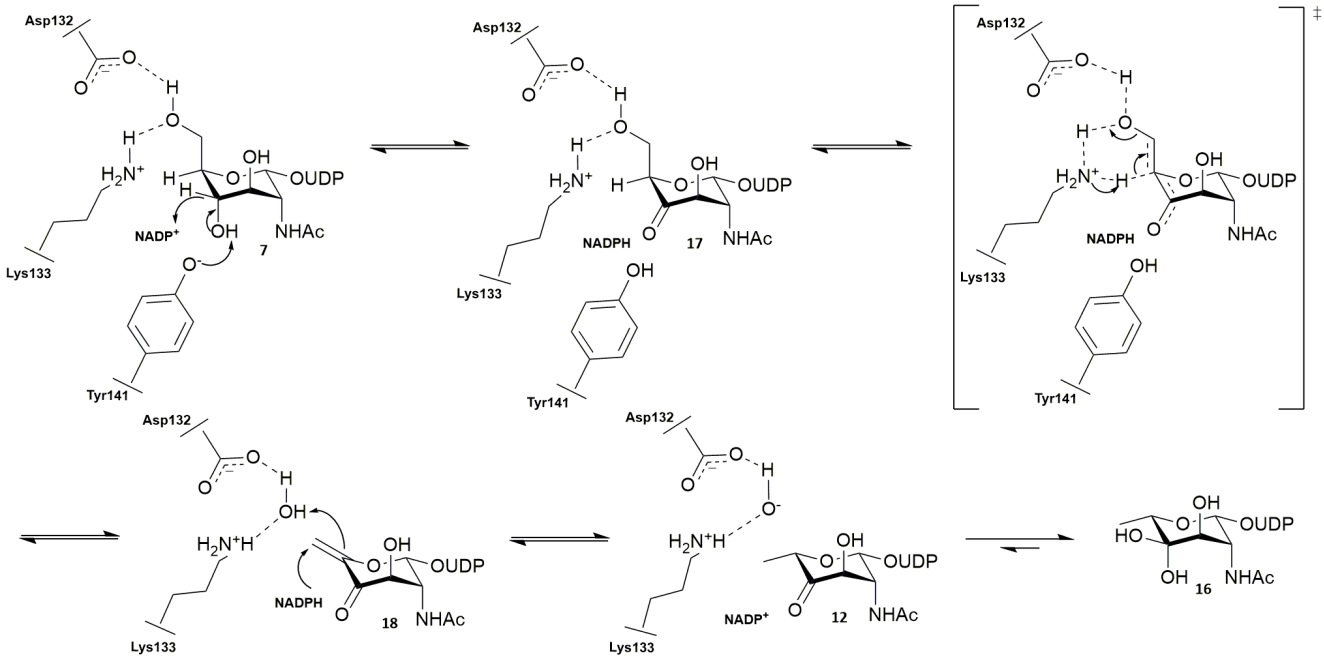


Figure 4. *H. pylori* PseB (PDB 2GN4) crystal structure in complex with the UDP-GlcNAc substrate **7** (green) and the tightly bound NADPH co-factor (red) and close-ups of the substrate and co-factor binding sites displaying the numerous hydrogen bonds between molecules and enzyme.



Scheme 3. PseB catalysed oxidation, dehydration and reduction of the substrate UDP-GlcNAc **7** to form the initial PseB product UDP-4-keto-6-deoxy-L-IdoNAc **12** which is in equilibrium with the hydrated form **16** in aqueous conditions.

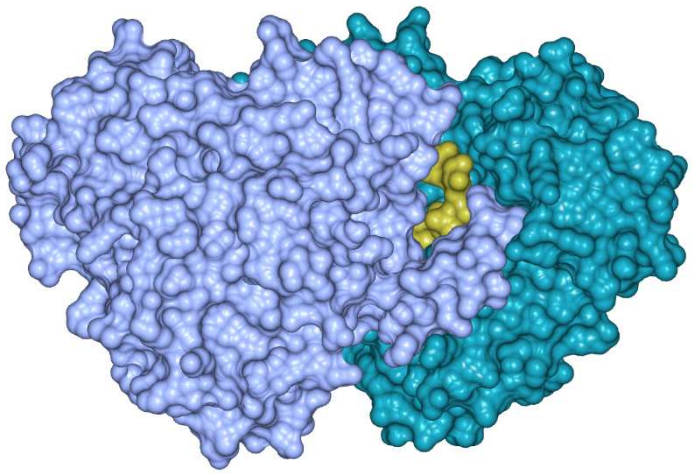


Figure 5. The *H. pylori* PseC (PDB 2FNU) homodimer surface structure (chain A dark cyan, chain B ice blue), with the active site depicted at the dimer interface with the bound external aldimine intermediate **19** (gold surface).

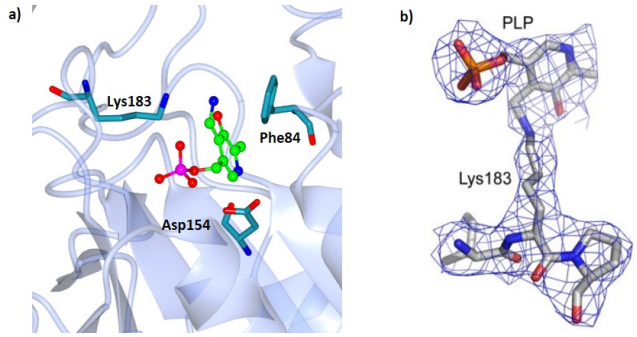


Figure 6. *H. pylori* PseC crystal studies (PDB 2FN6) **a)** the co-factor binding site in complex with the PLP co-factor **19**, highlighting Type 1 aminotransferase conserved residues and **b)** electron density of the PLP-enzyme internal aldimine intermediate **20** (figure adapted from the original paper).

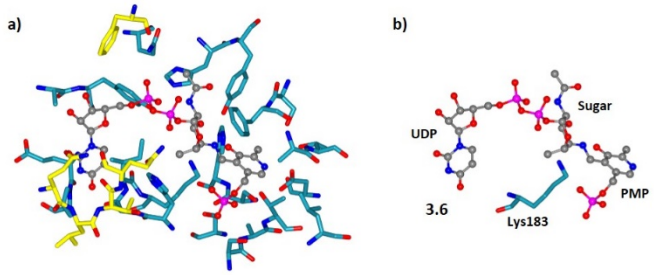
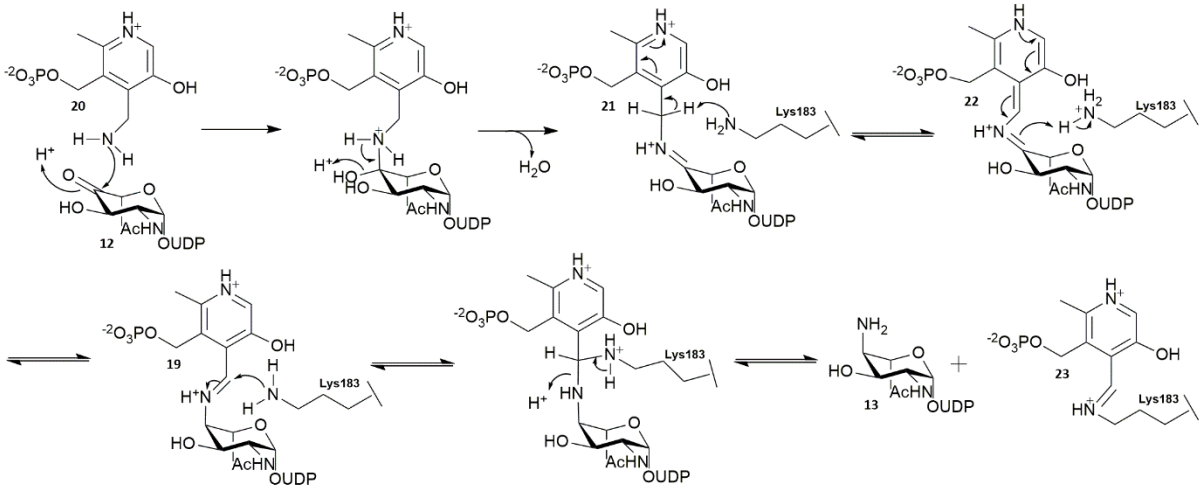


Figure 7. The PMP-substrate external aldimine (ball and stick model) **a)** in complex with the surrounding H. pylori PseC residues (cyan chainA, yellow chain B) and **b)** with the catalytic Lys183 residue.



Scheme 4. The proposed PseC mechanism transferring an amino group from the PMP co-factor, generated *in situ*, to the keto-sugar **13**. Nucleophillic attack of the PMP group to C4 followed by release of water generates a substrate-co-factor ketimine intermediate **21**. The catalytic base residue Lys183 is then employed to aid in the mechanistic progression of amino transfer through the quinoid intermediate **22**, and the external aldimine **19** before release of the product UDP-4-amino-4,6-dideoxy-β-L-AltNac **13**.

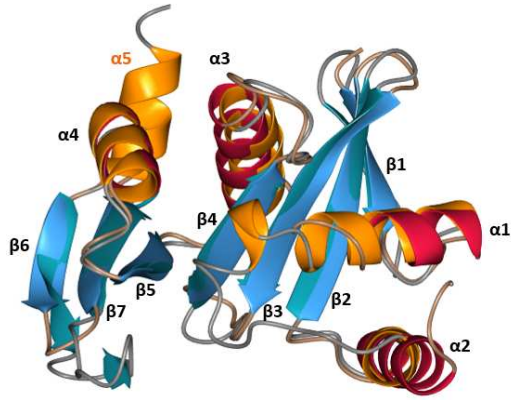


Figure 8 Overlay of the *C. jejuni* (PDB 4XPL) and *H. pylori* (PDB 4RI1) PseH crystal structures, highlighting the conservation of secondary structure features and the single additional *H. pylori* alpha helix (α5). *C. jejuni* PseH; α helix - crimson, β sheet – light blue, loop – pale brown. *H. pylori* PseH; α helix – dark orange, β sheet – dark cyan, loop – grey..

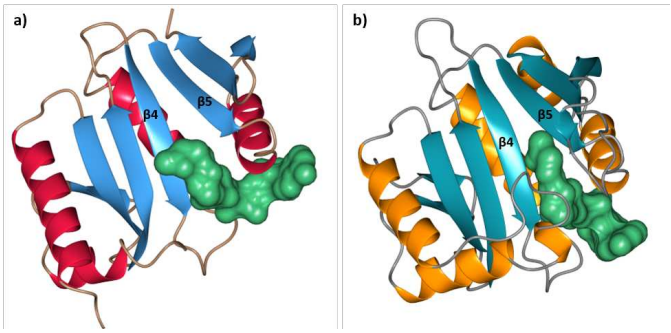
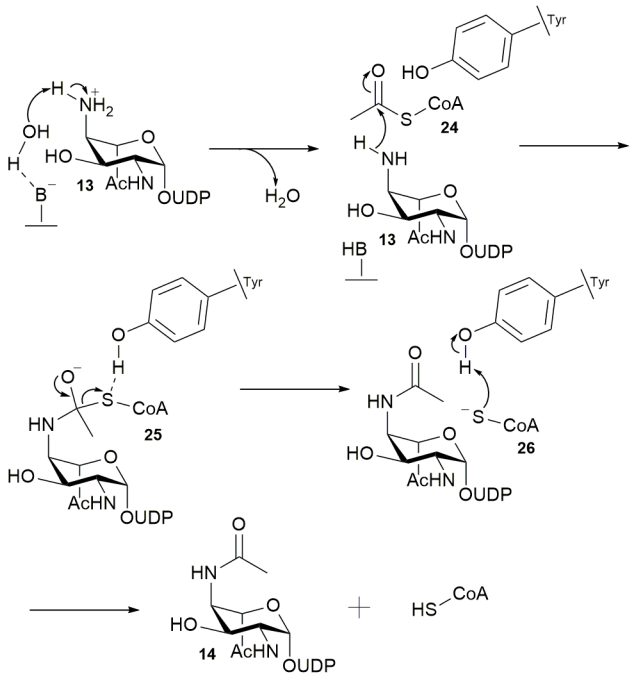


Figure 9 PseH crystal structures **a)** *C. jejuni* (PDB 4XPL) and **b)** *H. pylori* (PDB 4RI1), co-crystallised with the acetyl-coA co-factor (green surface) with the thio-ester buried within the cleft between the β-4 and β-5 sheets.



Scheme 5 Acetyltransfer mechanism for synthesis of UDP-4-acetamido-4,6-dideoxy-β-L-AltNAc **14** showing the proposed water mediated amine deprotonation to afford nucleophilic attack on acetyl-CoA **24**. An active site tyrosine is also depicted as it has been proposed as the catalytic acid utilised to stabilise intermediates and the anionic CoA byproduct **26**.

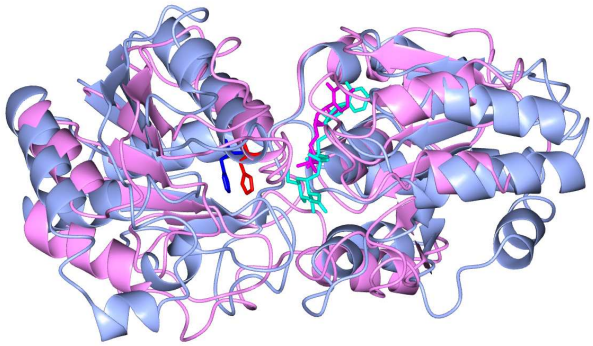
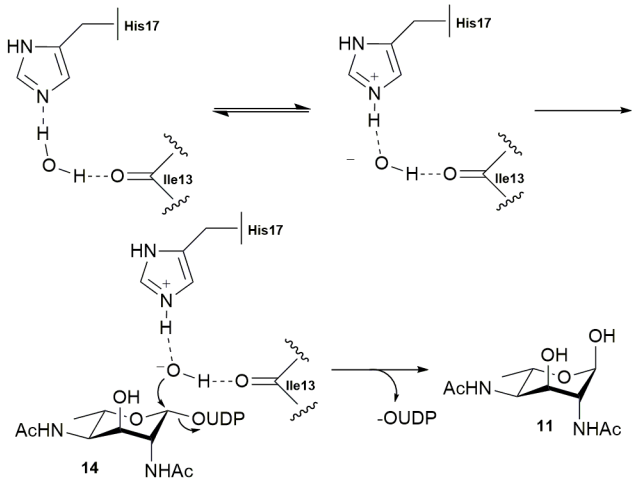
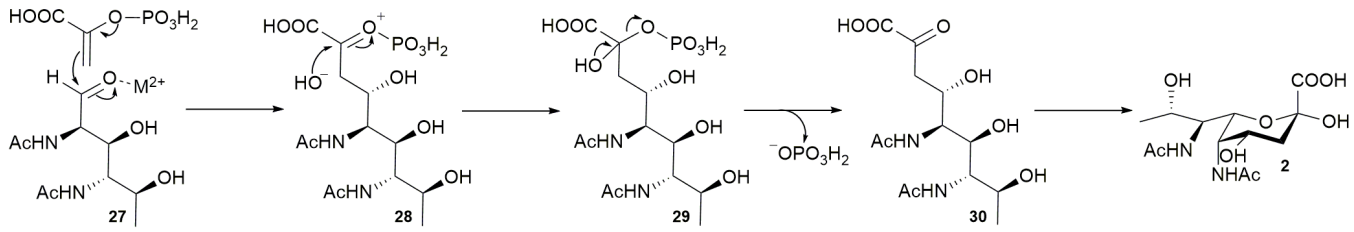


Figure 10 Overlay of ligand bound crystal structures with the putative catalytic histidine residues highlighted (PseG His17 blue, MurG His19 red) proximal to the ligand binding site. PseG (pink) with bound UDP (magenta, PDB 3HBN) and MurG (ice blue) with bound UDP-GlcNAc (cyan, PDB 1NLM).



Scheme 6 The proposed mechanism for PseG; His17 is predicted to behave as a catalytic base, deprotonating a water molecule anchored by the Ile13 backbone carbonyl to generate a hydroxide nucleophile. The hydroxide attacks the anomeric carbon of the substrate **14** hydrolysing the UDP group to generate the product **11** with inversion of stereochemistry (molecular modelling of the substrate into the crystal structures suggest it binds in a twist-boat conformation however this has been omitted for simplicity of the scheme).



Scheme 7 *C. jejuni* PseI pseudaminic acid synthetase PEP condensation mechanism; PEP undergoes nucleophilic attack on the open form of the substrate **27** to afford an oxocarbenium **28** . Attack by a hydroxyl generates the tetrahedral intermediate **29** which releases an inorganic phosphate to afford the desired nonulosonic acid **30** that ring closes as the desired Pse5Ac7Ac **2**.