

Cite this: *RSC Chem. Biol.*, 2026, 7, 352Received 16th October 2025,  
Accepted 4th December 2025

DOI: 10.1039/d5cb00270b

rsc.li/rsc-chembio

## Protein mannosylation in actinobacteria an enigmatic post-translational modification

Cameron B. King  and Warren W. Wakarchuk \*

Protein glycosylation is a very common post-translational modification seen in all branches of biology. The functional roles for protein glycosylation are many and varied, essential in eukaryotes but seemingly dispensable in bacteria. One group of bacteria where protein glycosylation has been looked at for at least 50 years are the actinobacteria, a large and diverse group of bacteria which include well know pathogens like *Mycobacteria tuberculosis*, *Corynebacterium diphtheriae*, and well know species important in biotechnology like *Streptomyces lividans* and *Corynebacterium glutamicum*. Actinobacterial protein glycosylation is a form of protein O-mannosylation which is found widely in eukaryotes from single celled yeast to complex multicellular organisms but is much less understood at the functional level. Very few direct roles for protein O-mannosylation have been described in the literature. This review examines newer findings from the actinobacterial research literature which with the help of glycoprotein models suggests how the glycans might play a role in actinobacterial growth and physiology.

### Introduction

Protein O-mannosylation (POM) is a form of O-linked glycosylation (Fig. 1) which is commonly found in eukaryotic organisms from single celled yeast through to humans where it serves an essential function (reviewed in ref. 1). According to the Carbohydrate Active Enzymes Data Base (CAZY.org), the key enzyme, protein O-mannosyltransferase from glycosyltransferase family GT-39, is also found in ~4000 species of bacteria, the vast majority of which are actinobacteria. In contrast to the eukaryotic systems, the actinobacterial system has been much less studied despite being so widespread in this diverse actinobacterial bacterial family.

The study of bacterial POM started at least 50 years ago with the discovery that a lipid-linked donor was required for glycoprotein synthesis in *Mycobacterium smegmatis*.<sup>2</sup> Since the early discovery of these O-mannosylated proteins, many glycoproteins from various *Mycobacterium tuberculosis* (MTB) samples have subsequently been reported. However, it was not until 1995 when the covalent linkage of mannose to an amino acid, was proven.<sup>3</sup> The delay in this identification was mainly because of the plethora of glycolipids, and peptidoglycan fragments that confounded the early analysis.

Since the seminal demonstration of covalent mannosylation, a few more manno-proteins have been characterized at the molecular level<sup>4–7</sup> but no one has yet published a concise enumeration of glycoproteins in any given species. There have

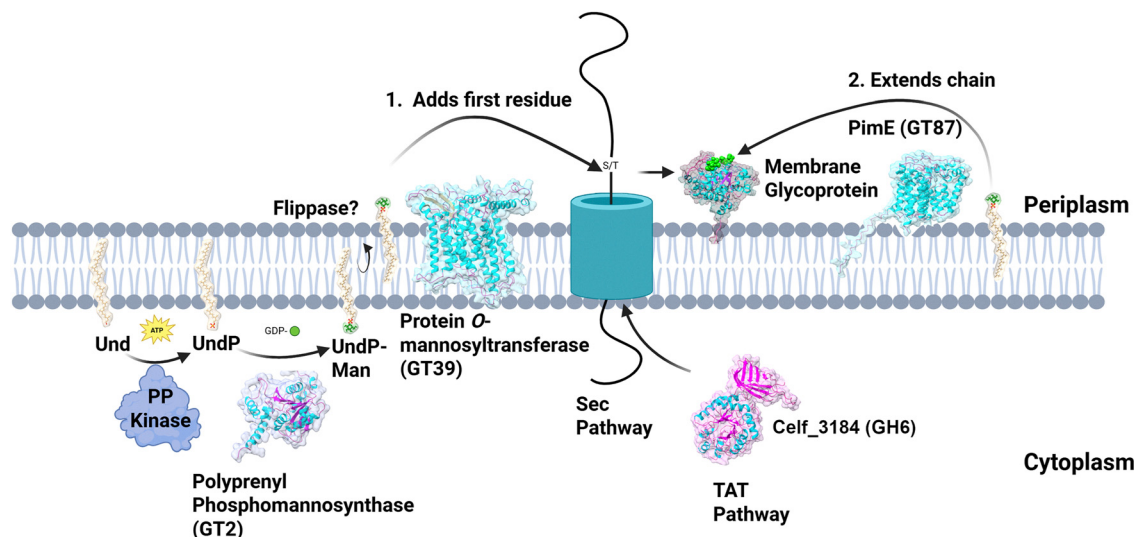
been a few surveys of mannoprotein content in *M. tuberculosis* (MTB),<sup>8–10</sup> *Streptomyces coelicolor*,<sup>11</sup> *Corynebacterium glutamicum*,<sup>12</sup> *Cellulomonas fimi*,<sup>13</sup> and one from *Mycobacterium smegmatis*.<sup>14</sup> Zheng<sup>15</sup> conducted a survey of *M. bovis*, Bacillus Calmette–Guérin and claimed to have identified more than 700 glycoproteins. However, compared to other surveys, the Zheng data seem at odds with the more typical  $\leq 100$  proteins. There are also some surveys from MTB that only “predicted” O-glycopeptides<sup>16</sup> or even N-glycans.<sup>17</sup> It is hard to reconcile the presence of N-glycans as they should not be present without the much-studied bacterial oligosaccharyltransferase from the GT-66 family, which is not found in actinobacterial species.<sup>18</sup> In those latter two surveys, there may be some O-mannosylated proteins, but without physical details it is hard to discern whether the results have been validated at the biochemical level. Despite the lack of complete data from any species, there are some common O-mannosylated proteins shared between them and some will be discussed below. Key information is missing from all the work thus far, including a molecular mechanism for a function that these glycans serve for the carrier proteins and how that contributes to the normal biology of the bacteria possessing the biochemical machinery for protein mannosylation.

### The protein mannosyl-transferase from glycosyltransferase family GT-39

Work to understand the biochemistry of the GT-39 protein mannosyltransferase started 25 years ago with fungal enzymes.

University of Alberta, Department of Biological Sciences, Edmonton Alberta, Canada T6G 2E9. E-mail: wakarchu@ualberta.ca





**Fig. 1** Predicted schematic of bacterial protein *O*-mannosylation pathway. We have used AlphaFold predicted structures of *Corynebacterium glutamicum* ATCC13032, *Mycobacterium smegmatis* and *Cellulomonas fimi* orthologues to illustrate proteins involved. The transmembrane GT-39 glycosyltransferase (From *C. glutamicum*, protein *O*-mannosyltransferase, UniprotID Q8NRZ6) catalyzes the transfer of a single  $\beta$ -D-mannose residue from a lipid donor, polyprenol phosphomannose (UndP-Man). Synthesis of the lipid donor occurs in the cytoplasm through successive reactions catalyzed by a polyprenol kinase on a C45–C75 polyprenol lipid (usually undecaprenol – Und) then a GT-2 family mannosyltransferase (shown here is the *C. glutamicum* polyprenol-phosphate  $\beta$ -D-mannosyltransferase, UniprotID Q8NQF7) on the resulting polyprenol phosphate (UndP). Once synthesized the lipid donor, UndP-Man, is flipped to face the periplasm by a hypothesized flippase enzyme where the mannose is transferred to a serine or threonine of a Sec or TAT secreted protein (from *Cellulomonas fimi*, Celf\_3184 UniprotID Q8NLT5) in the periplasm/at the membrane surface. Additionally in the periplasm, the GT-87 enzyme PimE (From *M. smegmatis* UniprotID A0R2K8) extends the chain of mannose residues beyond the first on mannoproteins also using UndP-Man as a donor.

This seminal work is reviewed in.<sup>19</sup> Subsequent work on the bacterial orthologues showed some similarities with the fungal enzymes (Table 1). Eukaryotes have another more recently described protein *O*-mannosyltransferase, TMTC, in GT-105,<sup>20</sup> and there are bacterial orthologues but apparently none reported in actinobacteria. The GT-39 enzymes are inverting glycosyltransferases that use a prenol-lipid-linked sugar donor. A direct demonstration of enzymatic activity of the MTB GT-39 protein mannosyltransferase (MtbPMT) was shown in 2002 using cell free assays.<sup>21</sup> A cryo-EM structure for the *Saccharomyces cerevisiae* PMT1/2 enzyme complex was recently solved.<sup>22</sup> This has served as a model for the AlphaFold-derived models of the bacterial orthologues. The conservation of the transmembrane helix arrangement of this multi-pass membrane protein can be seen in Fig. 2.

There are several major differences between the yeast PMT and actinobacterial PMT: (a) the yeast enzyme is firstly a heterodimer, and (b) at the monomer level, differences appear in the loops from the catalytic centre and, where substrate recognition take place; (c) the secondary structure analysis indicates there is an extra helical segment in the *S. cerevisiae* PMT1 sequence (Fig. 3); (d) there is also the interface where the protein–protein interaction for dimer formation takes place, the  $\beta$ -trefoil MIR domain (MIR domains have a  $\beta$ -trefoil fold consisting of six  $\beta$ -hairpins arranged within a pseudo-threefold symmetry); (e) the yeast PMT enzyme monomer is also considerably larger ( $\sim 800$  residues compared with  $\sim 550$  residues for the actinobacterial enzyme). Currently, it is not known if the bacterial enzymes need to form dimers to be functional if so, the bacterial enzymes would form homodimers.

**Table 1** Percent sequence identity of PMT enzymes from select Actinobacteria<sup>a</sup> compared to *Saccharomyces cerevisiae* PMT1

PMT source – genome location/uniprot ID	PMT1	Mtb	Sco	Cg	Cfi
<i>S. cerevisiae</i> P33775 (PMT1)	100	22	26	26	22
<i>M. tuberculosis</i> Hv1002 (MtbPMT)	22	100	39	42	38
<i>S. coelicolor</i> Sco_3154 (ScoPMT)	26	39	100	34	43
<i>C. glutamicum</i> Cg_1014 (CgPMT)	26	42	34	100	31
<i>C. fimi</i> Celf_3080 (CfiPMT)	22	38	43	31	100

<sup>a</sup> These GT-39 enzymes are from bacterial species who have had a published survey of mannoproteins conducted, and they are compared to the *Saccharomyces* enzyme for which there is a cryo-EM based structure.

More recently, site-directed mutagenesis of the *S. coelicolor* PMT (ScoPMT) revealed an interesting phenomenon, in that many of the mutants did not produce detectable protein when analysed in western blots from PMT complementation experiments.<sup>23</sup> A similar problem was seen in work with *C. glutamicum* where the expression of the native protein (CgPMT) expressed *in trans* could not be detected by western blotting analysis, only *via* complementation of the enzyme's activity on reporter proteins.<sup>24</sup> Two other PMTs were examined in that study, one from *Cellulomonas fimi* (CfiPMT) and one from *Cellulomonas flavigena* (CflaPMT). Neither could be detected directly by western blot analysis. In the work on the *M. smegmatis* and MTB PMT, there was no indication of whether the PMT protein itself could be directly detected in



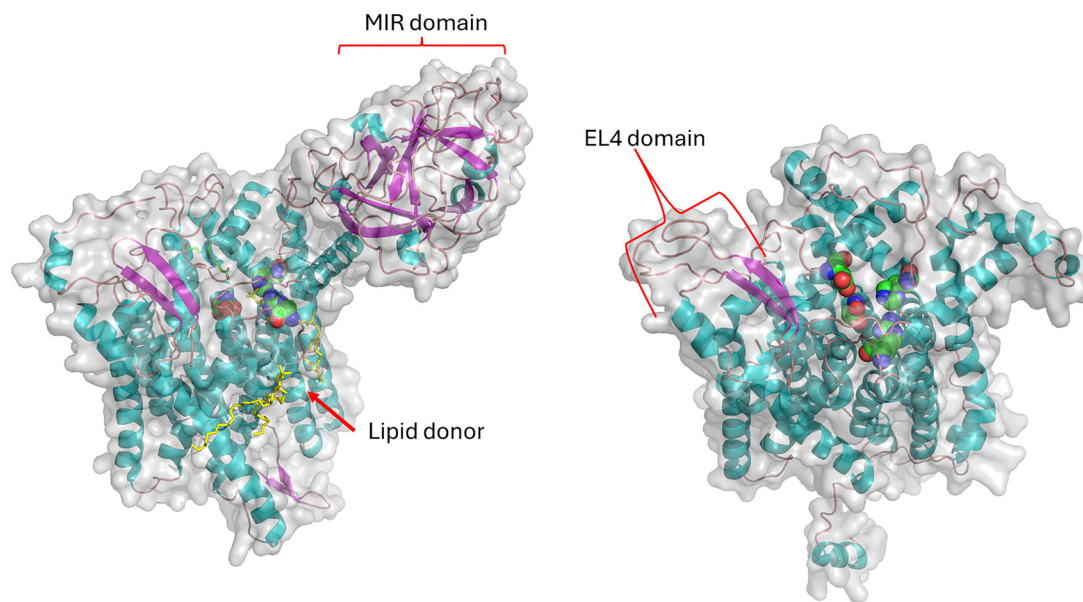


Fig. 2 View of the *S. cerevisiae* PMT1 (monomer) structure (left) and the *S. coelicolor* PMT (right) with the conserved catalytic residues highlighted as spheres. In the PMT1 structure the lipid donor is highlighted in yellow. The MIR  $\beta$ -trefoil domain is shown at the top of the yeast PMT1. The EL4 proline recognition domain is shown on the top of the ScoPMT. The yeast structure was derived from PDB 6P25, and *S. coelicolor* structure came from the AlphaFold model AF\_Q9RKD3-F1. Figures made with Pymol 3.1.

the complemented strains as those strains also used a reporter protein strategy.<sup>25</sup> The results suggest that the protein's expression and localization are sensitive to subtle changes in structure. It is also likely that the enzyme is present at very low levels, even as a recombinant protein. This lack of detectable over-expression would suggest that regulation may exist at the level of translation which limits how much PMT is made and inserted in the membrane.

Despite the challenges posed by the detection of PMT protein expression, a few key residues were identified in the ScoPMT, which were conserved with the yeast PMT. Site-directed mutagenesis suggested that the amino acids D113, H159, D233, and R510 were part of the active site.<sup>23</sup> (Fig. 2). In the 2D homology alignment against the yeast PMT1 secondary structure (Fig. 3), some regions of sequence conservation appear in the helical bundles and therefore likely key structural features. While these residues are conserved within this GT-39 enzyme family, it is also worth pointing out that overall, the sequence similarity between the bacterial orthologues is quite low. The highest sequence identity is between ScoPMT and CfiPMT at 43%.

Despite this low overall sequence identity, there was good complementation of the *C. glutamicum*  $\Delta$ PMT mutant by the CfiPMT even at only 31% protein sequence identity. However, the complementation was not complete, and this suggests distinct substrate specificity of each enzyme, as was observed by Saxena *et al.*<sup>24</sup> where complementation by orthologues was not always successful. Earlier work with the *S. coelicolor*  $\Delta$ PMT strain showed that the MtbPMT could not complement the glycosylation of the heterologously expressed MTB APA protein (Rv1860), nor the endogenous PTS protein, which is a  $\Phi$ 31

phage receptor.<sup>26</sup> This early observation coupled with recent ones suggests that PMTs require precise localization or other protein interactions to be active.

In a 2025 paper by Géraud *et al.*,<sup>27</sup> the overall topology of the MtbPMT was examined using N- and C-terminal fusion proteins in the surrogate host *M. smegmatis* with the  $\Delta$ PMT mutation. N-Terminal fusions of the reporter PhoA alkaline phosphatase were not active, whereas the C-terminal fusions were, showing that the C-terminal end is in the periplasmic space. In the same paper, the localization of the PMT was examined with fluorescence reporter fusions – and again the C-terminal end being a highly oxidizing environment, yielded low fluorescence compared to the N-terminal fusions, which were much brighter. This methodology shows that a GT-39 PMT enzyme can be detected in living cells, which means it may be possible to examine protein interactions to further define this complex post-translation event in actinobacteria.

The MtbPMT active site structure was also probed with a series of mutations which expand our understanding of those conserved residues mentioned above in the context of the ScoPMT. There was a newer finding from that study that the EL4 C-terminal domain provides substrate recognition similar to that of eukaryotic WW domain found in proline recognition domains (reviewed in ref. 28). This structural similarity between the EL4 and WW domains reflects the use of high proline content in many MTB glycopeptides.

For *M. tuberculosis* and *M. smegmatis*, a reporter protein approach was used to examine the MtbPMT and MsmegPMT knockouts.<sup>25</sup> The fasciclin domain protein (UniProt A0R2Q4\_MYCS2) was used as a secreted reporter protein. This 27 kDa major supernatant protein, proved to be an excellent



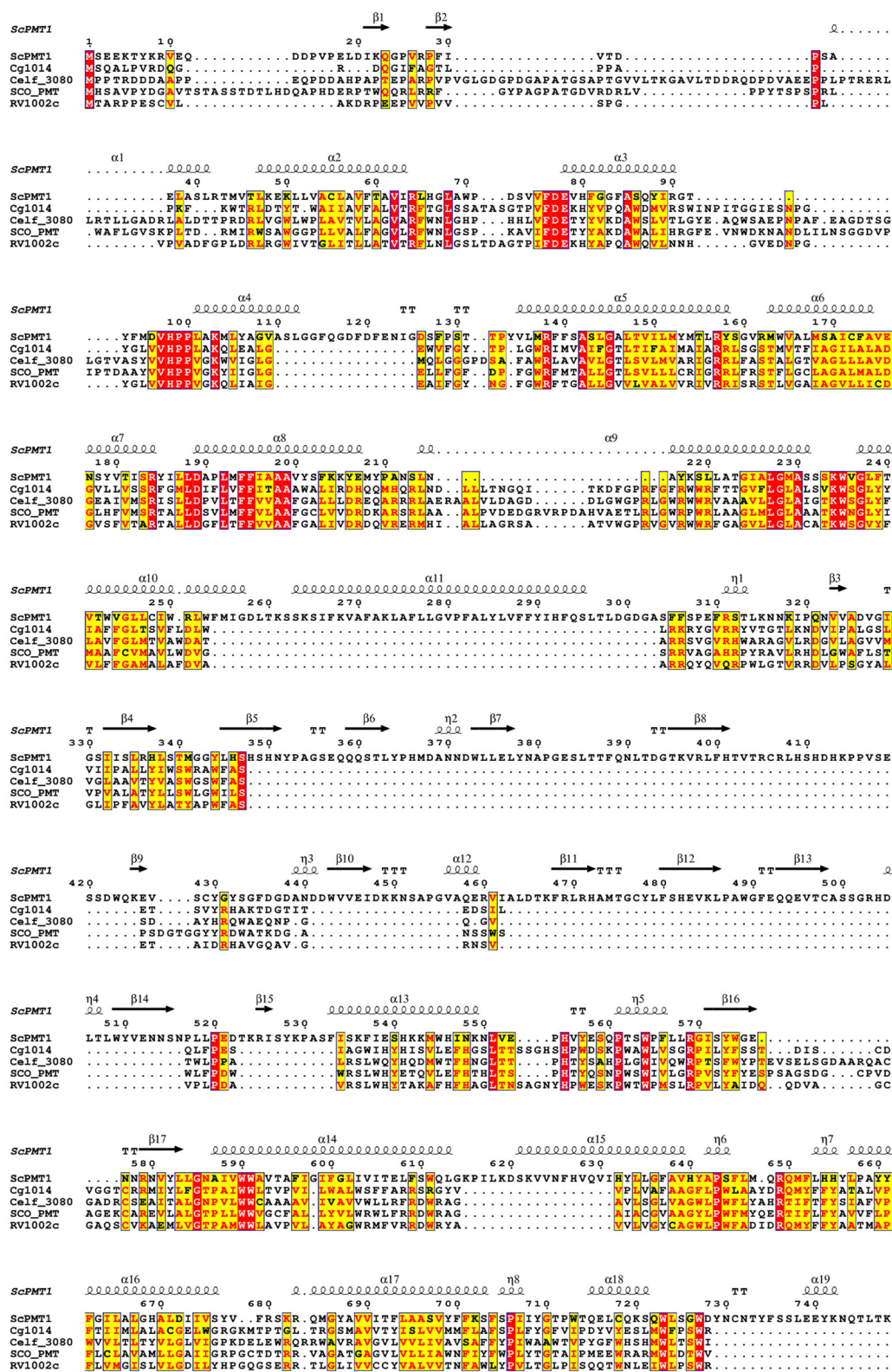


Fig. 3 2D sequence alignment against secondary structure elements of the yeast PMT1 structure. The sequence alignment was generated using the T-coffee server.<sup>30</sup>



protein for analysis, as the N-terminal peptide from aa30-55 contained the three sites of glycosylation. This reporter protein strategy is an exceptional tool to investigate structure-function of the pathway components. It should be noted that this 2013 study also identified the PimE protein as the elongating GT-87 family mannosyltransferase in *Mycobacterium* where di- and tri-saccharides are the end result. Other actinobacterial species lack a direct orthologue of PimE, but do harbour GT-87 family members which appear to perform the mannose chain elongation.

The consequences of the PMT gene knockouts have been examined in MTB, *M. smegmatis*,<sup>25</sup> *S. coelicolor*,<sup>11</sup> and *C. glutamicum*.<sup>12,24</sup> For *C. glutamicum* and *M. smegmatis*, there was no obvious growth defect, nor were there cell wall defects in the  $\Delta$ PMT mutant. However, in MTB and *S. coelicolor* there were some notable changes in lab growth and new  $\beta$ -lactam sensitivities. It should be mentioned that in the paper by Keenan *et al.* they state in reference to the  $\Delta$ PMT strains that “they (*M. smegmatis* and *C. glutamicum*) are strongly retarded in growth”, however in the reference given for work on *C. glutamicum*<sup>12</sup> and in our own work,<sup>24</sup> there was no evidence of a strong growth phenotype in *C. glutamicum* resulting from the PMT gene knockout.

MTB showed a strongly altered growth phenotype, especially on solid media, and a decreased virulence in a mouse model. There was no evidence of cell wall defects based on survival in the presence of chaotropes, which have been found to exacerbate cell wall defects.<sup>29</sup> The *S. coelicolor*  $\Delta$ PMT mutant showed an increase in susceptibility to some  $\beta$ -lactam antibiotics which points to peptidoglycan metabolism as a target pathway, and indeed when two glycoproteins involved in peptidoglycan metabolism were inactivated through gene knock out, an increased sensitivity to  $\beta$ -lactams was observed, similar to what was seen for the  $\Delta$ PMT strain.

## Target proteins of the PMT enzymes

The early reports of actinobacterial glycoproteins were likely dismissed as anomalies, as the existence of glycoproteins in bacteria was not widely recognized until modern mass spectral techniques proved covalent attachment and large-scale sequencing projects, revealing the genetic basis for protein glycosylation now known to be commonplace in the prokaryotic world.<sup>31-34</sup> Proteins specifically modified with mannose were observed in the 1970s in *M. smegmatis*,<sup>2</sup> and in that decade they were also observed in secreted cellulases from *Cellulomonas* species.<sup>35</sup> Characterization of the actual biochemical pathway for this post-translational modification took a longer time, and the breakthrough paper appeared only in 2005, in which the mycobacterial pathway was revealed to be like the eukaryotic GT-39 POMT pathway.<sup>36</sup> One differentiating feature of the bacterial proteins is that the mannose modification is only 1-3 mannose residues long and has no other sugars as part of the elongation process. Eukaryotic mannoproteins often have longer mannose glycans and in humans these can be capped by other sugars.<sup>1</sup>

Key questions for the bacterial PMT research remain from the work thus far: what are the targets of mannosylation and are these targets limited to only secreted proteins?

The application of proteomics technology has enabled a few surveys of proteins, as noted above. So far, the analysis of these surveys has largely been limited to secreted proteins. Or those found in culture supernatants and this analysis has not addressed how these proteins are represented in the wider actinobacterial community. To understand the broader implications of these proteins in the context of actinobacterial growth and physiology, it is important to look at the types of O-mannosylated proteins that have been identified.

### Secreted enzymes

Many members of the actinobacterial group produce large numbers of secreted glycosylated enzymes, which break down plant cell wall polysaccharides to provide ready access to fermentable sugars.<sup>37-39</sup> Perhaps the best studied mannosylated example is from *C. fimi*, where the POM initially appeared to be limited to the linker between domains of the GH10 family xylanase or GH6 family cellulase and a CBM2a cellulose-binding module.<sup>40,41</sup> This domain is a repeating sequence of Pro-Thr and is very heavily glycosylated with up to 32 mannose residues on the Celf\_3184 protein (Uniprot F4GZY2\_CELFA)<sup>24</sup> (Fig. 4). Using the Re-Glyco tool<sup>42</sup> gives us a model of how the glycan would be presented on this unstructured linker. From the model in Fig. 4, it is not hard to imagine that proteases would have reduced access to this linker with the glycan present. A related GH6 family protein from *C. flavigena* Cfla\_1896 (Uniprot D5UEY3\_CELFN) shows even more glycosylation on a related, but longer linker, with up to 65 mannose residues. Earlier work speculated that the glycosylated linker afforded some protection from proteases.<sup>40</sup> However, more recent work with other mannosylated linkers on fungal cellulases has suggested a role in actual binding to the cellulosic substrate.<sup>43</sup>

The linker region Celf\_3184 = PTTSPPTPTPTPTPTPTPTPTPTPTPTPTPTPT, where we assume most if not all T residues are

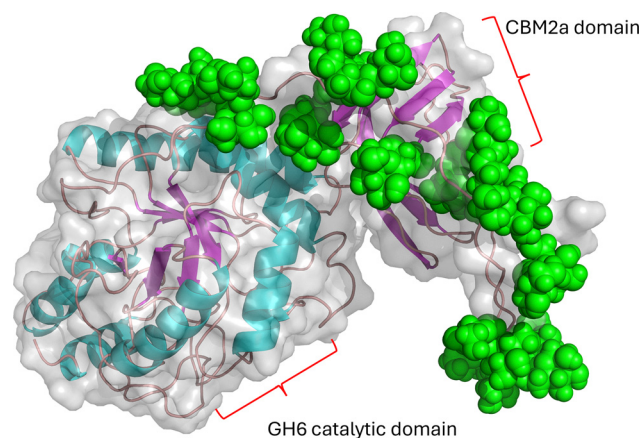


Fig. 4 AlphaFold model of Celf\_3184 after using Re-Glyco software to display possible mannose modifications.





## Proteins involved in peptidoglycan metabolism

One interesting group of common proteins from the mannosylation glycoproteome surveys are those involved in peptidoglycan metabolism. They have consistently shown up in the various glycoproteomics experiments that we are reviewing here (Table 3). As the metabolism of peptidoglycan requires proteins to work together and there are known protein complexes, the presence of glycosylation on some members of this group makes it tempting to speculate that a protein-carbohydrate interaction plays a role in the formation of these complexes.

### Penicillin-binding proteins

It is quite striking that several of these conserved actinobacterial proteins are involved in peptidoglycan metabolism. As mentioned above, there are at least two mannoproteins in *S. coelicolor* which are  $D_2D_2L$ -transpeptidases (Class C PBP). When these particular mannoproteins are deleted, these cells become more susceptible to cell-wall-active  $\beta$ -lactams. We identified and published a penicillin-binding protein from *Cellulomonas* and showed that the glycopeptide was part of the endo-peptidase domain of this PBP2-like protein (Class B PBP). The other actinobacteria we have been discussing here have a PBP1-like protein (Class A PBP), which is a two-domain protein comprised of the GT-51 glycosyltransferase domain as well as an endo-peptidase domain.

From the available data it does not appear that all the PBPs are glycoproteins, but rather that only select members are mannosylated. Further, they are not all in the same class of PBP. In MTB and *M. smegmatis*, PonA2 is Class A PBP1-like protein, in *C. fimi*, Celf\_0189 is a Class B PBP2-like protein, and in *S. coelicolor*, there are two Class C PBPs. In Fig. 6, the AlphaFold model of Celf\_0189 is shown as its re-glycosylated form based on published glycopeptide data. In this case the glycosylation appears to be in a place where it could influence protein-protein interactions. Fig. 7 shows the re-glycosylation model of SCO4847. As with the other models, the glycosylation appears to be on an unstructured loop, which is common for *O*-glycosylation. It is tempting to speculate that the

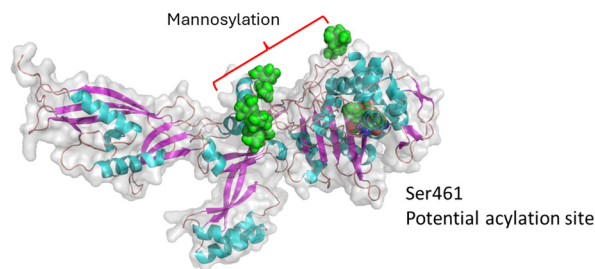


Fig. 6 AlphaFold model of Celf\_0189 after using Re-glyco software. Celf\_0189 is a *C. fimi* PBP2-like protein. The potential site of acylation by a  $\beta$ -lactam is Ser461 and is shown in a spheres representation.

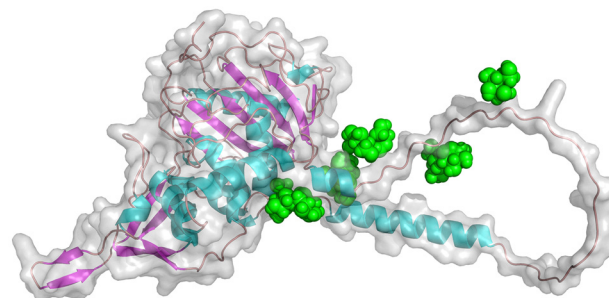


Fig. 7 AlphaFold model of SCO4847 after using Re-Glyco software. SCO4847 is a Class C PBP which is annotated as an endopeptidase.

combination of mobility of the loop and the presence of the mannose helps direct this protein to a binding partner(s) or to interact with the peptidoglycan substrate.

### PASTA domain kinases

The next group of proteins from the class of peptidoglycan active proteins is that of the PASTA domain kinases. These kinases are Ser/Thr active kinases, and PknB has been shown to be essential in MTB.<sup>49</sup> A PASTA domain kinase has been shown to be glycosylated in MTB, *S. coelicolor*, and *C. fimi* (Fig. 8).<sup>8,11,13</sup> To regulate cell growth and cell division in MTB, the PASTA domain kinases have been involved in binding mucopeptides

Table 3 Proteins involved in peptidoglycan metabolism

Organism	Penicillin binding protein	Pasta domain kinase	$D_2D_2L$ endopeptidases
<i>M. tuberculosis</i>	PonA2 – PBP1a-like RV3682 <sup>a</sup>	Rv0014c Uniprot-P9WI81	ND
<i>M. smegmatis</i>	PonA2 like Uniprot-A0R5I3_MYCS2	ND	ND
<i>S. coelicolor</i>	SCO4013 Uniprot-Q9ADP3	SCO3848 Uniprot-Q9XA16	SCO4394, SCO4847
<i>C. fimi</i>	Celf_0189 PBP2-like Uniprot-F4H5L1_CELFA	Celf_0029 Uniprot-F4H3V7_CELFA	ND
<i>C. glutamicum</i> <sup>b</sup>	PBP1-like Uniprot-Q8NLF6_CORGL	Uniprot-Q8NU98	Uniprot-Q8NSJ1_CORGL

ND – not determined. <sup>a</sup> The genome location and/or the uniprot ID is given for each protein. <sup>b</sup> This is included simply because *C. glutamicum* protein is homologous to the other orthologues.



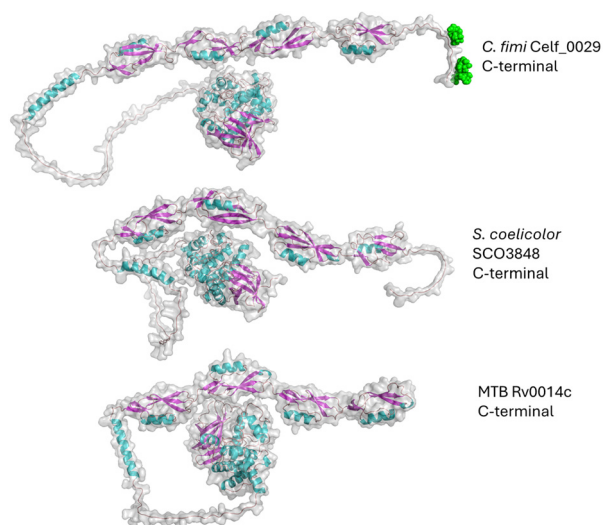


Fig. 8 AlphaFold models for PASTA domain kinases from *C. fimi*, *S. coelicolor*, and MTB.

and in a proposed model suggests PknB is targeted to the sites of peptidoglycan turnover.<sup>50</sup> In the case of the *C. fimi* PASTA domain kinase, the glycosylation is at the extreme end of the last PASTA domain. However, we do not have the glycopeptide information for the MTB enzyme.

We know from work with MTB PknB that all the PASTA domains are required for proper activity, and that they are responsible for localization within the cell. PknB is also an essential enzyme in MTB, as is the orthologue in *C. glutamicum*.<sup>51</sup> An interesting possibility is that the glycosylation on the PknB may play a role in interacting with PknB's various substrate proteins, or with muropeptides which help regulate its activity, and that this is why the MTB glycosylation mutant has a growth defect. This could also be why *S. coelicolor* has a growth defect. However, that does not appear to be the case for *C. glutamicum*, as the  $\Delta$ PMT mutant has no noticeable growth defect in a lab setting.

These enzymes are characterized by the presence of four repeated PASTA domains on the C-terminal end of the structure. The *C. fimi* protein has been through the Re-Glyco software to show where the mannose glycan (green spheres) is found, based on the reported glycopeptide data.

Looking at the models, it is easy to see the similarities between the *C. fimi* and *S. coelicolor* orthologues, where they both have a prominent tail on the C-terminal end. Because the MTB version lacks the C-terminal tail, it is not easy to see where the mannose could be on that version. Obviously, knowing where the glycan is on these other PASTA domain kinases would help with understanding why they carry glycans.

## GH3 glycoside hydrolases for peptidoglycan recycling

The MTB protein LpqI is annotated as a glycosidase from the GH3 family (Table 4). This protein was identified in the

Table 4 GH3 NagZ or LpqI orthologues

Organism	Gene/protein ID	Ref.
<i>M. tuberculosis</i>	LpqI, Rv0237 L7N6B0	8
<i>S. coelicolor</i>	Q93RU1_STRCO	Not yet identified as a glycoprotein
<i>C. fimi</i>	Q7WUL3-NAG3_CELFI	Not yet identified as a glycoprotein
<i>C. glutamicum</i>	Q8NLT5_CORGL	Not yet identified as a glycoprotein

glycoproteomics survey from Gonzalez-Zamorano *et al.*<sup>8</sup> The MTB protein has been extensively characterized, including a crystal structure, but from a truncated version expressed in *E. coli*.<sup>52</sup> This protein cleaves the GlcNAc–MurNAc glycan backbone of the peptidoglycan fragments, thereby helping to recycle the peptidoglycan. An orthologue has been cloned from *C. fimi*<sup>53</sup> but no investigation of its possible glycoprotein status was investigated. The *C. fimi* protein has also been shown to prefer phosphorylated substrates,<sup>54</sup> which would be unusual if the only purpose was GlcNAc–MurNAc hydrolysis. The *C. glutamicum* protein has also been partially characterized,<sup>55</sup> and it was shown to cleave simple *para*-nitrophenyl-*N*-acetylglucosamine. Like MTB, the *C. glutamicum* protein is an orphan GH3 family glycoside hydrolase in its parent organism, like LpqI in MTB, that lacks the ability to hydrolyze external glycans. Confirmation of the other members of this group as mannoproteins is needed, as well as showing that they indeed cleave the GlcNAc–MurNAc glycosidic bond. What still needs to be examined is the role that POM plays for this membrane-associated hydrolase.

## Protein O-mannosylation beyond discovery research

Recombinant expression of peptide-fusions and various mycobacterial proteins has been used to determine site preferences for the PMT and thus showed the power of expression in surrogate hosts. An early exploration of glycosylation sites using the algorithm Net-O-Glyc showed that eight of 11 predicted sites of MTB proteins could be glycosylated in *M. smegmatis*.<sup>56</sup> This same approach was employed in the seminal 2005 publication by VanderVen *et al.*, who used a non-glycoprotein from the Ag85 complex fused to an APA protein glycopeptide to investigate the secretion of mycobacterial glycoproteins.<sup>36</sup> Other early observations about POM were that certain surface proteins from MTB and *M. bovis* BCG that carried this modification might be vaccine targets. These immunodominant proteins are Ala-Pro rich antigen, APA (Rv1860); MPB83 (BQ2027\_MB2898)/MPB70 (BQ2027\_MB2900); and PstS1 (Rv0934). These proteins have been shown to act as adhesins and may provide a pathway to being taken up into macrophages where they hide from the immune system and proliferate.<sup>57</sup> They are also subject to upregulation under phosphate starvation, which leads to higher uptake by macrophages.<sup>58</sup> Since



they were discovered and their immunodominance observed, several papers have documented their production in alternate expression hosts.

One of the first papers to look at recombinant glycosylated APA proteins describes the use of the surrogate host *S. lividans* to produce the so-called 45 kDa and 47 kDa antigens (APA, Rv1860) as mannoproteins.<sup>59</sup> These glycoproteins had been shown to stimulate T-cells,<sup>60</sup> and immune serum from TB patients only reacted to the glycosylated forms of the proteins. This use of the surrogate host showed that native glycosylation was possible even with the heterologous protein. Optimizing the expression of APA in *S. lividans* showed that it is important to be careful when scaling up to maintain yields and protein quality.<sup>61</sup> This also suggests that regulating recombinant protein mannosylation is coupled to as-yet unknown metabolic cues in growing cells. It should also be noted that *S. lividans* was used more than a decade earlier to produce a mannosylated a “PT” sequence repeat linker in a *C. fimi* GH-10 family xylanase like the GH6 family endoglucanase noted earlier in this paper, and that the glycosylation was again true to the native pattern previously observed.<sup>41</sup> This shows that *S. lividans* is very versatile, as it can express both mycobacterial and *Cellulomonas* glycoproteins with accurate glycosylation.

The use of surrogate hosts has permitted a deeper understanding of the biochemistry of APA proteins as they interact with mammalian receptors. An interesting study from 2007 showed that the human pulmonary C-type lectin surfactant protein A required mannosylation on the 45/47 kDa APA and may be an adhesin that permits host cells to colonize by taking advantage of this innate immune receptor.<sup>62</sup> This study again made use of the *M. smegmatis* expression system to produce enough proteins for the receptor- and antibody-binding studies. *M. smegmatis* as a closer relative to MTB is excellent for providing authentic expression of these glycoproteins

A new surrogate host, *Rhodococcus erythropolis*, was introduced to express not only the MTB APA proteins but also the Pts1 and LprG proteins.<sup>63</sup> The MTB Pts1 protein is part of the phosphate transport system and an orthologue of the first *S. coelicolor* glycoprotein to be described in detail.<sup>5</sup> Saxena *et al.* noted that *Cellulomonas* also had a glycosylated orthologue of this protein.<sup>13</sup> In MTB, Pts1 is the 38 kDa antigen and is an immunodominant antigen<sup>64</sup> like the APA and LprG proteins. While this expression system did produce all three proteins with mannosylation, the yields were quite low, and it appears that the system would require a great deal of optimization before it could be as valuable as *S. lividans*.

More recently, the quest for vaccine candidate validation for APA has been examined in greater detail with protein produced with and without mannosylation.<sup>65</sup> For this paper, they used proteins produced in *E. coli* and *S. lividans* to provide the antigens. Immunization of mice revealed that the two proteins induce different cytokine responses – and that the mannosylated protein could stimulate a proliferative response from T-cells. The complex response of the cytokines and immune cells suggests that we still don't know the mechanism by which the mannosylated proteins stimulate the response, even though we

know some of the receptors that interact directly with the APA proteins.

The immunodominant mannosylated antigens and their possible role in host colonization have given rise to the idea of POM as a target for antibiotic development. From a chemical biology perspective, more work would have to be done to determine if POM is an ideal target for an antibiotic; other glycans in MTB have been studied much more closely to determine their suitability as target for antibiotics.<sup>66,67</sup> Recent research from Géraud *et al.*<sup>68</sup> has shown a reporter protein-based assay for the assessment of potential inhibitors. While the screen had only a few compounds, it demonstrated that an *in cellulo*-based assay for POM/PMT activity is feasible in Mycobacteria. A note of caution was issued regarding compound uptake however, as Mycobacteria have a very hydrophobic cell wall, which means that hydrophobic compounds are likely to have limited uptake through the cell-wall barrier.

## Conclusions

The literature on protein mannosylation in actinobacteria has been slowly increasing over the last 50 years. The predominant focus has been on the MTB cell-wall antigen proteins, with a secondary interest in secreted glycoside hydrolases. While structural details of the protein targets of POM are scarce, we have presented here a snapshot of models for discussion using modern software tools that were previously unavailable. The vast improvements in predicting protein structure by AlphaFold coupled with the ability to use web-based molecular dynamics allows us to “visualize” the products of POM. It is striking that we now have models which allow us to test structure-function in these proteins, a case in point being the PMT work from Géraud *et al.*,<sup>27</sup> where site-directed mutagenesis based on the AlphaFold model has provided new insight into the EL4 domain on the MTB PMT.

The use of re-glycosylated protein models also allows us to think about how these glycans can be used to interact with other proteins. This is especially true in the case of the group of proteins involved in peptidoglycan metabolism. Peptidoglycan-active proteins are known to form complexes, and glycan-protein interactions are common in other cellular contexts. This leads to a discussion of how the lack of POM leads to growth defects. The major defect appears to be cell-wall-related in both MTB and *S. coelicolor*, which certainly fits with the idea of some defect in peptidoglycan metabolism. While the lack of POM in MTB and *S. coelicolor* leads to some biological consequence, this is not the case for *C. glutamicum* and *M. smegmatis*. This sets up the next area for follow-up research in *C. glutamicum* and *M. smegmatis*: under what conditions would lack of POM have biological consequences?

MTB's use of POM to provide a means to enter macrophages (along with other glycans) continues to present an interesting intervention point for anti-infective strategies. The fact that we are still learning about those adhesin/antigenic proteins that were discovered decades ago makes it clear that we have more to discover about the functional role of POM for MTB. Without



a solid conclusion about POM's biological role in MTB outside of infectivity, what remains to be discovered in the nearly 4000 other species of bacteria that also harbour a GT-39 family PMT? Questions about the role of the glycosylation will need to be addressed with existing genetic manipulations, but finding a phenotype for lab-grown strains will require consideration about different growth conditions, and perhaps a variety of stressors to assist in finding the link between POM and Actinobacterial growth and physiology. Our hope is that this review will spark some interest in developing new ways to probe this enigmatic post-translational modification.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

There is no original data in this manuscript to make available.

## Acknowledgements

This work was partially funded by an NSERC grant RGPIN-2019-04640 to WW. Thanks to Dr HIRAK SAXENA for critical reading of the manuscript.

## References

- M. O. Sheikh, S. M. Halmo and L. Wells, *Glycobiology*, 2017, **27**, 806–819.
- J. C. Schultz and K. Takayama, *Biochim. Biophys. Acta*, 1975, **381**, 175–184.
- K. M. Dobos, K. Swiderek, K. H. Khoo, P. J. Brennan and J. T. Belisle, *Infect. Immun.*, 1995, **63**, 2846–2853.
- B. Coddeville, S. W. Wu, E. Fabre, C. Brassart, Y. Rombouts, A. Burguière, L. Kremer, K. H. Khoo, E. Ellass-Rochard and Y. Guérardel, *J. Proteomics*, 2012, **75**, 5695–5705.
- S. Wehmeier, A. S. Varghese, S. S. Gurcha, B. Tissot, M. Panico, P. Hitchen, H. R. Morris, G. S. Besra, A. Dell and M. C. Smith, *Mol. Microbiol.*, 2009, **71**, 421–433.
- M. Hartmann, A. Barsch, K. Niehaus, A. Puhler, A. Tauch and J. Kalinowski, *Arch. Microbiol.*, 2004, **182**, 299–312.
- S. L. Michell, A. O. Whelan, P. R. Wheeler, M. Panico, R. L. Easton, A. T. Etienne, S. M. Haslam, A. Dell, H. R. Morris, A. J. Reason, J. L. Herrmann, D. B. Young and R. G. Hewinson, *J. Biol. Chem.*, 2003, **278**, 16423–16432.
- M. González-Zamorano, G. Mendoza-Hernández, W. Xolalpa, C. Parada, A. J. Vallecillo, F. Bigi and C. Espitia, *J. Proteome Res.*, 2009, **8**, 721–733.
- G. T. Smith, M. J. Sweredoski and S. Hess, *J. Proteomics*, 2014, **97**, 296–306.
- P. Tucci, M. Portela, C. R. Chetto, G. González-Sapienza and M. Marín, *PLoS One*, 2020, **15**, e0221837.
- T. Keenan, A. Dowle, R. Bates and M. C. M. Smith, *mBio*, 2019, **10**(3), e01092-19.
- M. Mahne, A. Tauch, A. Puhler and J. Kalinowski, *FEMS Microbiol. Lett.*, 2006, **259**, 226–233.
- H. Saxena, N. Buenbrazo, W. Y. Song, C. Li, D. Brochu, A. Robotham, W. Ding, L. Tessier, R. Chen, J. Kelly and W. Wakarchuk, *Glycobiology*, 2023, **33**, 512–524.
- S. Yang, S. Sui, Y. Qin, H. Chen, S. Sha, X. Liu, G. Deng and Y. Ma, *APMIS*, 2022, **130**, 181–192.
- J. Zheng, L. Liu, C. Wei, B. Liu and Q. Jin, *IUBMB Life*, 2022, **74**, 221–234.
- A. G. Birhanu, S. A. Yimer, S. Kalayou, T. Riaz, E. D. Zegeye, C. Holm-Hansen, G. Norheim, A. Aseffa, M. Abebe and T. Tønjum, *Sci. Rep.*, 2019, **9**, 2927.
- S. Yari, P. Afrough, F. Yari, M. Ghazanfari Jajin, A. Fateh and A. Hadizadeh Tasbiti, *Sci. Rep.*, 2023, **13**, 22194.
- C. Lizak, S. Gerber, S. Numao, M. Aebi and K. P. Locher, *Nature*, 2011, **474**, 350–355.
- S. Strahl-Bolsinger and A. Scheinost, *J. Biol. Chem.*, 1999, **274**, 9068–9075.
- I. S. B. Larsen, Y. Narimatsu, H. J. Joshi, L. Siukstaite, O. J. Harrison, J. Brasch, K. M. Goodman, L. Hansen, L. Shapiro, B. Honig, S. Y. Vakhrushev, H. Clausen and A. Halim, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 11163–11168.
- H. N. Cooper, S. S. Gurcha, J. Nigou, P. J. Brennan, J. T. Belisle, G. S. Besra and D. Young, *Glycobiology*, 2002, **12**, 427–434.
- L. Bai, A. Kovach, Q. You, A. Kenny and H. Li, *Nat. Struct. Mol. Biol.*, 2019, **26**, 704–711.
- N. D. M. Holman, A. J. Wilkinson and M. C. M. Smith, *Microbiology*, 2021, **167**, 001103.
- H. Saxena, R. Patel, J. Kelly and W. Wakarchuk, *Glycobiology*, 2025, **35**(1), cwae095.
- C. F. Liu, L. Tonini, W. Malaga, M. Beau, A. Stella, D. Bouyssie, M. C. Jackson, J. Nigou, G. Puzo, C. Guillhot, O. Burlet-Schiltz and M. Rivière, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 6560–6565.
- L. E. Córdova-Dávalos, C. Espitia, G. González-Cerón, R. Arreguín-Espinosa, G. Soberón-Chávez and L. Servín-González, *FEMS Microbiol. Lett.*, 2014, **350**, 72–82.
- N. Géraud, C. Rivière, C. Falcou, G. Cioci, C. Froment, V. Gervais, J. Marcoux, M. Gilleron, J. Nigou, E. Fabre and M. Rivière, *Commun. Biol.*, 2025, **8**, 1175.
- A. Zarrinpar and W. A. Lim, *Nat. Struct. Biol.*, 2000, **7**, 611–613.
- E. Gelman, J. D. McKinney and N. Dhar, *Antimicrob. Agents Chemother.*, 2012, **56**, 3610–3614.
- J. F. Taly, C. Magis, G. Bussotti, J. M. Chang, P. Di Tommaso, I. Erb, J. Espinosa-Carrasco, C. Kemena and C. Notredame, *Nat. Protoc.*, 2011, **6**, 1669–1682.
- J. H. Anonsen, A. Vik, W. Egge-Jacobsen and M. Koomey, *J. Proteome Res.*, 2012, **11**, 5781–5793.
- A. Alemka, H. Nothaft, J. Zheng and C. M. Szymanski, *Infect. Immun.*, 2013, **81**, 1674–1682.
- L. Balonova, L. Hernychova, B. F. Mann, M. Link, Z. Bilkova, M. V. Novotny and J. Stulik, *J. Proteome Res.*, 2010, **9**, 1995–2005.
- J. M. Chan, A. Gori, A. H. Nobbs and R. S. Heyderman, *Front. Microbiol.*, 2020, **11**, 593356.



- 35 P. Beguin and H. Eisen, *Eur. J. Biochem.*, 1978, **87**, 525–531.
- 36 B. C. VanderVen, J. D. Harder, D. C. Crick and J. T. Belisle, *Science*, 2005, **309**, 941–943.
- 37 M. L. Langsford, N. R. Gilkes, W. W. Wakarchuk, D. G. Kilburn, R. C. Miller and R. A. J. Warren, *J. Gen. Microbiol.*, 1984, **130**, 1367–1376.
- 38 G. Elegir, G. Szakács and T. W. Jeffries, *Appl. Environ. Microbiol.*, 1994, **60**, 2609–2615.
- 39 S. Zhang, G. Lao and D. B. Wilson, *Biochemistry*, 1995, **34**, 3386–3395.
- 40 M. L. Langsford, N. R. Gilkes, B. Singh, B. Moser, R. C. Miller Jr, R. A. J. Warren and D. G. Kilburn, *FEBS Lett.*, 1987, **225**, 163–167.
- 41 A. M. MacLeod, N. R. Gilkes, L. Escote-Carlson, R. A. Warren, D. G. Kilburn and R. C. Miller, Jr., *Gene*, 1992, **121**, 143–147.
- 42 C. M. Ives, O. Singh, S. D'Andrea, C. A. Fogarty, A. M. Harbison, A. Satheesan, B. Tropea and E. Fadda, *Nat. Methods*, 2024, **21**, 2117–2127.
- 43 L. Chen, M. R. Drake, M. G. Resch, E. R. Greene, M. E. Himmel, P. K. Chaffey, G. T. Beckham and Z. Tan, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 7612–7617.
- 44 J. Russell, S. K. Kim, J. Duma, H. Nothhaft, M. E. Himmel, Y. J. Bomble, C. M. Szymanski and J. Westpheling, *Biotechnol. Biofuels*, 2018, **11**, 259.
- 45 D. Chung, N. S. Sarai, B. C. Knott, N. Hengge, J. F. Russell, J. M. Yarbrough, R. Brunecky, J. Young, N. Supekar, T. Vander Wall, D. W. Sammond, M. F. Crowley, C. M. Szymanski, L. Wells, P. Azadi, J. Westpheling, M. E. Himmel and Y. J. Bomble, *ACS Sustainable Chem. Eng.*, 2019, **7**, 4792–4800.
- 46 C. M. Sasseti, D. H. Boyd and E. J. Rubin, *Mol. Microbiol.*, 2003, **48**, 77–84.
- 47 S. Pandey, A. Sharma, D. Tripathi, A. Kumar, M. Khubaib, M. Bhuwan, T. K. Chaudhuri, S. E. Hasnain and N. Z. Ehtesham, *PLoS One*, 2016, **11**, e0150288.
- 48 S. Pandey, D. Tripathi, M. Khubaib, A. Kumar, J. A. Sheikh, G. Sumanlatha, N. Z. Ehtesham and S. E. Hasnain, *Front. Cell. Infect. Microbiol.*, 2017, **7**, 38.
- 49 Y. Chawla, S. Upadhyay, S. Khan, S. N. Nagarajan, F. Forti and V. K. Nandicoori, *J. Biol. Chem.*, 2014, **289**, 13858–13875.
- 50 M. Mir, J. Asong, X. Li, J. Cardot, G.-J. Boons and R. N. Husson, *PLoS Pathog.*, 2011, **7**, e1002182.
- 51 M. Fiuza, M. J. Canova, I. Zanella-Cléon, M. Becchi, A. J. Cozzone, L. M. Mateos, L. Kremer, J. A. Gil and V. Molle, *J. Biol. Chem.*, 2008, **283**, 18099–18112.
- 52 P. J. Moynihan, I. T. Cadby, N. Veerapen, M. Jankute, M. Crosatti, G. V. Mukamolova, A. L. Lovering and G. S. Besra, *Nat. Commun.*, 2019, **10**, 2647.
- 53 C. Mayer, D. J. Vocadlo, M. Mah, K. Rupitz, D. Stoll, R. A. Warren and S. G. Withers, *FEBS J.*, 2006, **273**, 2929–2941.
- 54 S. S. Macdonald, M. Blaukopf and S. G. Withers, *J. Biol. Chem.*, 2015, **290**, 4887–4895.
- 55 C. Matano, S. Kolkenbrock, S. N. Hamer, E. Sgobba, B. M. Moerschbacher and V. F. Wendisch, *BMC Microbiol.*, 2016, **16**, 177.
- 56 J. L. Herrmann, R. Delahay, A. Gallagher, B. Robertson and D. Young, *FEBS Lett.*, 2000, **473**, 358–362.
- 57 M. Esparza, B. Palomares, T. García, P. Espinosa, E. Zenteno and R. Mancilla, *Scand. J. Immunol.*, 2015, **81**, 46–55.
- 58 P. Espinosa-Cueto, A. Magallanes-Puebla and R. Mancilla, *BMC Immunol.*, 2020, **21**, 34.
- 59 M. Lara, L. Servín-González, M. Singh, C. Moreno, I. Cohen, M. Nimtz and C. Espitia, *Appl. Environ. Microbiol.*, 2004, **70**, 679–685.
- 60 C. Horn, A. Namane, P. Pescher, M. Rivière, F. Romain, G. Puzo, O. Bâzu and G. Marchal, *J. Biol. Chem.*, 1999, **274**, 32023–32030.
- 61 R. A. Gamboa-Suasnavart, N. A. Valdez-Cruz, L. E. Cordova-Dávalos, J. A. Martínez-Sotelo, L. Servín-González, C. Espitia and M. A. Trujillo-Roldán, *Microb. Cell Fact.*, 2011, **10**, 110.
- 62 A. Ragas, L. Roussel, G. Puzo and M. Rivière, *J. Biol. Chem.*, 2007, **282**, 5133–5142.
- 63 A. J. Vallecillo, C. Parada, P. Morales and C. Espitia, *Microb. Cell Fact.*, 2017, **16**, 12.
- 64 C. Espitia, M. Elinos, R. Hernández-Pando and R. Mancilla, *Infect. Immun.*, 1992, **60**, 2998–3001.
- 65 J. A. Martínez-Sotelo, A. J. Vallecillo, C. Parada, E. Segura, J. Campuzano, M. Silva-Miranda, L. Servín-González and C. Espitia, *Curr. Microbiol.*, 2024, **81**, 197.
- 66 K. A. Abrahams and G. S. Besra, *Parasitology*, 2018, **145**, 116–133.
- 67 L. J. Alderwick, H. L. Birch, A. K. Mishra, L. Eggeling and G. S. Besra, *Biochem. Soc. Trans.*, 2007, **35**, 1325–1328.
- 68 N. Géraud, C. Falcou, J. Parra, C. Froment, D. Rengel, O. Burlet-Schiltz, J. Marcoux, J. Nigou, M. Rivière and E. Fabre, *Glycobiology*, 2023, **33**, 1139–1154.

