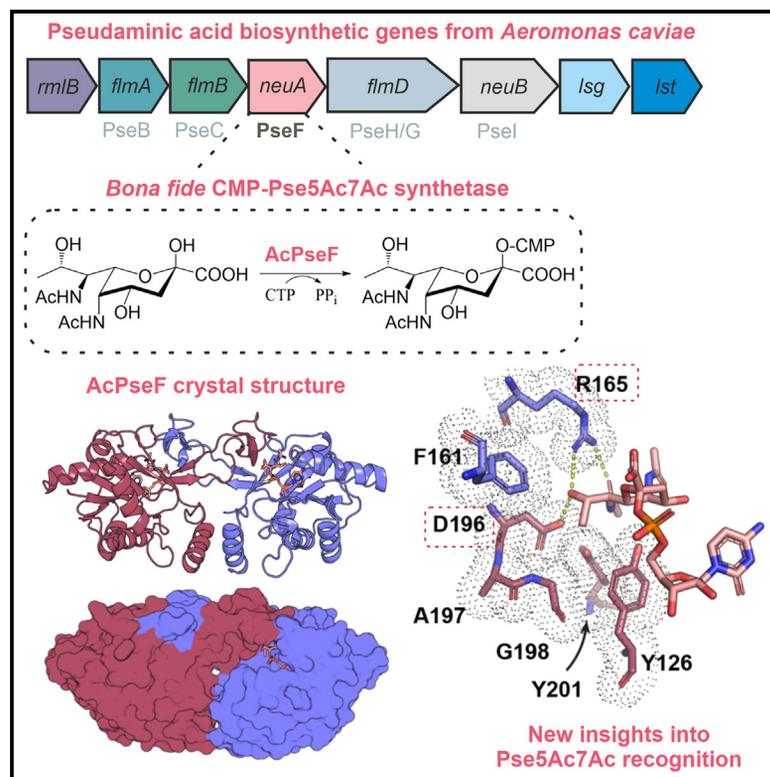


# Structure

## Structural dissection of the CMP-pseudaminic acid synthetase, PseF

### Graphical abstract



### Authors

Tessa Keenan, Andrew R. Cowan, Emily K.P. Flack, ..., Gavin H. Thomas, Glyn R. Hemsworth, Martin A. Fascione

### Correspondence

g.r.hemsworth@leeds.ac.uk (G.R.H.), martin.fascione@york.ac.uk (M.A.F.)

### In brief

Keenan, Cowan et al. determined crystal structures of the CMP-Pse5Ac7Ac synthetase, AcPseF, in the ligand-free and product-bound forms. They provide molecular-level insights into how a Pse5Ac7Ac-processing enzyme might recognize and discriminate Pse5Ac7Ac from other ulosonic acids, which may inform the development of novel chemical tools or even antimicrobial inhibitors.

### Highlights

- Crystal structures of AcPseF in the Apo and product-bound states were determined
- AcPseF displays metal-dependent activity over a broad pH and temperature range
- Pse5Ac7Ac is positioned within a hydrophobic pocket in the AcPseF active site
- AcPseF binds the CMP-Pse5Ac7Ac side chain in the lowest energy *tg* conformation



## Article

# Structural dissection of the CMP-pseudaminic acid synthetase, PseF

Tessa Keenan,<sup>1,4</sup> Andrew R. Cowan,<sup>2,4</sup> Emily K.P. Flack,<sup>3</sup> Natasha E. Hatton,<sup>1</sup> Abigail J. Walklett,<sup>1</sup> Gavin H. Thomas,<sup>3</sup> Glyn R. Hemsworth,<sup>2,\*</sup> and Martin A. Fascione<sup>1,5,\*</sup>

<sup>1</sup>Department of Chemistry, University of York, York YO10 5DD, UK

<sup>2</sup>Astbury Centre for Structural Molecular Biology and School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

<sup>3</sup>Department of Biology, University of York, York YO10 5DD, UK

<sup>4</sup>These authors contributed equally

<sup>5</sup>Lead contact

\*Correspondence: [g.r.hemsworth@leeds.ac.uk](mailto:g.r.hemsworth@leeds.ac.uk) (G.R.H.), [martin.fascione@york.ac.uk](mailto:martin.fascione@york.ac.uk) (M.A.F.)

<https://doi.org/10.1016/j.str.2024.09.017>

## SUMMARY

Pseudaminic acid is a non-mammalian sugar found in the surface glycoconjugates of many bacteria, including several human pathogens, and is a virulence factor thought to facilitate immune evasion. The final step in the biosynthesis of the nucleotide activated form of the sugar, CMP-Pse5Ac7Ac is performed by a CMP-Pse5Ac7Ac synthetase (PseF). Here we present the biochemical and structural characterization of PseF from *Aeromonas caviae* (AcPseF), with AcPseF displaying metal-dependent activity over a broad pH and temperature range. Upon binding to CMP-Pse5Ac7Ac, AcPseF undergoes dynamic movements akin to other CMP-ulosonic acid synthetases. The enzyme clearly discriminates Pse5Ac7Ac from other ulosonic acids, through active site interactions with side-chain functional groups and by positioning the molecule in a hydrophobic pocket. Finally, we show that AcPseF binds the CMP-Pse5Ac7Ac side chain in the lowest energy conformation, a trend that we observed in the structures of other enzymes of this class.

## INTRODUCTION

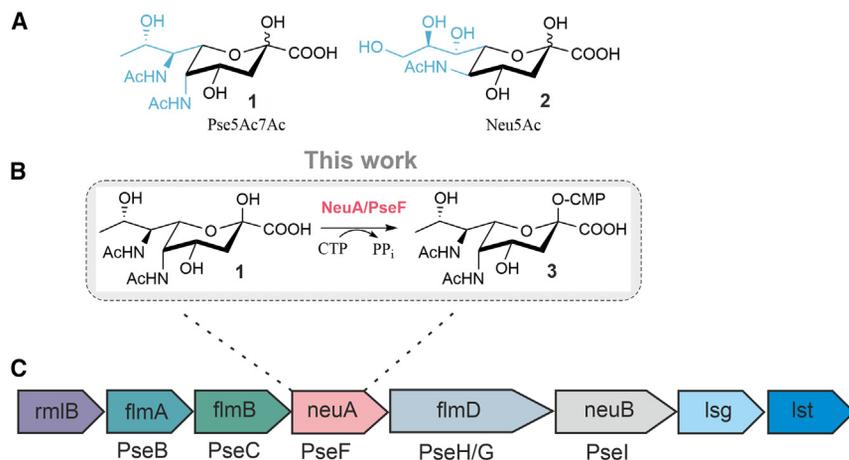
5,7-Diacetyl pseudaminic acid (Pse5Ac7Ac, **1**) and its analogs are non-mammalian sugars found on the surface glycoconjugates of many bacteria, including several pathogenic species on the WHO global priority pathogens list (*Pseudomonas aeruginosa*, *Acinetobacter baumannii*, *Campylobacter jejuni*, and *Helicobacter pylori*).<sup>1–9</sup> As a key component of structural glycans (lipopolysaccharide and capsular polysaccharide) and glycoproteins (flagellin and pilin), pseudaminic acid is an important virulence factor and even required for the correct assembly of flagella and motility in some species.<sup>1,4,10</sup> Additionally, Pse glycans have a proposed role in host immune evasion that is particularly intriguing as their structural similarity to abundant eukaryotic sialic acids such as *N*-acetyl neuraminic acid Neu5Ac, **2** (Figure 1A) has been suggested to facilitate “molecular mimicry” and dampening of the host immune response.<sup>11,12</sup>

The biosynthesis of CMP-Pse5Ac7Ac **3**, the glycosyl donor substrate of Pse glycosyltransferases, is well understood in bacteria, beginning with UDP-*N*-acetylglucosamine (UDP-GlcNAc) and proceeding through an efficient multi-enzyme pathway (Figure S9).<sup>4,13,14</sup> The final step in the biosynthetic pathway involves the nucleotide activation of Pse5Ac7Ac **1** by PseF, a CMP-

Pse5Ac7Ac synthetase (Figure 1B). PseF catalyzes the formation of CMP-Pse5Ac7Ac **3** using cytidine triphosphate (CTP) as a cofactor while releasing pyrophosphate as a by-product. In *C. jejuni*, the *psef* gene is located in the flagellum glycosylation locus and *psef* defective strains were shown to be non-motile.<sup>15</sup> Similarly, in the opportunistic pathogen *Aeromonas caviae*, *neuA* (which encodes AcPseF as referred to herein) forms part of a Pse5Ac7Ac biosynthetic gene cluster (*flm* locus) located in a genomic island encoding *O*-antigen biosynthesis and flagellin glycosylation, with mutants defective in the gene displaying loss of motility, flagella, and LPS *O*-antigen (Figure 1C).<sup>16,17</sup> We previously demonstrated the *in vitro* activity of AcPseF as a synthetase in the chemoenzymatic synthesis of CMP-Pse5Ac7Ac **3** and an azide-functionalized analog,<sup>18–20</sup> and *A. baumannii* PseF (AbPseF) was also recently demonstrated to be active on Pse5Ac7Ac, in addition to formyl and hydroxybutyryl-functionalized analogs.<sup>21</sup> Yet, despite our increasing understanding of the activity of Pse processing enzymes, very little is understood at the molecular level about how Pse is recognized by macromolecules, with currently no structures in complex with a Pse sugar in the PDB.

Akin to PseF, CMP sialic acid synthetases use CTP to activate their substrates for subsequent use. Several crystal structures of bacterial CMP-ulosonic synthetases have been published, such





**Figure 1. The biosynthesis of CMP-Pse5Ac7Ac is mediated by a CMP-Pse5Ac7Ac synthetase**

(A) Structural comparison of Pse5Ac7Ac **1** and Neu5Ac **2**.

(B) Reaction mediated by NeuA, a CMP-Pse5Ac7Ac synthetase to afford nucleotide donor  $\alpha$ -CMP-Pse5Ac7Ac **3**.

(C) Organization of the CMP-Pse5Ac7Ac biosynthetic genes in *A. caviae*, adapted from Tabei et al.<sup>17</sup> *rmlB*, dTDP-D-glucose-4,6-dehydratase; *flmA*, dehydratase/epimerase; *flmB*, aminotransferase; *neuA*, CMP-Pse5Ac7Ac synthetase; *flmD*, bifunctional N-acetyltransferase and nucleotidase; *neuB*, Pse5Ac7Ac synthetase; *lsg*, flippase; *lst*, glycosyltransferase.

as the CMP-Neu5Ac synthetases from *Neisseria meningitidis* (NmCSS) and *Vibrio cholerae* (VbCSS),<sup>22–24</sup> and the CMP-KDO synthetase from *Escherichia coli* (EcCKS), which have facilitated mechanistic studies of these enzymes.<sup>25,26</sup> Overall, sialic acid synthetases appear to have a conserved mechanism involving an “open” and “closed” conformational transition, which occurs upon binding of the CTP and sialic acid substrates in the active site. Catalysis by these enzymes is dependent on two  $Mg^{2+}$  ions positioned by two active-site aspartic acid residues (Asp209 and Asp211 in NmCSS) constituting the DXD motif,<sup>24</sup> which appears to be critical for the correct positioning of the CTP  $\alpha$ -phosphate for activation. In CMP-Neu5Ac synthetases, the  $Mg^{2+}$  ions are also thought to co-ordinate to the C2-OH group in Neu5Ac, facilitating deprotonation of this hydroxyl by a general base hydroxide ion.<sup>24</sup> A similar mechanism has been proposed for CMP-Kdo synthetases.<sup>27</sup> *In silico* study of the *H. pylori* CMP-Pse5Ac7Ac synthetase (HpPseF) has predicted the overall fold of the enzyme and putative active site residues, based on published NmCSS structures,<sup>28</sup> with the authors postulating the enzyme may follow a similar catalytic mechanism to other CMP-sialic acid synthetases.

Herein we report the biochemical characterization and structural analysis of the CMP-Pse5Ac7Ac synthetase from *A. caviae* (AcPseF). Our structures in the ligand-free state (Apo) and in complex with the product CMP-Pse5Ac7Ac **3** reveal how this sugar interacts with this protein for the first time and spotlight the importance of conformational changes in this enzyme family. We highlight residues of likely importance for recognition of Pse5Ac7Ac that may help to differentiate the sugar from other nonulosonic acids. Finally, we show that the Pse5Ac7Ac moiety adopts the low energy *trans-gauche* (*tg*) side chain conformation around the C6–C7 bond, which could have important implications for rational design of PseF inhibitors.

## RESULTS

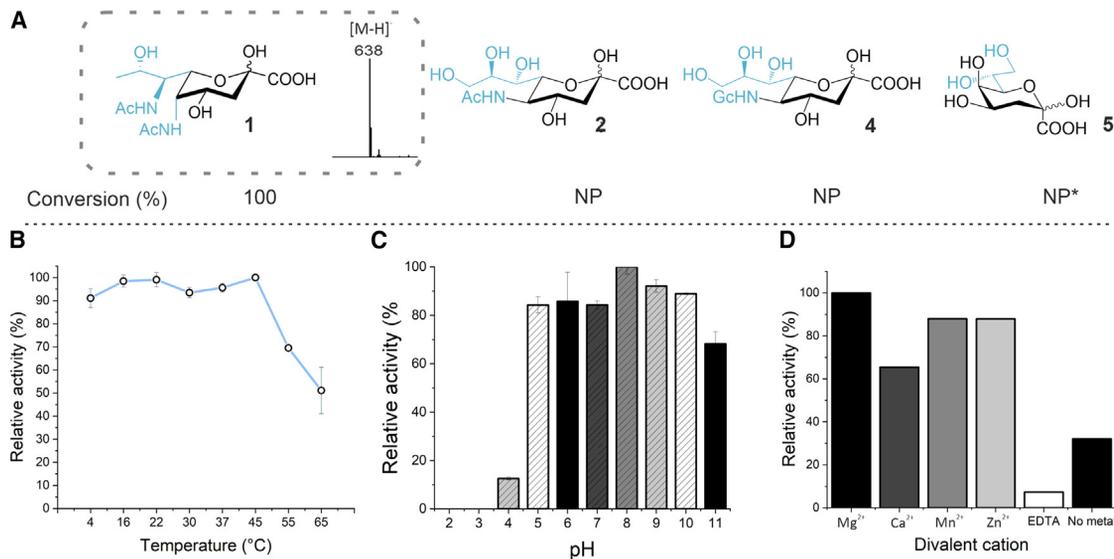
### Biochemical and kinetic characterization of AcPseF

Recombinant AcPseF was produced in *E. coli* BL21 (DE3) and purified to homogeneity. To investigate AcPseF as a *bona fide* CMP-Pse5Ac7Ac synthetase, the activity of the enzyme toward Pse5Ac7Ac **1**, Neu5Ac **2**, Neu5Gc **4**, and KDO **5** was determined

in the presence of cofactor cytidine triphosphate (CTP) using our previously established methodology.<sup>18</sup> Following analysis by negative high performance liquid chromatography electrospray ionisation mass spectrometry (ESI-LC-MS), we observed full conversion of Pse5Ac7Ac **1** to CMP-Pse5Ac7Ac **3** ( $[M-H]^-$   $m/z$  638) after 30 min (Figures 2A and S2). No conversion of Neu5Ac **2** or Neu5Gc **4** was observed (Figures S4 and S5), while trace-level conversion of KDO **5** was observed in reactions with 140  $\mu$ M AcPseF (Figure S3). Overall, our results suggest that AcPseF is indeed a dedicated CMP-Pse5Ac7Ac synthetase.

To determine the conditions for optimal AcPseF activity, the effects of temperature, pH, and metal ions were investigated. The optimum temperature was found to be 45°C, while the enzyme displayed high activity between 4°C and 45°C, which rapidly declined at higher temperatures (Figure 2B). AcPseF exhibited activity over a broad pH range, with >80% relative activity observed in the pH range 5–10 and the highest activity observed at pH 8 (Figure 2C). AcPseF displayed metal-dependent activity, with the highest activity in the presence of  $Mg^{2+}$ ,  $Mn^{2+}$ , and  $Zn^{2+}$ . Activity observed in the absence of divalent cation was consistent with preloading of the enzyme active site, most likely with  $Mg^{2+}$  which we found to be required for enzyme stability and was therefore added to the protein purification buffers. Overall, AcPseF displays high activity over a broad pH and temperature range, which could make the enzyme an attractive target for future application in the enzymatic synthesis of carbohydrates using multi-enzyme systems.

Employing conditions within the optimum activity range for AcPseF (pH 7.4, 37°C, 5 mM  $MgCl_2$ ), the apparent Michaelis-Menten kinetic parameters were then determined for Pse5Ac7Ac **1** and CTP. A discontinuous kinetic assay was performed where CMP-Pse5Ac7Ac **3** product formation was quantified over a 10 min time course, following separation of the reaction products by high-performance liquid chromatography (HPLC) and monitoring the absorbance at 270 nm. First, enzyme titration curves were prepared by incubating a fixed concentration of substrate with increasing concentrations of AcPseF, to estimate the concentration of enzyme required to observe the reaction during initial velocity after 10 min (Figure S8). Kinetic parameters were then measured under pseudo first order conditions, in which one substrate was held at a constant



**Figure 2. Substrate specificity and biochemical characterization of AcPseF**

- (A) Activity of AcPseF toward various sialic acids.  
 (B) Temperature-dependent activity profile.  
 (C) pH-dependent activity profile.  
 (D) Divalent cation screen. Error bars represent the SEM of three replicates.

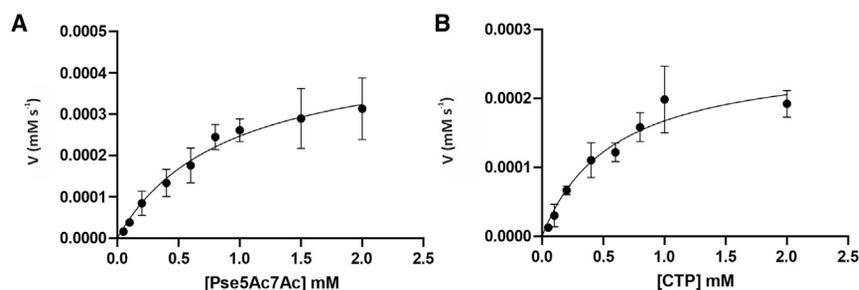
concentration while the other substrate concentration was varied (Figure 3; Table 1). At constant CTP, the apparent kinetic parameters for Pse5Ac7Ac **1** were  $k_{\text{cat}} = 4.308 \text{ s}^{-1}$ ,  $K_M = 0.895 \text{ mM}$ , and  $k_{\text{cat}}/K_M = 5.210 \text{ s}^{-1} \text{ mM}^{-1}$ . At variable CTP, the Michaelis-Menten parameters at constant concentration of Pse5Ac7Ac **1** were  $k_{\text{cat}} = 4.891 \text{ s}^{-1}$ ,  $K_M = 0.586 \text{ mM}$ , and  $k_{\text{cat}}/K_M = 8.350 \text{ s}^{-1} \text{ mM}^{-1}$ . Overall, the AcPseF kinetic parameters for Pse5Ac7Ac **1** are strikingly similar to those previously reported for AbPseF ( $k_{\text{cat}} = 3.690 \text{ s}^{-1}$ ,  $K_M = 0.783 \text{ mM}$ , and  $k_{\text{cat}}/K_M = 4.712 \text{ s}^{-1} \text{ mM}^{-1}$ ) a homolog from *A. baumannii* that shares ~61% amino acid sequence similarity (Figure S10),<sup>21</sup> but notably the overall catalytic efficiency is ~5- to 20-fold less efficient than bacterial CMP-Neu5Ac synthetases.<sup>29</sup>

### Overall structure of AcPseF

AcPseF was crystallized in the absence of any substrate and in complex with the product CMP-Pse5Ac7Ac **1**, with each structure determined to 1.9 Å and 1.5 Å, respectively (Table 2). Apo AcPseF crystallized with a single protomer in the asymmetric unit and the continuous polypeptide could be modeled from residues 13 to 246, though there were some regions for which the electron density was less well defined as a likely result of local

flexibility in the protein. The PseF homodimer was formed across the crystallographic 2-fold axis (Figures 4A and S11) consistent with previous findings that the enzyme forms a dimer in solution<sup>19</sup> and confirmed using PDB ePISA<sup>30</sup> analysis revealing a buried surface area of 3380 Å<sup>2</sup>. Each protomer consists of two domains; a nucleotide binding domain (His20 – Thr157, His195 – Arg248) and a dimerization domain (Arg158 – Tyr194), with dimerization mediated by domain swapping analogous to what has been observed with CMP-Neu5Ac synthetases.<sup>24</sup>

The CMP-Pse5Ac7Ac:PseF complex crystallized differently with a dimer in the asymmetric unit. Again, there was some flexibility evident in the model and Asp231 in chain A and residues 227–229 in chain B were omitted from the final model due to a lack of electron density. These regions contain the DXD motif that has been implicated in metal binding in related enzymes (discussed in more detail later). This structure confirmed that upon binding of CMP-Pse5Ac7Ac **3**, PseF adopts a more “closed” conformation around the ligand as revealed by comparisons against the apo structure (Figures 4A and S12). Subtle conformational changes occur in the enzyme, particularly in flexible regions consisting of residues Arg29 – Arg37, Arg89 – Thr100, and Ile214 – Thr219 contained within the nucleotide



**Figure 3. Michaelis-Menten plots of reaction rates measured for AcPseF**

Plots with varying concentrations of Pse5Ac7Ac **1** (A) and CTP (B), respectively, are shown. Error bars represent SEM of triplicate (3 independent experiments,  $n = 3$ ) or quadruplet (2 independent experiments with 2 technical replicates,  $n = 4$ ) samples.

**Table 1. Michaelis-Menten kinetic parameters for AcPseF**

Variable substrate	$K_M$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_M$ ( $s^{-1} mM^{-1}$ )
Pse5Ac7Ac	$0.895 \pm 0.215$	$4.308 \pm 0.550$	5.210
CTP	$0.586 \pm 0.126$	$4.891 \pm 0.443$	8.350

binding domain, and Ile163 – Glu191 encompassing the dimerization domain. This also revealed an asymmetry between the two protomers within our CMP-Pse5Ac7Ac:PseF complex structure, with chains A and B adopting slightly different conformations. This is most obvious by superposing the chains on top of each other (Figure 4B, root-mean-square deviation [RMSD] of chain A vs. chain B = 0.85 Å), with the major difference between the chains located in a flexible region harboring the conserved Lys33 (discussed in more detail with reference to the active site later).

The open vs. closed conformations observed in our structures echo similar movements that have been observed in NmCSS,<sup>24</sup> but with some differences evident as well. In NmCSS, structures of the Apoenzyme and complexes with CTP and the CMP-Neu5Ac product all adopted the open conformation whereas our product complex appears to be more closed. This is most obvious when we superpose the individual chains of our CMP-Pse5Ac7Ac:PseF complex onto chain A of the Apo and CMP-Neu5Ac complex structures of NmCSS where we can observe that the nucleotide domains superpose relatively well, but the

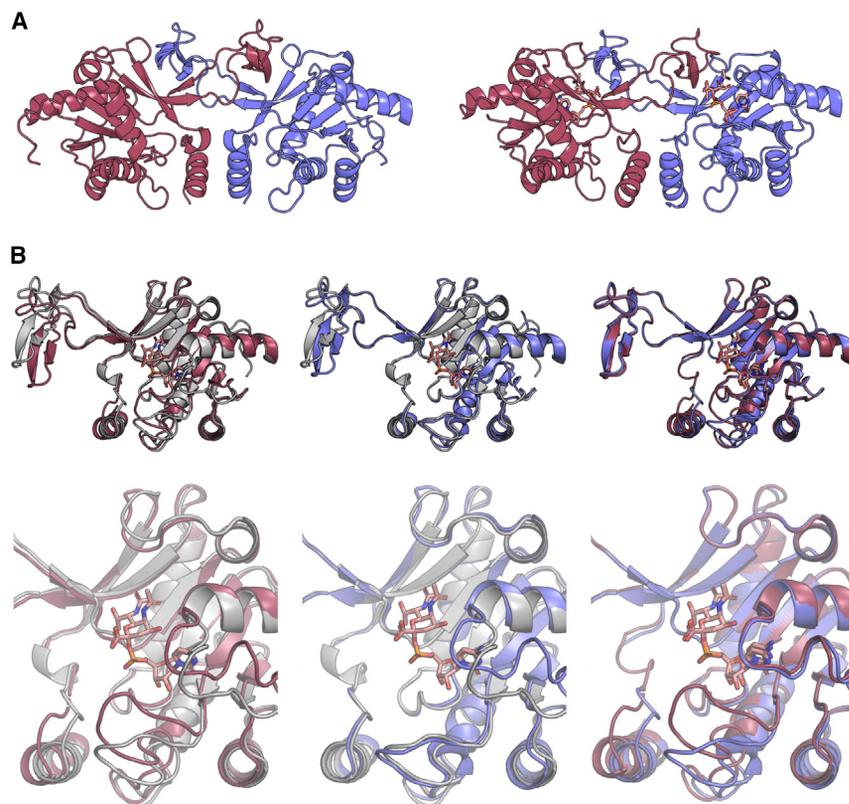
positioning of the loop from the opposing dimer (not used in the superposition) is significantly different (Figure S13A). In NmCSS, the only time a closed state could be observed was when the protein was co-crystallized with CMP (resulting from the hydrolysis of CTP) and an inhibitor Neu5Ac2en.<sup>24</sup> Superpositions of the individual chains of our CMP-Pse5Ac7Ac:PseF complex with chain A of the CMP-Neu5Ac2en:NmCSS complex reveals the closest overall structural match with the loop reaching over from the opposing dimer now matching more closely between the structures even though it was not included in the superposition (Figure S13B). Our Pse5Ac7Ac ligand has thus induced a closed conformation that the product did not appear to do in NmCSS. This analysis highlights that there are significant dynamic movements of these proteins during catalysis, and those observed in other family members are likely recapitulated in AcPseF. The conformations of chain A and B in our structure also appear to be subtly different, particularly near the ribose and phosphate moieties of the product that we will discuss in more detail in the following. Given this analysis we consider that the distinctions we observe between chains are not crystal

**Table 2. Data collection and refinement statistics for Apo-PseF and CMP-Neu5Ac7ac-PseF**

	Apo-AcPseF	CMP-Neu5Ac7Ac:PseF
<b>Data collection</b>		
Space group	P6 <sub>5</sub> 22	P2 <sub>1</sub>
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	109.5, 109.5, 87.1	42.6, 79.0, 74.0
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 120.0	90.0, 92.5, 90.0
Resolution (Å)	64.22–1.90 (1.94–1.90) <sup>a</sup>	42.57–1.50 (1.53–1.50) <sup>a</sup>
$R_{merge}$	0.22 (4.415) <sup>a</sup>	0.044 (0.120) <sup>a</sup>
$R_{pim}$	0.035 (0.683) <sup>a</sup>	0.028 (0.076) <sup>a</sup>
<i>CC</i> (1/2)	0.999 (0.917) <sup>a</sup>	0.997 (0.975) <sup>a</sup>
Mean( <i>I</i> / $\sigma$ )	14.8 (1.5) <sup>a</sup>	15.0 (6.2) <sup>b</sup> <sup>a</sup>
Completeness (%)	100 (100) <sup>a</sup>	99.7 (98.4) <sup>a</sup>
Multiplicity	40.0 (41.8) <sup>a</sup>	3.5 (3.5) <sup>a</sup>
<b>Refinement</b>		
Resolution (Å)	64.22–1.90	42.57–1.50
No. reflections (all/free)	24803/1187	78042/4148
$R_{work}/R_{free}$	0.192/0.241	0.167/0.185
<i>B</i> -factors (Å <sup>2</sup> )		
Protein	43.9	16.8
Ligand	N/A	12.5
Water	43.2	27.2
<b>Root-mean-square deviations</b>		
Bond lengths (Å)	0.0165	0.0107
Bond angles (°)	2.174	1.729
PDB ID	9FTB	9FTC

<sup>a</sup>Values in parentheses are for highest-resolution shell.

<sup>b</sup>These data diffracted beyond the limits of the detector.



**Figure 4. The crystal structures of AcPseF and CMP-Pse5Ac7ac:AcPseF complex**

(A) For the Apo structure (left), the dimer around the 2-fold symmetry axis is shown with each protomer colored red and lilac. For the CMP-Pse5Ac7ac:AcPseF complex (right), a dimer was modeled in the asymmetric unit that is shown with chains A and B colored in red and lilac, respectively. The ligand is shown as sticks colored by element with carbons in peach.

(B) Superposition of Apo PseF (gray) with CMP-Pse5Ac7ac:PseF chain A (left – red, RMSD = 0.967 Å; over 1,348 atoms) and chain B (middle – lilac, RMSD = 0.824 Å; over 1,418 atoms). Alignment of CMP-Pse5Ac7ac bound AcPseF chain A vs. chain B (right – RMSD = 0.435 Å; over 1,354 atoms). The image below each respective alignment shows a close-up view on the nucleotide binding domain, where differences are evident in flexible loop regions surrounding the active site.

artifacts, but likely represent snapshots along the conformational landscape that the protein adopts through catalysis that need to be considered if using rational design to make inhibitors of these enzymes.

#### AcPseF active site architecture

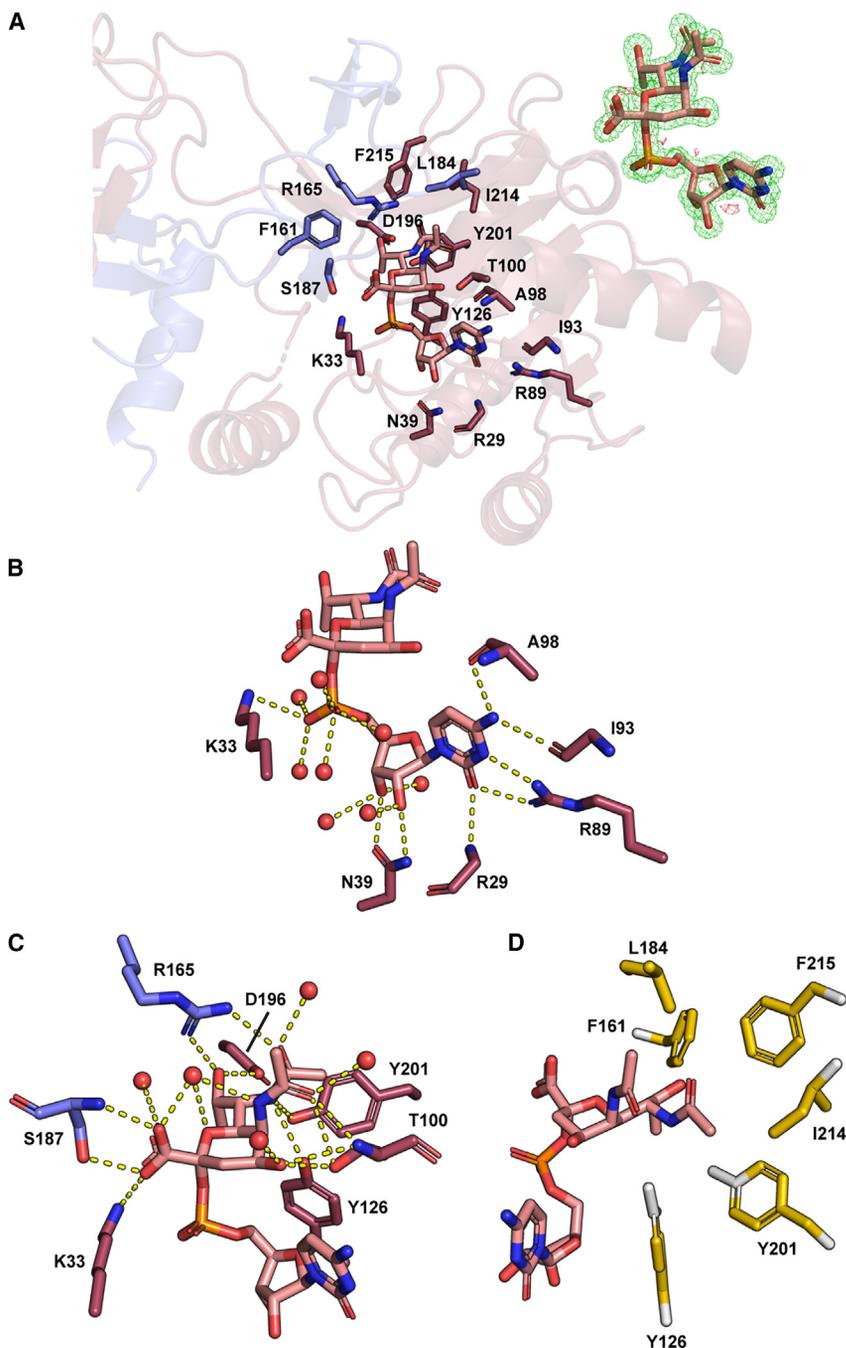
Close examination of the residues surrounding the CMP-Pse5Ac7Ac product shows that the ligand sits in a large pocket making both direct and water-mediated interactions with the protein (Figures 5A and S14A). The asymmetry in the overall structure of each protomer is also reflected in the interactions observed between protein and ligand in each chain. Considering the nucleotide portion of the product first there are relatively few differences between the two active sites in our structures when it comes to interactions with the cytosine ring. In chain A (Figure 5B) four hydrogen bonds (H-bonds) were observed between active site residues and the cytosine moiety of the ligand; one between Ala98 backbone carbonyl oxygen and the cytosine 4-NH<sub>2</sub>; two between Arg89 side chain amines and the cytosine ring nitrogen, and keto substituents, respectively; and one between Arg29 main chain amine and the cytosine keto substituent. In chain B (Figure S14B) the H-bond between Arg29 and the cytosine keto substituent is absent, which is reflective of the slight difference in overall conformation of the chain. Indeed, the electron density surrounding the cytosine ring in chain B showed significantly more disorder in this region likely reflecting the fact that this chain is more open and thus the ligand has more freedom to move in the crystal (Figure S14A).

Examination of the ribose ring and phosphate of the nucleotides revealed more obvious differences in active site architec-

ture between protomers. In chain A (Figure 5C) the Asn39 side chain (ND2 and OD1) H-bonds to the ribose 2'-OH and 3'-OH groups, while these H-bonds are absent in chain B (Figure S14C) and instead a network of water molecules form compensatory H-bonds. Additionally, the positively charged Lys33 side chain amino group ion-pairs with the negatively charged phosphate in chain A, while in chain B the density for the Lys33 side chain was absent and the main chain was more distant from the ligand.

Finally, the Pse5Ac7Ac moiety is located in a largely hydrophobic pocket formed by residues Tyr126, Phe161, Leu184, Tyr201, Ile214, and Phe215 that surrounds the non-polar C5-C9 regions of the sugar in both chains (Figures 5D and S14D). This is not dissimilar to the CMP-Neu5Ac bound NmCSS structure in which hydrophobic residues surround the C5 N-acetyl group of the sugar,<sup>24</sup> with mutational analysis demonstrating they were crucial for NmCSS activity.<sup>31</sup> There are additionally some direct H-bonds between protein and ligand with differences evident once more between protomers. In the chain A active site, direct H-bonds can be observed between the carbohydrate and Lys33, Thr100, Tyr126, Asp196, and Tyr201, while interactions are also formed with the loop that reaches over from chain B making H-bonds with Arg165, Ser187, and the main chain amide of Arg188. In chain B, the hydrophobic pocket is largely identical but H-bonds to Lys33 and Thr100 have been lost due to the difference in conformation of the chain. Interestingly, the H-bonds to the residues that reach over from the other protomer are still formed in this active site.

Following comparison against the AcPseF structure in the absence of substrate, it is evident that the nucleotide binding site is prepared for interacting with the cytosine moiety with only small movements evident here upon closure. The conformational changes in the protein are thus mainly driven by polar interactions with the phosphoribose moiety and Pse5Ac7Ac portions of the substrate. The slightly different conformations of each



**Figure 5. Active site view of CMP-Pse5Ac7Ac bound in chain A of AcPseF**

(A) The residues that bind to the product via either sidechain or main chain interactions are shown as sticks. Chain A and chain B are colored in red and lilac, respectively. The inset shows the  $F_o - F_c$  omit map for the CMP-Pse5Ac7Ac ligand contoured at  $4.0\sigma$ , with positive density in green and negative density in red.

(B) Interactions with the nucleotide portion of the ligand in chain A, with water molecules shown as red spheres and hydrogen bonds as yellow dashes.

(C) Interactions with the Pse5Ac7Ac portion of the ligand in the chain A active site.

(D) The hydrophobic pocket surrounding Pse5Ac7Ac where all carbon atoms not bound to nitrogen and oxygen atoms are highlighted in yellow.<sup>46</sup> See Figure S14D for equivalent views for the active site in chain B.

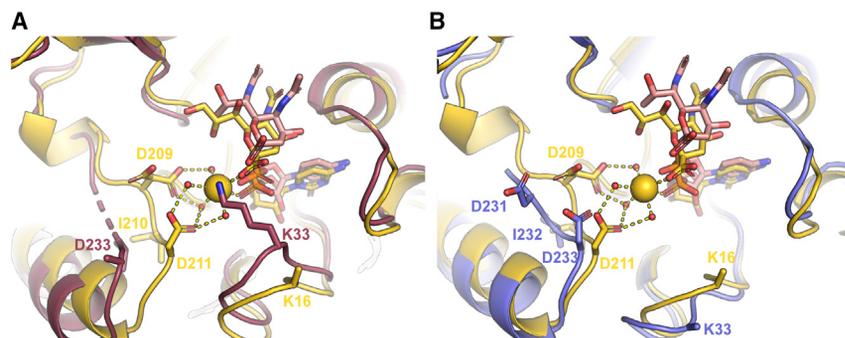
chain appear to represent a more open (chain B) and closed (chain A) conformation as reflected by the smaller number of direct interactions between protein and ligand in the slightly more open state. These differences may thus reflect the motions that are required for product release following catalysis by this enzyme.

#### AcPseF metal dependency

Our biochemical analysis clearly demonstrated that AcPseF activity is metal-dependent with  $Mg^{2+}$  being the preferred metal cofactor (Figure 2D). Though we included  $Mg^{2+}$  in all of our sam-

ples, we did not see any evidence of metals being present in either of the Apo or Pse5Ac7Ac:PseF structures. Indeed this is a distinct difference to the structure of the CMP-Neu5Ac:NmCSS in which a single  $Ca^{2+}$  ion coordinated by Asp209 and Asp211 that constitute the DXD motif, bridged the negative charges provided by the phosphate group and the carboxylate group of Neu5Ac.<sup>24</sup> In chain A of our CMP-Pse5Ac7Ac:PseF structure (the more closed conformation), the Lys33 amino group takes up a position equivalent to the  $Ca^{2+}$  ion present in the CMP-Neu5Ac:NmCSS structure (Figure 6A). As a result, the DXD motif has moved away from the active site and become disordered due to its mobility in the crystal. Meanwhile, in chain B, there were only water molecules that hydrogen bonded to the phosphate and carboxylate groups of the product (Figure 6B). Lys33 in this chain has moved away from the active site and has also become rather disordered such that we were unable to model the sidechain of this residue into any electron density (Figure S16). This conformation is thus more open than that observed in chain A.

The binding of metals has been studied extensively in NmCSS and related ulosonic acid-synthetases with metals most commonly observed in structures containing CTP or CDP.<sup>22–26</sup> Indeed, in these structures the metals are coordinated by the DXD motif but crucially interact with the  $\beta$ - and  $\gamma$ -phosphates where they are thought to help position them appropriately for catalysis. Mutations in Asp209 and Asp211 constituting the DXD motif in NmCSS have previously been shown to reduce enzymatic activity intimating their importance for catalysis. This effect was partially rescued by increasing the amount of  $Mg^{2+}$  in the reaction that presumably allowed correct positioning of



**Figure 6. Metal binding site in AcPseF compared to NmCSS**

Comparison of (A) chain A (red) and (B) chain B (lilac) from our CMP-Pse5Ac7Ac structure against CMP-Neu5Ac:NmCSS (gold) around the NmCSS  $\text{Ca}^{2+}$  binding site. The position of the DXD motif (residues 209–211) from NmCSS is shown with residues represented by sticks with equivalent residues from PseF also shown where possible. The position of Lys33 and the equivalent sidechain from NmCSS is also shown as sticks. Coordination around the  $\text{Ca}^{2+}$  is shown as dashed lines with relevant water molecules shown that bridge to the aspartates of the DXD motif. Water molecules from our PseF structure have not been included for clarity.

substrates without the coordinating aspartic acid residues being present.<sup>31</sup> The prevailing thinking in the field is, therefore, that the metals are most important before and during catalysis and likely leave with the pyrophosphate group post reaction.<sup>24</sup> While we were unable to capture a precatalytic state of AcPseF with CTP bound, our results align well with this dogma. The conformations we observed with the DXD containing loop moved out and the distinct positioning of Lys33 in the two chains in the dimer may represent distinct post-catalytic states that the enzyme can adopt to assist with the leaving of the pyrophosphate and metal ions before further opening of the structure allows the release of the CMP-Pse5Ac7Ac product.

#### Features determining specificity to Pse5Ac7Ac

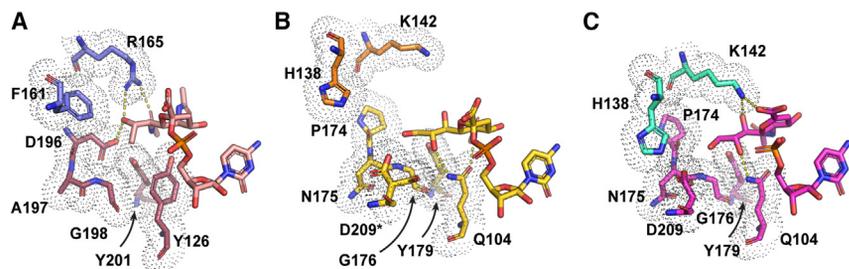
Our analysis highlights significant commonalities, as well as some interesting differences, between the molecular motions present in our structures as compared to those that have been observed in CSS and CKS enzymes previously.<sup>24–26</sup> Considering that our biochemical analysis revealed that AcPseF is specific for Pse5Ac7Ac and will not accept other nonulosonic acids like Neu5Ac as a substrate, we were also interested in how the enzyme might discriminate between sugars. Pse5Ac7Ac and Neu5Ac differ structurally in several ways, notably (1) the inverted stereochemistry of the C5 N-acetyl group, (2) the side chain Pse5Ac7Ac contains a C7 N-acetyl group, where Neu5Ac contains an OH group with the inverted stereochemistry, (3) inverted stereochemistry of the C8-OH groups, and (4) Pse5Ac7Ac side chain terminates with the C9-CH<sub>3</sub> while Neu5Ac has an extended side chain that terminates with a primary alcohol. The inverted stereochemistry of the C5 N-acetyl in Neu5Ac is unlikely to be responsible for the lack of AcPseF activity toward the substrate, as it could potentially be accommodated in space taken up by the C7 N-acetyl of Pse5Ac7Ac **1**, which is absent in Neu5Ac (Figures 7 and S15). If the stereochemistry of Pse5Ac7Ac at the C5 N-acetyl were inverted, this would result in a clash with the C7 N-acetyl. Instead AcPseF recognition of Pse5Ac7Ac appears to be primarily through the side chain of the sugar, with crucial interactions made between Arg165 and both the C7 N-acetyl and C8-OH groups, respectively (Figures 7A and S15A). These interactions could not be made with Neu5Ac that has an inverted C7-OH instead of the N-acetyl and inverted stereochemistry of the C8-OH. In contrast, NmCSS recognition of Neu5Ac **2** appears to be mainly through Lys142 that interacts with C7-OH and C5-N-acetyl (Figures 7B and 7C). Considering the extended C9 primary alcohol contain-

ing side chain of Neu5Ac, we would anticipate this to clash with AcPseF Phe161, suggesting that the active site is not large enough to accommodate the sugar and potentially explaining the lack of activity toward it in our biochemical studies (Figure 2A).

The active site surrounding the C5 and C7 N-acetyl groups (Figure 5A) is spacious which may shed light on why AcPseF is able to use N7-azide functionalized Pse5Ac7Ac substrates and, the related enzyme AbPseF, can use Pse5Ac7Ac analogs modified with bulky hydroxybutyryl groups at the N5 and N7 positions.<sup>20,21</sup> Nearly all of the Pse5Ac7Ac binding residues in the AcPseF active site are conserved in AbPseF (Figure S11), suggesting that AbPseF might have a similarly spacious active site region around the N-acetyl groups, a pocket that may represent fertile ground for the design of specific inhibitors of PseF.

#### Pseudaminic acid adopts the low energy *tg* conformation

Glycosyl donor reactivity is influenced by the carbohydrate side chain conformation: *gauche gauche* (*gg*), *gauche, trans* (*gt*), or *trans, gauche* (*tg*); and these can affect both the stereoselectivity<sup>32</sup> and rate of glycosidic bond formation and hydrolysis.<sup>33–35</sup> Crich and co-workers, who have extensively studied nonulosonic acid side chain conformation, established that in solution the Pse5Ac7Ac side-chain adopts the predicted lowest energy *tg* conformation around the exocyclic C6–C7 bond, consistent with experimental Pse5Ac7Ac <sup>3</sup>J<sub>H6,H7</sub> coupling constants of 10.5 Hz,<sup>36</sup> in contrast to the lowest energy *gg* conformation exclusively adopted by Neu5Ac in solution.<sup>32,35,37</sup> Through a survey of >300 crystal structures in complex with glycosides, Crich and co-workers also established that enzymes that proceed through oxocarbenium ion formation, such as glycosyl hydrolases and transferases, display a preference for a specific side chain conformation.<sup>38</sup> This is generally a conformation best able to additionally stabilize the transition state *en route* to oxocarbenium ion formation, which for pyranoses is often a higher energy rather than a lower energy conformation. Importantly, knowledge of these preferences has been successfully exploited in the design of more effective conformationally restricted enzyme inhibitors.<sup>39–41</sup> For nonulosonic acid processing enzymes the situation may be more complicated, however, with neuraminidases exclusively binding the *gg* conformer that is both the lowest energy in solution and best able to facilitate oxocarbenium ion formation through electrostatic interactions. This preference could therefore reinforce Crich's side chain



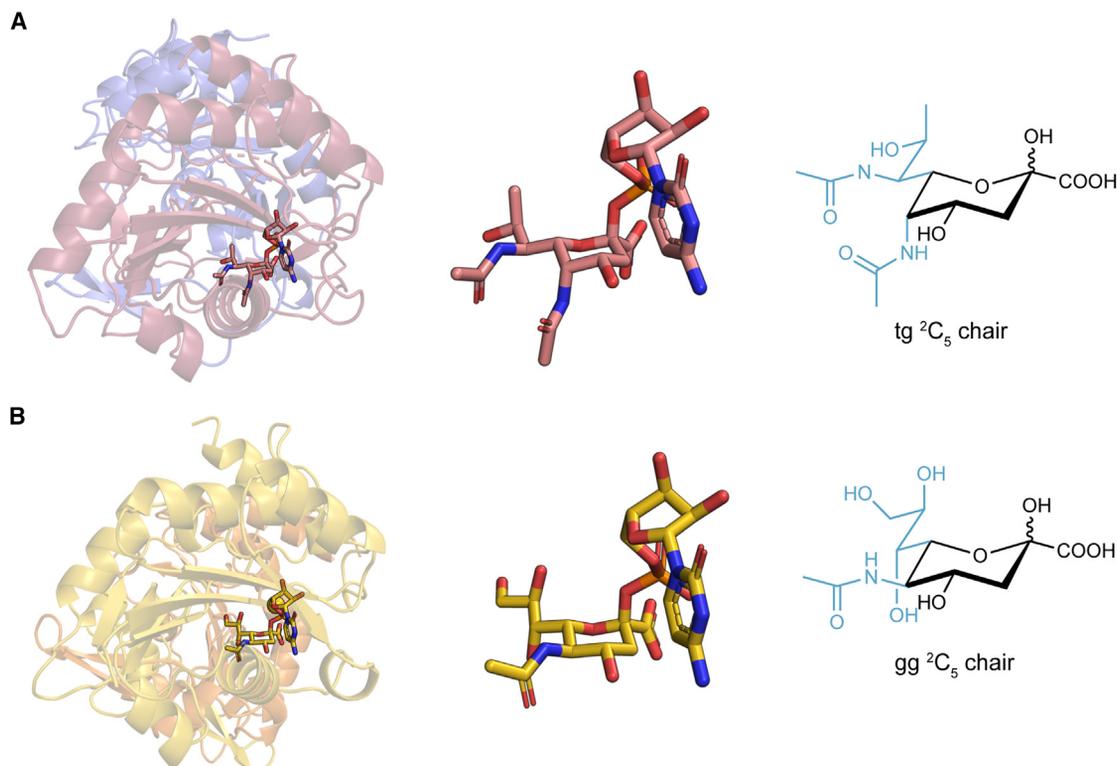
hydrogen bonds between the protein and ligand. Any waters and metal ions have been omitted for clarity. (A) Shows only the active site in chain A of our CMP-Pse5A7Ac:AcPseF structure, the equivalent comparison for chain B is shown in Figure S15.

transition state stabilization hypothesis, or simply reflect a preference to bind sialic acids in their most solution abundant, lowest energy conformation, as lectins have seemingly evolved to do.<sup>42</sup> Analogously, we observe that the side chain conformation of CMP-Pse5Ac7Ac, in our ligand bound AcPseF structure, adopts the lowest energy *tg* conformation (Figure 8A), while the CMP-Neu5Ac side chain in both the ligand bound NmCSS (Figure 8B, PDB: 6CKM and 6CKL)<sup>24</sup> and murine CSS structures (1QWJ) is also orientated in the lowest energy *gg* conformation,<sup>43</sup> and CMP-Kdo also adopts its predicted lowest energy *tg* side chain conformation in the active site of EcCKS (PDB: 3KED).<sup>44</sup> Although a small sample size, these observations might suggest that akin to lectins,<sup>42</sup> CMP-ulosonic acid synthetases have also evolved to bind the lowest energy conformations of their substrates in solution. As oxocarbenium ion formation is not involved in the proposed synthetase mechanism<sup>24</sup> it could be argued that

there is seemingly little benefit to adopting a higher energy side chain conformation to gain additional transition state stabilization of this oxocarbenium ion.

## DISCUSSION

Significant advances have been made in understanding the biosynthesis of Pse5Ac7Ac **1** and Pse processing enzymes,<sup>13,19,45</sup> yet molecular level insights into how the sugar is recognized by macromolecules have remained elusive. Having previously established *in vitro* activity of AcPseF,<sup>18,19</sup> a CMP-pseudaminic acid synthetase from *A. caviae*, we elected to structurally characterize the enzyme with a view to revealing novel insights into how a protein interacts with Pse5Ac7Ac **1**. We have demonstrated that AcPseF is a *bona fide* CMP-Pse5Ac7Ac synthetase, which displays metal-dependent



**Figure 8. Side chain conformations of Pse5Ac7Ac and Neu5Ac**

AcPseF shown in (A) and NmCSS (6CKM) shown in (B). Again, color schemes are retained from earlier figures.

activity over a broad pH and temperature range. Our structural data suggest that AcPseF undergoes dynamic movements upon CMP-Pse5Ac7Ac binding, strongly reminiscent of other CMP-ulosonic acid synthetases. Active site “closure” appears to be driven by polar interactions with the phosphoribose moiety and Pse5Ac7Ac portions of the substrate. The AcPseF active site is conveniently structured to recognize and accommodate the side chain of Pse5Ac7Ac, both through direct interactions with side-chain functional groups and by positioning of the molecule in a hydrophobic pocket. While demonstrating a clear propensity to discriminate Pse5Ac7Ac from other ulosonic acids, the considerable space surrounding the C5- and C7-N-acetyl groups suggests that the enzyme may tolerate Pse analogs derivatized at these positions. Finally, we show that AcPseF binds the CMP-Pse5Ac7Ac side chain in its lowest energy conformation, a trend that we observed in the structures of other enzymes of this class. Taken together, this work provides valuable insight into the recognition and binding of Pse5Ac7Ac by AcPseF, which could not have been gained purely through *in silico* study of the enzyme. We anticipate that our structure may inform the design of novel chemical tools or even antimicrobial inhibitors.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr Martin A. Fascione ([martin.fascione@york.ac.uk](mailto:martin.fascione@york.ac.uk)).

### Materials availability

The plasmid used for the expression of AcPseF in this study and key compounds including Pse5Ac7Ac and CMP-Pse5Ac7Ac can be requested from the [lead contact](#).

### Data and code availability

The PDB files reported in this study have been deposited in the Protein DataBank under accession codes 9FTB and 9FTC and are publicly available as of the date of publication. The accession codes are also listed in the [key resources table](#). The paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

We thank Dr Ed Bergstrom and The York Centre of Excellence in Mass Spectrometry. The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1). This work was supported by both The University of York and The University of Leeds, the BBSRC (White Rose DTP PhD awards to A.J.W. 2596660 and A.R.C. BB/T007222/1), the EPSRC (PhD award to E.K.P.F), and a Horizon Europe Guarantee Consolidator award (to M.A.F.) (selected by the ERC, funded by UKRI; EP/X023680/1). We thank Dr Jonathan Shaw (University of Sheffield) for the AcPseF\_pET28a plasmid.

## AUTHOR CONTRIBUTIONS

T.K. performed the biochemical assays, kinetic assays and chemo-enzymatic synthesis. T.K., A.R.C., G.R.H., and E.K.P.F. performed protein production and X-ray crystallography. N.E.H. synthesized 6-deoxy-AltdiNac. A.J.W. synthesized CMP-Pse5Ac7Ac. T.K., G.R.H., M.A.F., and A.R.C. wrote the manuscript. M.A.F., G.R.H., and G.H.T. supervised the project. All authors commented on the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
  - Bacterial strains
- [METHOD DETAILS](#)
  - General methods
  - Chemical and chemoenzymatic synthesis
  - High performance liquid chromatography coupled to mass spectrometry (LC-MS)
  - High performance liquid chromatography (HPLC)
  - PseI protein production
  - AcPseF expression and purification
  - Substrate specificity screening
  - Temperature, pH and metal-ion screening
  - Michaelis-Menten kinetic analysis
  - X-ray Crystallography
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.str.2024.09.017>.

Received: June 28, 2024

Revised: August 21, 2024

Accepted: September 16, 2024

Published: October 10, 2024

## REFERENCES

1. Thibault, P., Logan, S.M., Kelly, J.F., Brisson, J.-R., Ewing, C.P., Trust, T.J., and Guerry, P. (2001). Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **276**, 34862–34870.
2. Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N.H., and Prestegard, J.H. (2022). *Essentials of Glycobiology*.
3. Salah Ud-Din, A.I.M., and Roujeinikova, A. (2018). Flagellin glycosylation with pseudaminic acid in *Campylobacter* and *Helicobacter*: prospects for development of novel therapeutics. *Cell. Mol. Life Sci.* **75**, 1163–1178.
4. Schirm, M., Soo, E.C., Aubry, A.J., Austin, J., Thibault, P., and Logan, S.M. (2003). Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori*. *Mol. Microbiol.* **48**, 1579–1592.
5. Knirel, Y.A., Vinogradov, E.V., Shashkov, A.S., Dmitriev, B.A., Kochetkov, N.K., Stanislavsky, E.S., and Mashilova, G.M. (1987). Somatic antigens of *Pseudomonas aeruginosa*: The structure of the O-specific polysaccharide chain of the lipopolysaccharide from *P. aeruginosa* O13 (Lányi). *Eur. J. Biochem.* **163**, 627–637.
6. Kocharova, N., Shashkov, A., Dmitriev, B., and Kochetkov, N. (1986). Antigenic polysaccharides of bacteria. 18. Structure of O-specific polysaccharide chains of *Pseudomonas aeruginosa* O5 (Lányi) lipopolysaccharide. *Bioorg. Khim.* **12**, 1384–1390.
7. Sof'ya, N.S., Popova, A.V., Shashkov, A.S., Shneider, M.M., Mei, Z., Arbatsky, N.P., Liu, B., Miroshnikov, K.A., Volozhantsev, N.V., and Knirel, Y.A. (2015). Structure of a new pseudaminic acid-containing capsular polysaccharide of *Acinetobacter baumannii* LUH5550 having the KL42 capsule biosynthesis locus. *Carbohydr. Res.* **407**, 154–157.

8. Kenyon, J.J., Marzaioli, A.M., Hall, R.M., and De Castro, C. (2014). Structure of the K2 capsule associated with the KL2 gene cluster of *Acinetobacter baumannii*. *Glycobiology* *24*, 554–563.
9. WHO (2017). WHO publishes list of bacteria for which new antibiotics are urgently needed. <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>.
10. Parker, J.L., Day-Williams, M.J., Tomas, J.M., Stafford, G.P., and Shaw, J.G. (2012). Identification of a putative glycosyltransferase responsible for the transfer of pseudaminic acid onto the polar flagellin of *Aeromonas caviae* Sch3N. *Microbiologyopen* *1*, 149–160.
11. Stephenson, H.N., Mills, D.C., Jones, H., Milioris, E., Copland, A., Dorrell, N., Wren, B.W., Crocker, P.R., Escors, D., and Bajaj-Elliott, M. (2014). Pseudaminic acid on *Campylobacter jejuni* flagella modulates dendritic cell IL-10 expression via Siglec-10 receptor: a novel flagellin-host interaction. *J. Infect. Dis.* *210*, 1487–1498.
12. Zunk, M., and Kiefel, M.J. (2014). The occurrence and biological significance of the  $\alpha$ -keto-sugars pseudaminic acid and legionaminic acid within pathogenic bacteria. *RSC Adv.* *4*, 3413–3421.
13. Schoenhofen, I.C., McNally, D.J., Brisson, J.-R., and Logan, S.M. (2006). Elucidation of the CMP-pseudaminic acid pathway in *Helicobacter pylori*: synthesis from UDP-N-acetylglucosamine by a single enzymatic reaction. *Glycobiology* *16*, 8C–14C.
14. Li, Z., Hwang, S., Ericson, J., Bowler, K., and Bar-Peled, M. (2015). 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-I-AltNAc for CMP-Pseudaminic Acid Synthesis in *Bacillus thuringiensis*. *J. Biol. Chem.* *290*, 691–704.
15. Guerry, P., Ewing, C.P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006). Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol. Microbiol.* *60*, 299–311.
16. Gryllos, I., Shaw, J.G., Gavin, R., Merino, S., and Tomás, J.M. (2001). Role of flm locus in mesophilic *Aeromonas* species adherence. *Infect. Immun.* *69*, 65–74.
17. 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., Tomás, J.M., et al. (2009). An *Aeromonas caviae* genomic island is required for both O-antigen lipopolysaccharide biosynthesis and flagellin glycosylation. *J. Bacteriol.* *191*, 2851–2863.
18. Flack, E.K.P., Chidwick, H.S., Guchhait, G., Keenan, T., Budhadev, D., Huang, K., Both, P., Mas Pons, J., Ledru, H., Rui, S., et al. (2020). Biocatalytic Transfer of Pseudaminic Acid (Pse5Ac7Ac) Using Promiscuous Sialyltransferases in a Chemoenzymatic Approach to Pse5Ac7Ac-Containing Glycosides. *ACS Catal.* *10*, 9986–9993.
19. Chidwick, H.S., Flack, E.K.P., Keenan, T., Walton, J., Thomas, G.H., and Fascione, M.A. (2021). Reconstitution and optimisation of the biosynthesis of bacterial sugar pseudaminic acid (Pse5Ac7Ac) enables preparative enzymatic synthesis of CMP-Pse5Ac7Ac. *Sci. Rep.* *11*, 4756.
20. Keenan, T., Chidwick, H.S., Best, M., Flack, E.K.P., Yates, N.D.J., Hatton, N.E., Warnes, M.E., and Fascione, M.A. (2024). Co-factor prosthesis facilitates biosynthesis of azido-pseudaminic acid probes for use as glycosyltransferase reporters. *Chem. Commun.* *60*, 1428–1431.
21. Guo, X., Cheung, Y.C., Li, C., Liu, H., Li, P., Chen, S., and Li, X. (2024). Investigation on the substrate specificity and N-substitution tolerance of PseF in catalytic transformation of pseudaminic acids to CMP-Pse derivatives. *Chem. Sci.* 5950–5956.
22. Bose, S., Purkait, D., Joseph, D., Nayak, V., and Subramanian, R. (2019). Structural and functional characterization of CMP-N-acetylneuraminase synthetase from *Vibrio cholerae*. *Acta Crystallogr. D Struct. Biol.* *75*, 564–577.
23. Mosimann, S.C., Gilbert, M., Dombrowski, D., To, R., Wakarchuk, W., and Strynadka, N.C. (2001). Structure of a sialic acid-activating synthetase, CMP-acylneuraminase synthetase in the presence and absence of CDP. *J. Biol. Chem.* *276*, 8190–8196.
24. Matthews, M.M., McArthur, J.B., Li, Y., Yu, H., Chen, X., and Fisher, A.J. (2020). Catalytic cycle of *Neisseria meningitidis* CMP-Sialic acid synthetase illustrated by high-resolution protein crystallography. *Biochemistry* *59*, 3157–3168.
25. Jelakovic, S., and Schulz, G.E. (2001). The structure of CMP: 2-keto-3-deoxy-manno-octonic acid synthetase and of its complexes with substrates and substrate analogs. *J. Mol. Biol.* *312*, 143–155.
26. Jelakovic, S., and Schulz, G.E. (2002). Catalytic mechanism of CMP: 2-keto-3-deoxy-manno-octonic acid synthetase as derived from complexes with reaction educt and product. *Biochemistry* *41*, 1174–1181.
27. Heyes, D.J., Levy, C., Lafite, P., Roberts, I.S., Goldrick, M., Stachulski, A.V., Rossington, S.B., Stanford, D., Rigby, S.E.J., Scrutton, N.S., and Leys, D. (2009). Structure-based mechanism of CMP-2-keto-3-deoxy-manno-octulonic acid synthetase. *J. Biol. Chem.* *284*, 35514–35523.
28. Wahid, S.U.H. (2017). Structural and functional characterization of the *Helicobacter pylori* cytidine 5'-monophosphate-pseudaminic acid synthase PseF: molecular insight into substrate recognition and catalysis mechanism. *Adv. Appl. Bioinform. Chem.* *10*, 79–88.
29. Li, Y., Yu, H., Cao, H., Muthana, S., and Chen, X. (2012). *Pasteurella multocida* CMP-sialic acid synthetase and mutants of *Neisseria meningitidis* CMP-sialic acid synthetase with improved substrate promiscuity. *Appl. Microbiol. Biotechnol.* *93*, 2411–2423.
30. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* *372*, 774–797.
31. Horsfall, L.E., Nelson, A., and Berry, A. (2010). Identification and characterization of important residues in the catalytic mechanism of CMP-Neu5Ac synthetase from *Neisseria meningitidis*. *FEBS J.* *277*, 2779–2790.
32. Dhakal, B., and Crich, D. (2018). Synthesis and stereocontrolled equatorially selective glycosylation reactions of a pseudaminic acid donor: Importance of the side-chain conformation and regioselective reduction of azide protecting groups. *J. Am. Chem. Soc.* *140*, 15008–15015.
33. Jensen, H.H., Nordström, L.U., and Bols, M. (2004). The disarming effect of the 4, 6-acetal group on glycoside reactivity: torsional or electronic? *J. Am. Chem. Soc.* *126*, 9205–9213.
34. Moumé-Pymbock, M., Furukawa, T., Mondal, S., and Crich, D. (2013). Probing the Influence of a 4, 6-O-Acetal on the Reactivity of Galactopyranosyl Donors: Verification of the Disarming Influence of the trans-gauche Conformation of C5–C6 Bonds. *J. Am. Chem. Soc.* *135*, 14249–14255.
35. Siyalapitiya Arachchige, S., and Crich, D. (2022). Side chain conformation and its influence on glycosylation selectivity in hexo- and higher carbon furanosides. *J. Org. Chem.* *87*, 316–339.
36. Tsvetkov, Y.E., Shashkov, A.S., Knirel, Y.A., and Zähringer, U. (2001). Synthesis and NMR spectroscopy of nine stereoisomeric 5, 7-diacetamido-3, 5, 7, 9-tetra-deoxyxynon-2-ulosonic acids. *Carbohydr. Res.* *335*, 221–243.
37. Kancharla, P.K., and Crich, D. (2013). Influence of side chain conformation and configuration on glycosyl donor reactivity and selectivity as illustrated by sialic acid donors epimeric at the 7-position. *J. Am. Chem. Soc.* *135*, 18999–19007.
38. Quirke, J.C.K., and Crich, D. (2021). Side chain conformation restriction in the catalysis of glycosidic bond formation by Leloir glycosyltransferases, glycoside phosphorylases, and transglycosidases. *ACS Catal.* *11*, 5069–5078.
39. Tseng, P.S., Ande, C., Moremen, K.W., and Crich, D. (2023). Influence of side chain conformation on the activity of glycosidase inhibitors. *Angew. Chem.* *62*, e202217809.
40. Zechel, D.L., Boraston, A.B., Gloster, T., Boraston, C.M., Macdonald, J.M., Tilbrook, D.M.G., Stick, R.V., and Davies, G.J. (2003). Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to  $\beta$ -glucosidases. *J. Am. Chem. Soc.* *125*, 14313–14323.

41. Gloster, T.M., Madsen, R., and Davies, G.J. (2006). Dissection of Conformationally Restricted Inhibitors Binding to a  $\beta$ -Glucosidase. *Chembiochem* 7, 738–742.
42. McMahon, C.M., Isabella, C.R., Windsor, I.W., Kosma, P., Raines, R.T., and Kiessling, L.L. (2020). Stereoelectronic effects impact glycan recognition. *J. Am. Chem. Soc.* 142, 2386–2395.
43. Krapp, S., Münster-Kühnel, A.K., Kaiser, J.T., Huber, R., Tiralongo, J., Gerardy-Schahn, R., and Jacob, U. (2003). The crystal structure of murine CMP-5-N-acetylneuraminic acid synthetase. *J. Mol. Biol.* 334, 625–637.
44. Ngoje, P., and Crich, D. (2020). Stereocontrolled synthesis of the equatorial glycosides of 3-deoxy-D-manno-oct-2-ulosonic acid: Role of side chain conformation. *J. Am. Chem. Soc.* 142, 7760–7764.
45. Walklett, A.J., Flack, E.K.P., Chidwick, H.S., Hatton, N.E., Keenan, T., Budhadev, D., Walton, J., Thomas, G.H., and Fascione, M.A. (2024). The Retaining Pse5Ac7Ac Pseudaminyltransferase KpsS1 Defines a Previously Unreported glycosyltransferase family (GT118). *Angew. Chem.* 63, e202318523.
46. Hagemans, D., van Belzen, I.A.E.M., Morán Luengo, T., and Rüdiger, S.G.D. (2015). A script to highlight hydrophobicity and charge on protein surfaces. *Front. Mol. Biosci.* 2, 56.
47. Waterman, D.G., Winter, G., Gildea, R.J., Parkhurst, J.M., Brewster, A.S., Sauter, N.K., and Evans, G. (2016). Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. D Struct. Biol.* 72, 558–575.
48. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* 69, 1204–1214.
49. Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., et al. (2018). CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Struct. Biol.* 74, 68–84.
50. Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* 30, 1022–1025.
51. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444.
52. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
53. Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67, 355–367.
54. Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1002–1011.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Chemicals, peptides, and recombinant proteins</b>		
AcPseF	This paper	GenBank: WP_139737850.1
PseI	This paper	N/A
6-deoxy-AltdiNAc <b>S8</b>	This paper	N/A
Pse5Ac7Ac <b>1</b>	This paper	N/A
CMP-Pse5Ac7Ac <b>3</b>	This paper	N/A
N-Acetylneuraminic acid <b>2</b>	Carbosynth	Cat# MA00746
N-Glycolylneuraminic acid <b>4</b>	Carbosynth	Cat# MG05324
3-Deoxy-D-manno-2-octulosonic acid ammonium <b>5</b>	Carbosynth	Cat# MD04654
Cytidine triphosphate disodium salt	Merck	Cat# C1506
<b>Deposited data</b>		
AcPseF (Apo structure)	This paper	PDB:9FTB
AcPseF (CMP-Pse5Ac7Ac complex)	This paper	PDB:9FTC
<b>Recombinant DNA</b>		
AcPseF_pET28a	Dr Jonathan Shaw, University of Sheffield. Unpublished.	N/A
<b>Software and algorithms</b>		
GraphPad Prism v. 9.5.1	GraphPad Software Inc.	<a href="https://www.graphpad.com/">https://www.graphpad.com/</a>
DIALS	Waterman et al., 2016. <sup>47</sup>	<a href="https://dials.github.io/">https://dials.github.io/</a>
CCP4 suite	Potterton et al., 2018. <sup>49</sup>	<a href="http://www.ccp4.ac.uk/">http://www.ccp4.ac.uk/</a>
Coot	Emsley et al., 2010. <sup>52</sup>	<a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a>
AIMLESS	Evans and Murshudov, 2013. <sup>48</sup>	<a href="https://www.ccp4.ac.uk/">https://www.ccp4.ac.uk/</a>
MOLREP	Vagin and Teplyakov, 1997. <sup>50</sup>	<a href="https://www.ccp4.ac.uk/">https://www.ccp4.ac.uk/</a>
BUCANEER	Cowtan, 2006. <sup>54</sup>	<a href="https://www.ccp4.ac.uk/">https://www.ccp4.ac.uk/</a>
REFMAC5	Murshudov et al., 2011. <sup>53</sup>	<a href="https://www.ccp4.ac.uk/">https://www.ccp4.ac.uk/</a>
AlphaFold	Varadi et al. 2022. <sup>51</sup>	<a href="https://alphafold.ebi.ac.uk/">https://alphafold.ebi.ac.uk/</a>
Chemdraw 22.2	Revvity Signals	<a href="https://revvitysignals.com/products/research/chemdraw">https://revvitysignals.com/products/research/chemdraw</a>

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

#### Bacterial strains

For protein expression, commercially available chemically competent *E. coli* BL21(DE3) cells (*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHI ΔEcoRI-B int::(lac)::PlacUV5::T7 gene1 i21 Δnin5*) were used (NEB, C2527H). The bacterial cells were grown on LB agar at 37°C or in liquid LB medium at 37°C with shaking at 180 rpm, unless otherwise stated.

### METHOD DETAILS

#### General methods

All commercially available reagents were used as received and were supplied by Sigma-Aldrich, Fisher Scientific, VWR International, Carbosynth and TCI. NMR spectra were recorded on a Jeol ECX-400 (400 MHz) spectrometer. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard. Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, app = apparent, br = broad.

### Chemical and chemoenzymatic synthesis

The synthesis of 6-deoxy-AldiNac **S8** has been reported previously<sup>18</sup>; the methods and characterisation data have been reproduced herein and detailed in [Data S1](#). CMP-Pse5Ac7Ac **3** was prepared as previously described.<sup>19</sup> The chemoenzymatic synthesis of Pse5Ac7Ac **1** is detailed in [Data S1](#).

### High performance liquid chromatography coupled to mass spectrometry (LC-MS)

High Performance Liquid Chromatography-Electrospray Ionisation Mass Spectrometry (ESI-LC-MS) was accomplished using a Dionex UltiMate® 3000 LC system (ThermoScientific) equipped with an UltiMate® 3000 Diode Array Detector (probing 250–400 nm) in line with a Bruker HCTultra ETD II system (Bruker Daltonics), using Chromeleon® 6.80 SR12 software (ThermoScientific), Compass 1.3 for esquire HCT Build 581.3, 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) at The University York Centre of Excellence in Mass Spectrometry (CoEMS). All mass spectrometry was conducted in negative ion mode. Data analysis was performed using ESI Compass 1.3 DataAnalysis, Version 4.1 software (Bruker Daltonics). Samples were analysed on an Accucore HILIC column (2.6 µm particle size, 50 x 2.1 mm). Water, 0.1% formic acid by volume (solvent A), and acetonitrile, 0.1 % formic acid (solvent B) were used as the mobile phase at a flow rate of 0.3 ml/min at room temperature. A multi-step gradient of 15 min was programmed as follows: 95 % B for 1 min, followed by a linear gradient to 5 % B over 9.5 min, followed by 5 % B for an additional 0.5 min. Then a linear gradient to 95 % B over 1 min, followed by holding at 95 % B for 3 min.

### High performance liquid chromatography (HPLC)

Analytical HPLC was performed using an XBRIDGE HILIC column (3.5µm, 4.6 x 100 mm, Waters) on an Agilent 1260 Infinity II Quaternary System equipped with a 1260 Quaternary pump G7111B, 1260 Vial sampler, G7129A, multicolumn thermostat G7116A oven and a 1260 multiwavelength detector G7165A. Solvent A: H<sub>2</sub>O + 100 mM Ammonium acetate pH 5.35, solvent B: MeCN. 70% B ramping down to 65% B over 15 min, back up to 70% B over 0.5 min, hold at 70% B for 9.5 min. Flow rate: 0.25 mL/min, 40°C, 5 µL injection volume. UV absorbance monitored at 254 nm, 270 nm, 280 nm and 220 nm.

### PseI protein production

PseI (WP\_002870258.1) cloned into a pET15b expression vector (restriction enzymes NdeI and BamHI) was purchased from GenScript and introduced into chemically competent *E. coli* BL21(DE3) cells by heatshock. A starter culture was prepared by picking a single colony transformant into 50 mL of LB + ampicillin (100 µg/mL) and grown at 37°C with shaking (180 rpm, 16 h). The starter culture was diluted 1/100 into 1 L of fresh LB + ampicillin (100 µg/mL) and grown until an OD<sub>600</sub> of ~0.6 was reached. IPTG (final concentration of 0.5 mM) was added to induce recombinant protein expression and the cells were grown for 20 h at 37°C. Cells were harvested by centrifugation (6000 × g, 20 min, 4°C) and the pellet stored at -80°C. For protein isolation, cell pellets were thawed on ice and suspended in ice cold lysis buffer (50 mM HEPES buffer pH 7.5, 0.15 M NaCl, protease inhibitor, benzonase). Cells were lysed on ice by sonication (30s on/30s off for 12 min) and the lysate was centrifuged (20000 × g, 50 min, 4°C). The supernatant was purified using a HisTrap HP Ni<sup>2+</sup> affinity column, pre-equilibrated with binding buffer (50 mM HEPES buffer pH 7.5, 0.15 M NaCl, 20 mM Imidazole). Protein was eluted using elution buffer (50 mM HEPES buffer pH 7.5, 0.15 M NaCl, 500 mM Imidazole). Fractions containing PseI (expected Mw. = 38.6 kDa) were pooled and desalted using a HiPrep™ 26/10 Desalting column (GE Healthcare) in desalting buffer (50 mM HEPES buffer pH 7.5, 0.15 M NaCl).

### AcPseF expression and purification

A plasmid encoding PseF (AcPseF\_pET28a) from *Aeromonas caviae* (WP\_139737850.1)<sup>19</sup> was introduced into chemically competent *E. coli* BL21(DE3) cells by heat shock and selected on LB agar with kanamycin (50 µg/mL) at 37°C for 16 h. Starter cultures were prepared by picking single clones into LB with kanamycin (50 µg/mL) and grown at 37°C for 16 h with shaking (180 rpm). 10 mL of starter culture was added to 1L of LB with kanamycin (50 µg/mL) and grown to an OD<sub>600</sub> of 0.6 – 0.7 at 37°C with shaking (180 rpm). IPTG was added to a final concentration of 0.1 mM and the cultures were grown at 16°C for a further 20 h with shaking (180 rpm). Cells were harvested by centrifugation (6000 × g, 4°C, 20 min) and the pellet was stored at -80°C until use. For protein purification, cell pellets were thawed on ice and resuspended in lysis buffer (50 mM sodium phosphate buffer pH 7.4, 50 mM NaCl, 1 mM MgCl<sub>2</sub>, benzonase (0.4U/mL) and 1 x protease inhibitor). The cells were lysed by sonication on ice and the lysate was clarified by centrifugation (18 000 rpm, 4°C, 50 min). The lysate was loaded onto a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated with binding buffer (50 mM sodium phosphate buffer pH 7.4, 400 mM NaCl, 1 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol, 20 mM Imidazole). After washing the column with 20 column volumes (CV) of binding buffer, recombinant AcPseF was eluted using a 0 – 100% gradient of elution buffer (50 mM sodium phosphate buffer pH 7.4, 400 mM NaCl, 1 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol, 500 mM Imidazole). The protein was desalted and exchanged into desalting buffer (50 mM sodium phosphate buffer pH 7.4, 50 mM NaCl, 1 mM MgCl<sub>2</sub>) using a HiPrep™ 26/10 Desalting column (GE Healthcare). Further purification was carried out by using gel filtration chromatography (HiLoad 16/600 Superdex 200 pg) in desalting buffer. Purity was assessed by SDS-PAGE ([Figure S1](#)). AcPseF was then concentrated using a spin concentrator (MWCO – 10 kDa) to approximately 20 mg/mL.

### Substrate specificity screening

Reactions were assembled on a 30  $\mu\text{L}$  scale, containing 1 mM sugar (**1-2**, **4-5**), 5 mM CTP, 5 mM  $\text{MgCl}_2$ , 50 mM sodium phosphate buffer pH 7.5 and AcPseF (0.4 mg/mL or 4 mg/mL). Alongside the test reactions, no enzyme and no sugar control reactions were assembled. Reactions were incubated at 37°C for 30 min. Following incubation, reactions were diluted 1 in 3 in MeCN. Samples were centrifuged at 13 000x g for 2 min, and the supernatant was analysed by HILIC negative ESI-LC-MS.

### Temperature, pH and metal-ion screening

#### Preparation of a CMP-Pse5Ac7Ac **3** standard curve by LC-MS

A 10 mM stock of **3** (in  $\text{dH}_2\text{O}$ ) was diluted 1/10 in 60% v/v MeCN/ $\text{dH}_2\text{O}$  to prepare a 1 mM stock solution. A serial 2-fold dilution was then performed with the 1 mM stock of **3** in 60% v/v MeCN/ $\text{dH}_2\text{O}$ , to prepare a series of standards ranging from 1 mM to 3.9  $\mu\text{M}$ . (1, 0.5, 0.25, 0.125, 0.0625, 0.03125, 0.0156, 0.0078, 0.0039 mM). The standards were analysed by HILIC LC-MS and **3** was monitored by measuring the UV absorbance at 280 nm. **3** typically eluted from the column at 5.6 min. In the UV traces, the peak corresponding to **3** was integrated and the peak area was plotted as a function of concentration (mM) to produce a CMP-Pse5Ac7Ac standard curve (Figure S6).

#### Temperature-dependant activity screen

Reactions on a 20  $\mu\text{L}$  scale contained 1 mM Pse5Ac7Ac **1**, 5 mM CTP, 5 mM  $\text{MgCl}_2$ , 100 mM sodium phosphate buffer pH 7.5 and AcPseF (0.4 mg/mL). Reactions were assembled in triplicate in PCR tubes, containing all components except for AcPseF enzyme. Reactions were pre-incubated at the reaction temperature (4°C, 16°C, 22°C, 30°C, 37°C, 45°C, 55°C or 65°C) for 5 min, in a PCR thermocycler. Following the addition of AcPseF, the reactions were incubated for a further 20 mins in a PCR thermocycler, at the designated reaction temperature. 40  $\mu\text{L}$  of MeCN was added to stop the reactions. Samples were centrifuged at 13 000 x g for 2 min, and the supernatant was analysed by HILIC negative ESI-LC-MS. In the UV-Vis 280 nm traces for each reaction, the peak corresponding to CMP-Pse5Ac7Ac **3** was integrated and the concentration was determined from the standard curve (Figure S6).

#### pH-dependant activity profile

Reactions on a 20  $\mu\text{L}$  scale contained 1 mM Pse5Ac7Ac **1**, 5 mM CTP, 5 mM  $\text{MgCl}_2$ , 100 mM buffer (Glycine-HCl pH 2 Glycine-HCl pH 3, NaOAc pH 4, NaOAc pH 5, NaOAc pH 6, Sodium phosphate pH 7, Tris-HCl pH 8, Glycine-OH pH 9, Glycine-OH pH 10, Glycine-OH pH 11) and AcPseF (0.4 mg/mL). Reactions were assembled in triplicate in PCR tubes, containing all components except for AcPseF enzyme. Reactions were pre-incubated at 37°C, for 5 min in a PCR thermocycler. Following the addition of AcPseF, the reactions were incubated for a further 30 mins in a PCR thermocycler at 37°C. 40  $\mu\text{L}$  of MeCN was added to stop the reactions. Samples were centrifuged at 13 000 x g for 2 min, and the supernatant was analysed by HILIC negative ESI-LC-MS. In the UV-Vis 280 nm traces for each reaction, the peak corresponding to CMP-Pse5Ac7Ac **3** was integrated and the concentration was determined from the standard curve (Figure S6).

#### Metal dependant activity

Reactions on a 30  $\mu\text{L}$  scale contained 1 mM Pse5Ac7Ac **1**, 5 mM CTP, 5 mM divalent cation, 50 mM sodium phosphate buffer pH 7.5 and AcPseF (0.4 mg/mL). Reactions were assembled in PCR tubes, containing all components except for AcPseF enzyme. Reactions were pre-incubated at 37°C, for 5 min in a PCR thermocycler. Following the addition of AcPseF, the reactions were incubated for a further 30 mins at 37°C. Samples were diluted in 60  $\mu\text{L}$  of MeCN and centrifuged at 13 000 x g for 2 min, and the supernatant was analysed by HILIC negative ESI-LC-MS. In the UV-Vis 280 nm traces for each reaction, the peak corresponding to CMP-Pse5Ac7Ac **3** was integrated and the absolute concentration was determined from the standard curve (Figure S6).

### Michaelis-Menten kinetic analysis

#### Standard curve

A 10 mM stock of CMP-Pse5Ac7Ac **3** (in  $\text{dH}_2\text{O}$ ) was diluted 1/10 in 50% v/v MeCN/ $\text{dH}_2\text{O}$  to prepare a 1 mM stock solution. A serial 2-fold dilution was performed with the 1 mM CMP-Pse5Ac7Ac **3** stock in 50% v/v MeCN/ $\text{dH}_2\text{O}$ , to prepare a series of standards ranging from 1 to 0.00098 mM. (1, 0.5, 0.25, 0.125, 0.0625, 0.03125, 0.0156, 0.0078, 0.0039, 0.0019, 0.00098 mM). Standards were analysed by HPLC and **3** was monitored by measuring the absorbance at 270 nm. CMP-Pse5Ac7Ac **3** retention time = 10.7 min. A calibration curve was generated by plotting the peak area as a function of concentration (Figure S7).

#### Enzyme titration curves to determine optimal enzyme concentration for kinetic experiments

A two-fold serial dilution of AcPseF was prepared in 50 mM sodium phosphate pH 7.4, 50 mM NaCl buffer to afford enzyme stocks at the following concentrations: 0.136, 0.271, 0.543, 1.086, 2.172, 4.344, 8.688, 17.375 and 34.750  $\mu\text{M}$ . Reactions were assembled in 50 mM sodium phosphate buffer pH 7.4 containing 5 mM  $\text{MgCl}_2$  and either 1 mM Pse5Ac7Ac **1** and 2.5 mM CTP (a representative example of Pse5Ac7Ac kinetic conditions) or 4 mM Pse5Ac7Ac and 0.8 mM CTP (a representative example of CTP kinetic conditions) in a reaction volume of 24  $\mu\text{L}$ . Following pre-incubation of the reaction mixture at 37°C for 5 min, 6  $\mu\text{L}$  of AcPseF stock solution (final concentration 0.027 – 6.95  $\mu\text{M}$ ) was added and the reactions were incubated for 10 min at 37°C. Reactions were quenched with acetonitrile (30  $\mu\text{L}$ ), incubated on ice for 5 min and centrifuged (13 000 x g, 5 min, 4°C). The supernatant was analysed by HPLC monitoring the absorbance at 270 nm and conversion to product **3** was determined by using a calibration curve (Figure S7). Curves of CMP-Pse5Ac7Ac (mM) against enzyme concentration were plotted (Figure S8). Initial velocity (i.e region of linearity in the curve) was observed after 10 min at 0.027 – 0.207  $\mu\text{M}$  AcPseF for the reaction with 1 mM Pse5Ac7Ac and 2.5 mM CTP, and 0.027 – 0.108  $\mu\text{M}$

AcPseF in reactions with 4 mM Pse5Ac7Ac and 0.8 mM CTP. Therefore, the concentration of AcPseF for the kinetic experiments was set at 0.109  $\mu\text{M}$  for Pse5Ac7Ac **1** (where CTP is held at 2.5 mM) kinetics and 0.054  $\mu\text{M}$  for CTP (where Pse5Ac7Ac is held at 2.5 mM) kinetics.

### Michaelis-Menten kinetics experiments

Michaelis-Menten kinetic analyses were performed for Pse5Ac7Ac **1** and CTP under pseudo first-order conditions in which the concentration of one substrate was varied while the other was held constant, and vice versa. For the Pse5Ac7Ac **1** kinetics, reactions were assembled in 50 mM sodium phosphate buffer pH 7.4 containing 5 mM  $\text{MgCl}_2$ , 109 nM AcPseF, constant 2.5 mM CTP and varying concentrations of Pse5Ac7Ac (0.05 – 2 mM) in a final reaction volume of 100  $\mu\text{L}$ . For the CTP kinetics, reactions were assembled in 50 mM sodium phosphate buffer pH 7.4 containing 5 mM  $\text{MgCl}_2$ , 54.5 nM AcPseF, constant 4 mM Pse5Ac7Ac **1** and varying concentrations of CTP (0.05 – 2 mM) in a final reaction volume of 100  $\mu\text{L}$ . All reaction components minus AcPseF were assembled & pre-incubated at 37°C for 5 min. AcPseF was then added to the reaction, which was further incubated for 10 min. Samples (15  $\mu\text{L}$ ) were taken over the time course (2.5, 5, 7.5 & 10 min) and quenched with acetonitrile (15  $\mu\text{L}$ ). After incubating on ice for 5 min, samples were centrifuged (13 000  $\times$  g, 5 min, 4°C) and the supernatant analysed by HPLC. The concentration of CMP-Pse5Ac7Ac **3** was determined using a calibration curve (Figure S7) and plotted as a function of time to generate curves of initial velocity ( $v^0$ ). The apparent kinetic parameters,  $k_{\text{cat}}$ ,  $K_m$ , and  $k_{\text{cat}}/K_m$  were calculated using the Prism 9 software package (GraphPad Scientific Software).

### X-ray Crystallography

Initial crystallisation conditions were identified using commercially available screens from Hampton in a 96-well sitting drop screening format. The following provided crystals which led to the solving of the Apo structure. Screening hits were optimised in a 48-well sitting drop format and optimal crystals were obtained at 20°C by mixing 0.5  $\mu\text{L}$  of a protein solution (22 mg/mL in 50 mM sodium phosphate buffer pH 7.4, 50 mM NaCl, 1 mM  $\text{MgCl}_2$ ) with 0.5  $\mu\text{L}$  of reservoir solution (1% w/v tryptone, 0.05 M HEPES sodium pH 6.5 – 8.0, 8 – 16% w/v polyethylene glycol (PEG) 3350 and 0.001 M sodium azide) in plates containing 150  $\mu\text{L}$  of reservoir solution per well. Crystals were observed within 2 weeks. All crystals were fished via a cryoprotectant, which comprised of reservoir solution with 13% v/v glycerol (75  $\mu\text{L}$  of glycerol + 500  $\mu\text{L}$  reservoir solution) into liquid nitrogen. Data were collected from AcPseF Apo crystals at beamline IO3, Diamond Light Source (Didcot, UK). The data were processed using DIALS and scaled using AIMLESS<sup>47,48</sup> before subsequent processing was performed in the CCP4i2 suite of programs.<sup>49</sup> The AcPseF Apo data were phased using molecular replacement in MOLREP, using an AlphaFold model for PseF from *Campylobacter jejuni* (Q2M5Q2) as a search model.<sup>50,51</sup>

The CMP-Pse5Ac7Ac:PseF complex was prepared by mixing 20 mg/ml of purified AcPseF with a small volume of concentrated CMP-Pse5Ac7Ac **3** to give a final concentration of 10 mM. Both the protein and ligand were prepared in the same buffer - 50 mM sodium phosphate buffer pH 7.4, 50 mM NaCl, 1 mM  $\text{MgCl}_2$ . Screening hits were optimised in 24-well plates via 2  $\mu\text{L}$  hanging drops which were prepared above a 500  $\mu\text{L}$  reservoir solution. Crystals were obtained at 18°C by mixing equal volumes of protein and reservoir solution containing 0.2 M sodium nitrate, 0.1 M Bis Tris propane pH 6.5, 15 – 25% w/v PEG 3350. Resulting crystals were cryo-cooled in liquid nitrogen following a brief incubation in mother liquor containing 20% ethylene glycol. Diffraction data were subsequently collected at Diamond Light Source, beamline I04. The autoprocessed data from Xia2/XDS were then scaled using AIMLESS and all subsequent data processing performed in the CCP4i2 suite of programmes. The CMP-Pse5Ac7Ac:PseF structure was then solved using molecular replacement with PHASER, starting from a truncated model of our Apo AcPseF structure encompassing the nucleotide binding domain (residues 13 – 157 and 195 – 246). This model was extended using BUCANEER before further iterative cycles of manual model building and refinement which were performed in COOT and REFMAC5 respectively for both the Apo and CMP-Pse7Ac5Ac structures.<sup>52–54</sup> Data collection and refinement statistics are reported in Table 2.

### QUANTIFICATION AND STATISTICAL ANALYSIS

Data collection and refinement statistics for all structures described in this work are available in Table 2, as well as the specific PDB entries for each dataset. Kinetic experiments were carried out in triplicate (3 independent experiments,  $n=3$ ) or quadruplet (2 independent experiments with 2 technical replicates,  $n=4$ ) and data fitting and statistical analysis was performed using GraphPad Prism version 7.05. Error bars representing standard error of the mean (SEM) are shown on the Michaelis Menten kinetic curves.