ArnT proteins that catalyse the glycosylation of lipopolysaccharide share common features with bacterial N-oligosaccharyltransferases

Tavares-Carreón, F., Mohamed, Y. F., Andrade, A., & Valvano, M. A. (2016). ArnT proteins that catalyse the glycosylation of lipopolysaccharide share common features with bacterial N-oligosaccharyltransferases. Glycobiology, 26(3), 286-300. DOI: 10.1093/glycob/cwv095

Published in:
Glycobiology

Document Version:
Peer reviewed version

ArnT proteins that catalyse the glycosylation of lipopolysaccharide share common features with bacterial N-oligosaccharyltransferases

Faviola Tavares-Carreón[a], Yasmine Fathy Mohamed[b,c], Angel Andrade[a,d], and Miguel A. Valvano[a,b].

Department of Microbiology and Immunology, University of Western Ontario, London, Ontario, Canada, N6A 5C1,[a] Centre for Infection and Immunity, Queen's University Belfast, BT9 5GZ, Belfast, United Kingdom,[b] Department of Pharmaceutical Microbiology, Faculty of Pharmacy, Alexandria University, Egypt,[c] and Departamento de Microbiología, Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, México[d]

Address correspondence to Miguel A. Valvano, m.valvano@qub.ac.uk

Running title: Structure-function of the ArnT glycosyltransferase

Keywords: Antimicrobial peptides/ lipid A modification/membrane proteins/ oligosaccharyltransferases.

Abstract

ArnT is a glycosyltransferase that catalyses the addition of 4-amino-4-*deoxy*-L-arabinose (L-Ara4N) to the lipid A moiety of the lipopolysaccharide. This is a critical modification enabling bacteria to resist killing by antimicrobial peptides. ArnT is an integral inner membrane protein consisting of 13 predicted transmembrane helices and a large periplasmic C-terminal domain. We report here the identification of a functional motif with a canonical consensus sequence $DEXRYAX_{(5)}MX_{(3)}GXWX_{(9)}YFEKPX_{(4)}W$ spanning the first periplasmic loop, which is highly conserved in all ArnT proteins examined. Site-directed mutagenesis demonstrated the contribution of this motif in ArnT function, suggesting that these proteins have a common mechanism. We also demonstrate that the *Burkholderia cenocepacia* and *Salmonella enterica* serovar Typhimurium ArnT C-terminal domain is required for polymyxin B resistance *in vivo*. Deletion of the C-terminal domain in *B. cenocepacia* ArnT resulted in a protein with significantly reduced *in vitro* binding to a lipid A fluorescent substrate and unable to catalyse lipid A modification with L-Ara4N. An *in silico* predicted structural model of ArnT strongly resembled the tertiary structure of *Campylobacter lari* PglB, a bacterial oligosaccharyltransferase involved in protein N-glycosylation. Therefore, distantly related oligosaccharyltransferases from ArnT and PglB families operating on lipid and polypeptide substrates, respectively, share unexpected structural similarity that could not be predicted from direct amino acid sequence comparisons. We propose that lipid A and protein glycosylation enzymes share a conserved catalytic mechanism despite their evolutionary divergence.

Introduction

The lipopolysaccharide (LPS) resides on the outer leaflet of the Gram-negative bacterial outer membrane and consists of three distinct structural regions: O antigen, core oligosaccharide, and lipid A. Both O antigen and core are made of covalently linked sugars, whereas lipid A has fatty acids bonded to a glucosamine disaccharide, which is embedded in the outer membrane (Nikaido 2003). Lipid A is the major signaling component of LPS and elicits innate immune responses ranging from mild inflammation to lethal sepsis (Miller, et al. 2005, Mogensen 2009). The lipid A backbone structure is highly conserved in most bacteria, but it is also remodelled by the addition or subtraction of components in response to environmental cues (Needham and Trent 2013). For example, covalent modification of phosphate groups in the lipid A by the addition of positively charged moieties, such as phosphoethanolamine (Kim, et al. 2006, Raetz, et al. 2007), galactosamine (Wang, et al. 2009), glucosamine (Marr, et al. 2008) and 4-amino-4-deoxy-L-arabinose (L-Ara4N) (Molinaro, et al. 2015, Trent, et al. 2001), reduce the overall negative charge of the molecule. These covalent modifications confer resistance to cationic antimicrobial peptides such as polymyxin B (PmB) and aminoglycoside antibiotics, and are associated with enhanced bacterial virulence (Needham and Trent 2013).

Burkholderia cenocepacia is an opportunistic pathogen infecting immune compromised individuals (Mahenthiralingam, et al. 2008, Vandamme and Dawyndt 2011), especially those with cystic fibrosis. The modification of lipid A of *B. cenocepacia* with L-Ara4N is constitutive and essential for *B. cenocepacia* cell viability (Ortega, et al. 2007), as L-Ara4N-modified lipid A is required for LPS export to the outer membrane (Hamad, et al. 2012, Ortega, et al. 2007). Further, L-Ara4N ligation to lipid A confers PmB resistance to *B. cenocepacia* and other bacteria such as *Salmonella enterica* serovar Typhimurium*, Pseudomonas aeruginosa, Bordetella bronchiseptica*, *Yersinia enterocolitica, Yersinia pestis,* and *Proteus mirabilis* (Boll, et al. 1994, Ernst, et al. 1999, Gunn, et al. 1998, Guo, et al. 1997, Kaca, et al. 1990, Loutet and Valvano 2011, Rolin, et al. 2014).

L-Ara4N is formed as an undecaprenyl-phosphate (Und-P)-linked precursor that is translocated across the plasma membrane and subsequently transferred to lipid A by ArnT (Raetz, et al. 2007) (Figure 1). ArnT is an integral membrane protein that belongs to the glycosyltransferase C superfamily (GT-C) (Liu and Mushegian 2003). The GT-C family proteins utilize Und-P-linked sugar substrates (Lairson, et al. 2008) and consists of proteins characterized by multiple transmembrane helices, a large periplasmic loop close to the N-terminus, and an extended C-terminal domain also located in the periplasm (Igura, et al. 2008, Korkegian, et al. 2014, Kus, et al. 2008, Lizak, et al. 2011, Maeda, et al. 2001, Strahl-Bolsinger, et al. 1993, Tavares-Carreon, et al. 2015). The precise mechanism of lipid A glycosylation with L-Ara4N by ArnT is unknown, but the structural information on the anomeric nature of the Und-P-α-L-Ara4N donor (Trent, et al. 2001) and the β configuration of the L-Ara4N linkage to the lipid A phosphates (Zhou, et al. 2000) suggests that ArnT is an inverting glycosyltransferase. It is however unclear how ArnT recognizes Und-P-L-Ara4N and lipid A and which amino acid residues are involved in the ligation reaction. Conceivably, residues exposed to the periplasm are involved in the ligation of L-Ara4N to the lipid A phosphate groups. This may occur either by a direct role in catalysis or binding to both donor (Und-P-L-Ara4N) and acceptor (lipid A) molecules, as it has been observed for other glycosyltransferases that carry similar reactions such as the *E. coli* WaaL (Boll, et al. 1994, Perez, et al. 2008), *Campylobacter* PglB (Kowarik, et al. 2006, Lizak, et al. 2011), *Shigella flexneri* GtrIV protein of (Nair, et al. 2011), and Emb enzymes of *Mycobacterium tuberculosis* (Belanger, et al. 1996, Telenti, et al. 1997). Agreeing with this

notion, we have previously identified four conserved periplasmic residues predicted to be required for ArnT catalysis (Tavares-Carreon, et al. 2015). These four residues, corresponding to tyrosine-43, lysine-69, arginine-254, and glutamic acid-493 (according to the *B. cenocepacia* ArnT protein), reside at conserved positions and are required for the activity of both *B. cenocepacia* and *S. enterica* ArnT proteins (Tavares-Carreon, et al. 2015).

In this work, we identify a signature motif, DEXR<u>Y</u>AX$_{(5)}$MX$_{(3)}$GXWX$_{(9)}$YFE<u>K</u>PX$_{(4)}$W (the tyrosine-43 and lysine-69 residues are underlined) , which is highly conserved in all members of the ArnT protein family and resides in the first periplasmic loop; this signature motif includes the residues that were previously found to be required for ArnT activity (Tavares-Carreon, et al. 2015). We also characterized the large ArnT C-terminal domain and show that sequential deletions of this domain result in *B. cenocepacia* and *S. enterica* ArnT proteins that could not restore PmB resistance in Δ*arnT* mutant strains. Importantly, we show that the ArnT C-terminal domain is required for lipid A binding, explaining why this region may be needed for ArnT function. Further, we show by molecular modelling that ArnT has similar predicted structural conformation as the bacterial N-oligosaccharyltransferase PglB (Lizak, et al. 2011), suggesting that that protein and lipid oligosaccharyltransferases of two distantly related protein families share common structural features.

Results

ArnT proteins share a functional motif that defines the members of this family

ArnT is classified within GT-C superfamily, particularly the GT83 glycosyltransferase family of the CAZy database (carbohydrates-active enzymes) (Lombard, et al. 2014). The GT-C superfamily includes large hydrophobic proteins with eight to thirteen predicted transmembrane helices. ArnT primary sequences from *B. cenocepacia* and *S. enterica* share an overall sequence identity of 29% and a similarity of 64.5% (Figure 2). Both proteins consist of two domains, an N-terminal transmembrane domain spanning approximately 437 residues, which is composed by 13 transmembrane helices, and a large soluble C-terminal domain, including the remaining 121 amino acids, oriented towards the periplasm (Tavares-Carreon, et al. 2015). GT-C family members have a modified DXD signature (conserved in glycosyltransferases) corresponding to EXD, DXE, DDX or DEX (Liu and Mushegian 2003). We identified a ^{39}DEX^{41} motif in ArnT sequence from *B. cenocepacia* and *S. enterica* (Figure 2, dashed-lined box). To determine whether the DEX motif is necessary for the activity of *B. cenocepacia* ArnT, we replaced the aspartic acid and glutamic acid residues by alanine, and assessed the abilities of mutant proteins to restore PmB resistance in the *B. cenocepacia* Δ*arnT-arnBC*$^{+}$*lptG*$_{D31H}$ suppressor strain MH55 (Table 1) (herein Δ*arnT*). This strain is highly sensitive to PmB, but viable due to a suppressor mutation that rescues export of lipid A-core to the outer membrane (Hamad, et al. 2012). Alanine replacements are commonly used to identify the functional relevance of predicted conserved aspartates in DXD motifs of many glycosyltransferases (Birch, et al. 2009, Breton and Imberty 1999, Korkegian, et al. 2014, Korres and Verma 2006, Liu and Mushegian 2003, Maeda, et al. 2001, Perrin-Tricaud, et al. 2011, Ramiscal, et al. 2010). We observed that *B. cenocepacia* ArnT$_{D39A}$ could not restore PmB resistance when expressed in Δ*arnT* (Figure 3A), similar to the negative control only containing the cloning vector (Figure 3A, pSCrhaB2). In contrast, ArnT$_{E40A}$ remained functional since expression of this protein restored PmB resistance at a similar level as that the wild-type strain (Figure 3A). The alanine replacement mutant proteins were detected in the membrane fraction by immunoblotting and they were expressed at wild type levels (Figure 3B). The corresponding residues in *S. enterica* ArnT, aspartic acid-32 and

glutamic acid-33, were also investigated. The *Salmonella* $ArnT_{Se\text{-}FLAG\text{-}10xHis}$ fusion resulted in a protein that could not complement *E. coli* Δ*arnT* strain, likely due to structural constrains associated with a proline at the C-terminal end of the protein (Bretscher, et al. 2006). This was resolved by replacing proline-547 by alanine, which did not alter ArnT function (Tavares-Carreon, et al. 2015). Therefore, the alanine replacements in aspartic acid-32 and glutamic acid-33 were made in the P547A $ArnT_{Se\text{-}FLAG\text{-}10xHis}$ protein. Both residues were required for function since the respective alanine substituted proteins could not restore PmB resistance in the *E. coli* Δ*arnT* strain (Figure 3C). Therefore, depending of the source of the ArnT proteins one or both residues in the DEX motif are required for function.

The functional data for *B. cenocepacia* Δ*arnT* were consistent with the chemical structure of the *B. cenocepacia* lipid A, as examined by mass spectrometry using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) in the negative ion mode. The mass spectrum of the wild-type lipid A sample showed the presence of L-Ara4N, as indicated by ion peaks corresponding to tetra-acylated lipid A with L-Ara4N in the monophosphate (at *m/z* 1495.3) and diphosphate (at *m/z* 1575.3) forms (Figure 3D) in agreement with the reported *B. cenocepacia* lipid A structure (Hamad, et al. 2012, Silipo, et al. 2005, Tavares-Carreon, et al. 2015). These ion peaks were absent in the lipid A spectrum of Δ*arnT* alone and Δ*arnT* expressing $ArnT_{D39A}$ (Figure 3E-F), while ion peaks containing L-Ara4N were restored in the lipid A sample from Δ*arnT* containing $ArnT_{E40A}$ (Figure 3G). These results indicate that the aspartic acid-39 in the DEX motif is critical for *B. cenocepacia* ArnT function.

Multiple sequence alignments of ArnT proteins from different bacteria were compared with our topology prediction data to identify transmembrane segments and loops regions. Previously, we reported two identifiable motifs, ^{42}RYA^{44} and ^{66}YFEKP^{70}, in which residues tyrosine-43 and lysine-69 were found to be necessary for ArnT function (Tavares-Carreon, et al. 2015). Also, a conserved tryptophan residue is present at position 75 (Tavares-Carreon, et al. 2015). All of these motifs including the DEX motif examined here reside in the first periplasmic loop of ArnT. We hypothesized that the three motifs ^{39}DEX^{41}, ^{42}RYA^{44}, ^{66}YFEKP^{70}, and W^{75} form a larger motif unique to ArnT protein family members (Figure 4A). The fingerprint sequence ^{39}DEXRYAX$_{(5)}$MX$_{(3)}$GXWX$_{(9)}$YFEKPX$_{(4)}$W^{75}, where X represent a variable but usually hydrophobic amino acid, was used to search the Swiss-Prot and TrEMBL databases with the ScanProsite tool (http://prosite.expasy.org/scanprosite/) resulting in 1,133 hits that corresponded to proteins putatively annotated as ArnT. From these 1,119 sequences were from different species of Proteobacteria, and the others were from Actinobacteria (5 sequences), *Enterococcus gallinarum* (1 sequence), and various uncultured bacteria and hits from metagenomic projects (8 sequences). An alignment using a subset of these proteins highlights the strong conservation of this motif (Figure 4A), which was also validated by a search using Weblogo server (Crooks, et al. 2004) (Figure 4B). Together, these results reveal a conserved signature motif in first periplasmic loop of ArnT that also contains functional residues and is unique to both functionally characterized and predicted ArnT proteins.

The C-terminal domain of ArnT is required for function

Alignment of the C-terminal domain of ArnT proteins showed that this portion of the protein is the least conserved in primary amino acid sequence. However, secondary structure predictions reveal strong similarities between the C-terminal domains of *B. cenocepacia* and *S. enterica* ArnT proteins with a conserved predicted α/β structure (Figure 2). To investigate the role of the periplasmic C-terminal region on ArnT function, we created a series of constructs encoding

derivatives of *B. cenocepacia* ArnT with sequential C-terminal deletions. These deletions were designed to gradually remove the terminal domain ($ArnT_{\Delta 10}$, $ArnT_{\Delta 40}$, $ArnT_{\Delta 80}$, and $ArnT_{\Delta CT}$) (Figure 2). The functionality of the truncated proteins was assessed *in vivo* by determining their ability to restore PmB resistance in *B. cenocepacia* Δ*arnT*. To facilitate detection of ArnT, all these proteins also had a C-terminal FLAG-10xHis epitope. The incorporation of the double-epitope in the wild type ArnT does not interfere with the functionality of the enzyme, as judged by the full complementation of PmB resistance in Δ*arnT* expressing $ArnT_{FLAG\text{-}10xHis}$ (Figure 5A) (Tavares-Carreon, et al. 2015). However, all C-terminal truncations failed to restore PmB resistance in Δ*arnT*, including the shortest truncation ($ArnT_{\Delta 10}$) (Figure 5A).

To investigate the basis of the PmB-sensitive phenotype of the C-terminal truncation strains, we first determined if the mutant proteins were localized in membranes. Membrane fractions from these strains were analysed by immunoblot. $ArnT_{WT}$ migrated as a 66.3-kDa polypeptide, while $ArnT_{\Delta 10}$, $ArnT_{\Delta 40}$, $ArnT_{\Delta 80}$, and $ArnT_{\Delta CT}$ migrated as 64.9, 61.6, 56.8 and 52.3 kDa polypeptides, respectively (Figure 5B). These polypeptides were detectable in the membrane fraction indicating that the mutant proteins are localized in the membrane. Therefore, the functional defect of ArnT truncations is not due to membrane mislocalization.

We next investigated the effect of the C-terminal truncations on lipid A modification with L-Ara4N using MALDI-TOF mass spectrometry. The mass spectrum from the wild-type sample revealed the presence of L-Ara4N in the monophosphate and diphosphate forms (Figure 6A), which was absent in the Δ*arnT* lipid A spectrum (Figure 6B). Mass spectrometry analysis of lipid A from Δ*arnT* carrying $ArnT_{\Delta 10}$, $ArnT_{\Delta 40}$, $ArnT_{\Delta 80}$, and $ArnT_{\Delta CT}$ did not reveal any peaks corresponding to glycoforms containing L-Ara4N (Figure 6C-F). These results agree with the hypersensitivity to PmB of the strain expressing these truncated proteins (Figure 5A), confirming the C-terminal domain of ArnT is necessary for the lipid A modification by L-Ara4N.

We next asked whether deletion of the C-terminal domain of *S. enterica* ArnT (herein $ArnT_{Se}$) is functionally important. Two constructions were made resulting in the deletion of the complete C-terminal domain or the last 10 amino acids of $ArnT_{Se}$. The $ArnT_{Se}$ functionality was assayed by complementation of PmB resistance in the *E. coli* Δ*arnT* mutant (Tavares-Carreon, et al. 2015). The results indicated that none of these proteins restored PmB resistance, suggesting that the last 10 amino acids are also required for the *S. enterica* ArnT activity (Figure 7A). Immunoblot analysis of membrane fractions revealed a band migrating below the 63-kDa mass standard, which agrees with the predicted molecular mass for $ArnT_{Se}$ of 59-kDa. We could detect $ArnT_{Se\Delta 10}$ and $ArnT_{Se\Delta CT}$ as polypeptides with apparent masses of 57 and 47 kDa, respectively (Figure 7B). These results indicated that similar to $ArnT_{Bc}$, the truncation of the $ArnT_{Se}$ C-terminus results in a protein that is localized to the membrane but unable to confer PmB resistance, strongly suggesting that despite the differences in primary amino acid sequence, the C-terminal domain in both ArnT proteins is required for function.

An in silico *structural model of ArnT resembles the crystal structure of the PglB oligosaccharyltransferase*

To gain further insight about the ArnT structure we generated an *in silico* structural model of *B. cenocepacia* ArnT using I-Tasser (Roy, et al. 2010). Remarkably, among the best templates in the database automatically used by I-Tasser to generate the ArnT model, we found the crystal structure of the oligosaccharyltransferase PglB from *Campylobacter lari* (PDB entry: 3RCE) (Lizak, et al. 2011). Therefore, we refined the predicted ArnT model using the tertiary structure of PglB as the template. Structural comparison between PglB with the predicted ArnT model (I-

Tasser C-score of -0.59) revealed that despite their low sequence identity (~24%) both proteins share strong structural homology (Figure 8). The resulting tridimensional model of ArnT exhibited a similar architecture as PglB consisting primarily of two domains: an N terminal domain containing transmembrane helices connected by short loops and a hydrophilic C-terminal domain with a globular conformation. The predicted structural model of ArnT was supported by our previous topological results that identified ArnT residues by substituted cysteine scanning mutagenesis (Tavares-Carreon, et al. 2015). Further, examination of the C-terminal region reveals that the last 10 residues required for ArnT activity (Figure 5A) form a predicted anti-parallel β-sandwich structure, which is also present in PglB (Figure 8A-B). The β-sandwich contains mostly hydrophobic and some charged amino acid residues. To better characterize this region in ArnT we introduced alanine replacements in conserved and charged residues (Figure 9A). Alanine replacement of arginine-549, glutamic acid-554 and lysine-555 residues did not affect ArnT function, as determined by complementation of PmB resistance in *B. cenocepacia* Δ*arnT* (Figure 9B). However, alanine substitution of hydrophobic amino acids such as isoleucine-544 and a double mutant valine-551/isoleucine-552 resulted in proteins that could not restore PmB resistance (Figure 9C). To eliminate the possibility that the double epitope (FLAG-10xHis) located at the C-terminus could interfere with ArnT function, we introduced a stop codon before the FLAG-10xHis tag coding sequence in the double alanine replacement mutant valine-551/isoleucine-552. The $ArnT_{V551A/I552A}$ protein lacking the FLAG-10xHis epitope could not restore PmB resistance when expressed in Δ*arnT* (Figure 9C). The same result was obtained using the identical mutant protein containing FLAG-10xHis demonstrating that the C-terminal double tag was not involved in the phenotype observed for the double mutant V551-I552. The alanine replacement mutant in $ArnT_{V551A-I552A}$ was expressed at wild type levels and it appeared in the membrane fraction (Figure 9D). The predicted ArnT tridimensional structural model shows that isoleucine-544, valine-551, and isoleucine-552 face each other (Figure 8B). Thus, we propose that these residues are required for the overall structural stability of ArnT or proper folding of the C-terminal domain.

The C-terminal domain of ArnT is required for LPS binding

Elucidation of the tertiary structure of *C. lari* PglB revealed that both the transmembrane and the soluble periplasmic domains participate in substrate binding and catalysis (Lizak, et al. 2011). Also, the C-terminal part was found to pin the acceptor peptide to the active site in the X-ray structure (Lizak, et al. 2011). Based on structural comparison between the predicted ArnT model and the PglB structure we hypothesized that the pocket in ArnT that is comparable to the PglB peptide binding pocket is involved in binding the lipid A instead of a peptide. This region in the predicted ArnT model is the only one that has a pocket of a diameter such that could accommodate the lipid A molecule. To explore this hypothesis, we performed an *in vitro* binding assay using purified ArnT proteins and a resin-capture procedure previously described (Sestito, et al. 2014). ArnT was immobilized on the Ni^{2+}-NTA resin and then Alexa-conjugated LPS from *S. enterica* Serovar Minnesota was added. The Alexa-LPS-ArnT complex was eluted from resin and the fluorescence was measured. In contrast to parental ArnT, $ArnT_{\Delta 10}$ showed at least a 3-fold decrease in binding and virtually no binding was observed in $ArnT_{\Delta CT}$ (Figure 10A). However, the purified isolated C-terminal domain failed to bind LPS (Figure 10A), suggesting that the C-terminal domain alone is not sufficient for LPS binding. The levels of truncated protein were detected by immunoblotting (Figure 10B), indicating that all of them were expressed at wild type levels.

The reduction in the LPS binding for $ArnT_{\Delta CT}$ and $ArnT_{\Delta 10}$ could also reflect structural modifications imposed by the C-terminal truncations. We employed a direct approach using substituted cysteine accessibility by methoxypolyethylene glycol maleimide (PEG-mal) labelling to assess the ArnT topology in the mutant proteins. PEG-mal is a membrane-impermeable maleimide with a mass ~5 kDa that results in a clear size-shift of the label protein, which can be detected by immunoblot. PEG-mal modification using EDTA-permeabilized whole cells indicates that the cysteine has a periplasmic orientation. While PEG-mal modification under denaturing conditions (presence of SDS) suggests that the cysteine is within a transmembrane helix or faces the cytoplasm. We constructed cysteine replacements in asparagine-62, lysine-69, phenylalanine-247 and arginine-254; these residues are positioned in the periplasmic loops 1 and 2, respectively (Tavares-Carreon, et al. 2015). The cysteine replacement mutants were made in a cysteineless version of ArnT with the corresponding C-terminal deletion. PEGylation assays using the resulting proteins showed PEG-mal labelling of $ArnT_{N62C}$, $ArnT_{K69C}$, $ArnT_{F247C}$, and $ArnT_{R254C}$ under non-denaturing conditions (Figure 10C). This indicated that these residues are exposed to the periplasmic face as previously established (Tavares-Carreon, et al. 2015), demonstrating that the C-terminal deletions did not cause gross changes in the conformation of the membrane domain of ArnT.

Discussion

Glycosyltransferases catalyse the synthesis of glycoconjugates by transferring an activated sugar residue to an appropriate acceptor molecule. Most glycosyltransferases possess a DXD signature, which is typically found in the enzymes utilizing nucleotide-activated sugars as donor substrates (Ramakrishnan, et al. 2004). It is proposed that the negatively charged aspartic acid residues in the DXD motif and the phosphate group of the nucleotide group are held together by a divalent metal ion (Lairson, et al. 2008, Liu and Mushegian 2003, Wiggins and Munro 1998). A similar function could be envisaged for the DXD motif in ArnT and the phosphate group of Und-P-Ara4N. The GT-C glycosyltransferases proteins to which ArnT belongs present a relaxed version of the DXD motif. We show here that this relaxed motif is part of a larger periplasmic signature motif, $DEXRYAX_{(5)}MX_{(3)}GXWX_{(9)}YFEKPX_{(4)}W$, which is highly conserved in ArnT proteins from diverse proteobacterial species. Alanine replacement of the aspartic acid-39 residue in *B. cenocepacia* ArnT (aspartic acid-32 in *S. enterica* ArnT) resulted in a non-functional protein that lost the ability to modify lipid A with L-Ara4N. The functional importance of the aspartic acid in the DEX motif was established for several glycosyltransferases including the human PIG-M protein (Maeda, et al. 2001), *Mycobacterium smegmatis* PimE/EmbC proteins (Korkegian, et al. 2014, Morita, et al. 2006), and *Pyrococcus furiosus* AglB protein (Matsumoto, et al. 2013), underscoring the requirement for acidic residues in the active site GT-C glycosyltransferases, as previously suggested (Franco and Rigden 2003). However, the alanine replacement of the glutamic acid present in the DEX motif resulted in a non-functional *S. enterica* ArnT but did not affect the *B. cenocepacia* protein. The reasons for this difference are presently unclear.

Our previous work suggested that the ArnT C-terminal domain residing on the periplasm plays a role in ArnT function (Tavares-Carreon, et al. 2015), as expected for an enzyme using substrates only accessible in the periplasm (Raetz, et al. 2007, Trent, et al. 2001). However, the precise function of the C-terminal domain has not been investigated. We show here that that the last ten amino acids were required for the ability of ArnT in *B. cenocepacia* and *S. enterica* to modify lipid A with L-Ara4N and to confer PmB resistance. *In silico* modelling of ArnT

predicted that a structure that closely resembles that of oligosaccharyltransferases involved in protein N-glycosylation. In eukaryotes, protein N-glycosylation requires a multiprotein membrane complex (Knauer and Lehle 1999), from which STT3 is the catalytic subunit (Yan and Lennarz 2002). STT3 is homologous to the prokaryotic PglB and AglB proteins, which also catalyse N-glycosylation reactions (Feldman, et al. 2005, Igura, et al. 2008, Kelleher and Gilmore 2006, Knauer and Lehle 1999). ArnT, PglB, AglB and STT3 could not be aligned due to their very low sequence homology. However, remarkable similarities where found by comparing the established tri-dimensional structures of PglB and AlgB with an *in silico* ArnT structural model. Most notably, these enzymes show a transmembrane domain including a relatively large periplasmic loop, and a C-terminal globular domain exposed to the periplasmic space. Comparison between the tridimensional predicted structure of the C-terminal soluble domain of *B. cenocepacia* and the solved structure of the globular C-terminal domain of PglB, *Mycobacterium tuberculosis* EmbC, *Pyrococcus furiosus* AglB, and *Archaeoglobus fulgidus* AglB-S1 (Alderwick, et al. 2011, Igura, et al. 2008, Maita, et al. 2010, Matsumoto, et al. 2012) showed conserved features. These domains present a globular conformation mainly by α-helical domain surrounded by β-strands. The C-terminal region has been demonstrated as the peptide substrate binding domain for PglB, EmbC, *Corynebacterium glutamicum* RptA and *A. fulgidus* AglB-Long (Birch, et al. 2009, Korkegian, et al. 2014, Lizak, et al. , Matsumoto, et al. 2013). Remarkably, our *in vitro* binding assay reveals for the first time that the C-terminal domain of ArnT is involved in lipid A substrate binding. Thus, we propose that the C-terminal domain of ArnT is the lipid A binding domain. However, the isolated C-terminal domain of ArnT expressed without the transmembrane region was unable to bind lipid A, a finding also observed with the isolated C-terminal domain in N-glycosylation proteins, which is unable to bind the acceptor peptide (Korkegian, et al. 2014, Maita, et al. 2010, Matsumoto, et al. 2012), indicating that C-terminal domain is not sufficient for substrate binding. Substrate binding in these enzymes may require the transmembrane region, especially one or more periplasmic loops. We also found that the two β-strands located in the last ten amino acids of the C-terminal domain of ArnT are functionally important, likely by contributing to the structural integrity of the C-terminal fold. Comparison of the modelled C-terminal domain of ArnT with the solved structures C-terminal domain of *P. furiosus* AglB and *C. lari* PglB proteins also reveal β-strands-like folding which are structurally or functionally essential for these proteins (Igura, et al. 2008, Maita, et al. 2010).

The discovery of a new signature motif and a similar feature of the C-terminal domain in ArnT with the distantly related oligosaccharyltransferases PglB and AlgB, highlights a common fold architecture for these enzymes irrespective of the protein or lipid nature of their specific substrates. We speculate that the bacterial and eukaryotic glycosyltransferases with similar functions may have evolved in parallel or may be derived from an ancient common ancestor, either by gene exchanges between organisms or by convergent evolution. Collectively, our results strongly suggest that the predicted tertiary structure of ArnT share properties in common to the solved structures of bacterial N-glycosylation oligosaccharyltransferases, revealing an unexpected conservation of evolutionary divergent enzymes that glycosylate proteins and lipids in the extracytoplasmic compartment.

Materials and methods

Bacterial strains and growth conditions

Bacterial strains and plasmids are listed in Table I. Growth medium was supplemented as needed with the following chemicals (final concentrations): ampicillin (100 μg ml^{-1}), trimethoprim (100

μg ml^{-1}), gentamicin (25 μg ml^{-1}), PmB (10 μg ml^{-1}), L-arabinose 0.2%; w/v), and L-rhamnose (0.2%; w/v). Bacteria were grown aerobically at 37 °C in Luria-Bertani (LB) medium (Difco). For recombinant protein expression, bacteria were grown overnight in 5 mL of LB, diluted to A_{600} of 0.1, and incubated at 37°C for 2 h until cultures reached A_{600} of 0.5, at which point L-arabinose or L-rhamnose was added and cells incubated for 3 h at 37°C.

DNA methods and plasmid construction

Plasmid pFT1 encoding the *B. cenocepacia arnT*$_{\text{FLAG-10xHis}}$ was used as template (Tavares-Carreon, et al. 2015). Gene deletions eliminating 10, 40, 80, and 121 amino acid residues from the ArnT C-terminus (ArnT$_{\Delta 10}$, ArnT$_{\Delta 40}$, ArnT$_{\Delta 80}$, and ArnT$_{\Delta CT}$, respectively) were constructed by PCR amplification using the forward primer 6269 containing an *Eco*RI restriction site, which anneals upstream of the *arnT* coding sequence, and the reverse primers 6424, 6425, 6426, and 6739 that anneal downstream of *arnT* and contain a *Sal*I restriction site. PCR products were digested with *Eco*RI and *Sal*I and ligated into pFT1 also digested with the same enzymes.

pFT143 encoding the *S. enterica arnT*$_{\text{Se-FLAG-10xHis}}$ gene (Tavares-Carreon, et al. 2015) was used as a template to delete the whole C-terminal domain of ArnT (with 6475 and 7296 primers), and the last 10 C-terminal residues (with 6475 and 7295). PRC products were digested with *Eco*RI and *Sal*I and ligated into the same sites in pFT143. Novel amino acid codons for amino acid replacements in both ArnT proteins were introduced in the respective genes by PCR using primers containing the desired mutations, pFT1 or pFT143 as templates, and the *Pfu* Turbo AD polymerase (Agilent Technologies). Upon completion PCR reactions were incubated with *Dpn*I overnight at 37°C to digest parental plasmid DNA. The resulting DNA was introduced into *E. coli* DH5α by $CaCl_2$ transformation. Plasmids were sequenced with primers 252, 258 and 6385 for *B. cenocepacia arnT* or 7179 for *S. enterica arnT* (Table II) to confirm the presence of the desired mutations. For the cloning of the C-terminal domain of *B. cenocepacia arnT*, primers 7380 and 6270 with *Nco*I and *Sal*I restriction sites respectively were used for a PCR reaction. The amplicon was digested with *Nco*I and *Sal*I and ligated into pFT1 equally digested, giving rise to pFT232.

Complementation of ArnT function in B. cenocepacia

Genes expressing the ArnT replacement mutant proteins were subcloned into the *Nde*I-*Hin*dIII site of pSCrhaB2, and the recombinant constructs introduced into the PmB-sensitive Δ*arnT-arnBC*$^{+}$ *lptG*$_{\text{D31H}}$ (Δ*arnT*) suppressor strain by triparental mating (Craig, et al. 1989, Figurski and Helinski 1979). The ability of each construct to restore PmB resistance to wild type levels denoted ArnT function. Strains were grown overnight with 0.4% (w/v) rhamnose (to induce protein expression), diluted to 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5} and 4 μl of each dilution plated onto LB plates supplemented with 0.4% rhamnose and 10 μg ml^{-1} PmB.

Membrane preparation and immunoblotting

Bacterial pellets were resuspended in 50 mM Tris-HCl, pH 8.0, with protease inhibitors and lysed at 10, 000 PSI using a cell disruptor (Constant Systems, Kennesaw, GA). Cell debris was pelleted by centrifugation at 15,000 × *g* for 15 min at 4°C, and the clear supernatant was centrifuged at 40,000 × *g* for 30min at 4°C. The pellet containing total membranes was suspended 50 mM Tris-HCl, pH 8.0. Protein concentration was determined by the Bradford assay (Bio-Rad). Proteins were separated by SDS-PAGE on a 12% gel and transferred to a nitrocellulose membrane. Immunoblots were probed with anti-FLAG (Sigma) mouse monoclonal

antibodies. Secondary anti- mouse Alexa fluor-680 IgG antibody (Invitrogen) was used to detect fluorescence using an Odyssey infrared imaging system (LI-COR Biosciences).

Lipid A isolation and mass spectrometry

Lipid A L-Ara4N modifications were identified by mass spectrometry (MS) analysis using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF). Cells were grown overnight in LB medium (100 ml) supplemented with 50 μg/ml trimethoprim and 0.4 % rhamnose. Cultures were centrifuged (10,000 × *g*), washed twice with phosphate buffer (10 mM Na_2HPO_4, 1.7 mM KH_2PO_4) and freeze-dried. Lipid A was extracted from lyophilized cells and desalted as previously described (El Hamidi, et al. 2005). Ten milligrams of the lyophilized bacterial cells were resuspended in 400 μl of a mixture of isobutyric acid: 1 M Ammonium hydroxide (5:3, v/v) and were kept for 2 h at 100°C with vortexing every 15 min., then centrifuged at 2000 × *g* for 15 min and the supernatant was mixed with an equal volume of water and lyophilized. The sample was washed with methanol twice and then lipid A was solubilized in 80 μl of chloroform: methanol: water (3:1.5:0.25, v/v). Lipid A suspension was desalted with a few grains of ion-exchange resin (Dowex 50W-X8; H^+). 2 μl aliquot of desalted lipid A suspension was loaded on polished steel target, air dried and covered by 1 μl of 2,5-dihydroxybenzoic acid matrix (DHB) (Sigma Chemical) dissolved in 0.1 M citric acid aqueous solution and allowed to dry. The target was inserted in a Bruker Autoflex MALDI-TOF spectrometer. Data acquisition and analysis were performed using the Flex Analysis software.

Expression and purification of ArnT

$ArnT_{FLAG\text{-}10xHis}$, $ArnT_{\Delta CT\text{-}FLAG\text{-}10xHis}$, $ArnT_{\Delta 10\text{-}FLAG\text{-}10xHis}$ and C-terminal domain-FLAG-10xHis recombinant proteins were purified as previously reported (Ruan and Valvano 2013). Proteins were overexpressed in *E. coli* DH5α cells in a 1 litre of LB. Cells were grown to A_{600} of 0.6 at 37°C and induced with 0.2% (w/v) arabinose for 4 h. Cells were harvested by centrifugation and pellets were stored at -20°C before membrane preparation. All subsequent steps were carried out at 4°C. Membrane preparation was solubilized in Equilibration buffer (50 mM phosphate buffer pH 8.0, 150 mM NaCl, 10% (v/v) glycerol, and 1% (w/v) *N*-dodecyl-β-D-maltopyranoside (Sigma-Aldrich) for 4 h. All purification buffers contained *N*-dodecyl-β-D-maltopyranoside. ArnT was purified from membranes on a nickel-nitrilotriacetic acid affinity column. The column was washed twice in Equilibration buffer with 50 mM of imidazole and the protein was eluted with Equilibration buffer containing 250 mM imidazole. Elution fractions were monitored by 12% polyacrylamide SDS-PAGE. The pooled fractions containing purified protein were dialyzed twice against 200 volumes of Equilibration buffer and finally concentrated in a Vivaspin 15R column (10, 000 molecular weight cut off, Sartorious). Protein concentration was determined by Bradford assay with bovine serum albumin as standard.

LPS binding by resin capture method

The *in vitro* LPS binding assay was based on a previous protocol with a few modifications (Sestito, et al. 2014). Briefly, purified ArnT protein (200 μg) was incubated in 0.5 ml of Phosphate/NaCl buffer (50 mM phosphate buffer pH 8.0, 150 mM NaCl) with 0.06 ml of HisLink protein purification resin (Promega) for 90 min on a rotary shaker at 4 °C. Unbound protein was removed by decanting the resin-bound ArnT. Resin-bound was incubated in 0.5 ml of Phosphate/NaCl buffer and 10 μg of Alexa-LPS conjugated from *Salmonella* serovar Minnesota (Molecular Probes) at room temperature in the dark on rotary shaking for further 90

min. The resin was washed twice with 0.5 ml of Phosphate/NaCl buffer. The protein-LPS complex was eluted with 0.5 ml of Phosphate/NaCl buffer containing 300 mM imidazole. The eluted (0.5 ml) was divided in 100 µl and placed in a Microplate, white flat-bottomed (Greiner bio-one) and fluorescence was measured in a Fluoroskan Ascent FL Microplate Fluorometer (Thermo Labsystems) (λ_{ex}=485nm, λ_{em}=527).

Sulfhydryl labelling

Plasmids containing $ArnT_{FLAG\text{-}10xHis}$ cysteine substitutions with ΔC-terminal or Δ10 amino acids deletions were introduced into *E. coli* DH5α by $CaCl_2$ transformation. Cultures were grown and sulfhydryl labelling was performed as previously described (Tavares-Carreon, et al. 2015). Briefly, cell pellets were resuspended in 0.3 ml of HEPES/$MgCl_2$ buffer and divided in three portions: 0.1 ml of cell suspension was incubated with 1 mM methoxypolyethylene glycol maleimide (PEG-Mal) at room temperature for 1 h with 50 mM EDTA; 0.1 ml was treated with 2% SDS prior to labelling, and 0.1 ml was incubated with buffer as control. The reaction was stopped with 45 mM of dithiothreitol for 10 min and mixed with 3X dye buffer (50 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, and 0.1% Bromophenol blue), and incubated for 30 min at 45°C. Mild denaturation conditions were used to optimize ArnT detection, as these conditions work best for membrane proteins (Ruan and Valvano 2013).

Bioinformatics and in silico modelling of ArnT

Homology searches were performed with BLASTP and PSI-BLAST. Alignments of amino acid sequences were performed by ClustalW (Chenna, et al. 2003). ArnT was classified within GT83 glycosyltransferase family using the CAZy database (*c*arbohydrate-*a*ctive en*z*ymes) (www.cazy.org) (Lombard, et al. 2014). Graphical representation of the signature motif in ArnT homologues were made by Weblogo (http://weblogo.berkeley.edu) (Crooks, et al. 2004). Motif search was done on the Swiss-Prot and TrEMBL databases using the ScanProsite tool (de Castro, et al. 2006). A fasta file containing all the sequences obtained from this analysis is available from the authors upon request. For *in silico* modelling the *B. cenocepacia* ArnT amino acid sequence was submitted to the I-TASSER server with default parameters (Roy, et al. 2010). Figures were prepared with the PyMOL molecular graphics system, version 1.7.2.3 (Schrödinger, LLC, Portland, OR).

Supplementary data

Supplementary Figures S1 and S2.

Funding

This work was supported by grants from the Canadian Institutes of Health Research and Cystic Fibrosis Canada (to M.A.V.).

Acknowledgments

We thank Dr. O. El-Halfawy for technical assistance.

Conflict of interest

None declared.

Abbreviations

GT-C, glycosyltransferase C superfamily; L-Ara4N, 4-amino-4-*deoxy*-L-arabinose; LPS, lipopolysaccharide; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; MS, mass; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; Und-P, undecaprenyl-phosphate

References

Alderwick, LJ, Lloyd, GS, Ghadbane, H, May, JW, Bhatt, A, Eggeling, L, Futterer, K, and Besra, GS (2011) The C-terminal domain of the Arabinosyltransferase *Mycobacterium tuberculosis* EmbC is a lectin-like carbohydrate binding module. *PLoS Pathog* 7, e1001299.

Belanger, AE, Besra, GS, Ford, ME, Mikusova, K, Belisle, JT, Brennan, PJ, and Inamine, JM (1996) The embAB genes of *Mycobacterium avium* encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *Proc Natl Acad Sci U S A* 93, 11919-11924.

Birch, HL, Alderwick, LJ, Rittmann, D, Krumbach, K, Etterich, H, Grzegorzewicz, A, McNeil, MR, Eggeling, L, and Besra, GS (2009) Identification of a terminal rhamnopyranosyltransferase (RptA) involved in *Corynebacterium glutamicum* cell wall biosynthesis. *J Bacteriol* 191, 4879-4887.

Boll, M, Radziejewska-Lebrecht, J, Warth, C, Krajewska-Pietrasik, D, and Mayer, H (1994) 4-Amino-4-deoxy-L-arabinose in LPS of enterobacterial R-mutants and its possible role for their polymyxin reactivity. *FEMS Immunol Med Microbiol* 8, 329-341.

Breton, C, and Imberty, A (1999) Structure/function studies of glycosyltransferases. *Curr Opin Struct Biol* 9, 563-571.

Bretscher, LE, Morrell, MT, Funk, AL, and Klug, CS (2006) Purification and characterization of the L-Ara4N transferase protein ArnT from *Salmonella typhimurium*. *Protein Expr Purif* 46, 33-39.

Cardona, ST, and Valvano, MA (2005) An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* 54, 219-228.

Chenna, R, Sugawara, H, Koike, T, Lopez, R, Gibson, TJ, Higgins, DG, and Thompson, JD (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31, 3497-3500.

Craig, FF, Coote, JG, Parton, R, Freer, JH, and Gilmour, NJ (1989) A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. *J Gen Microbiol* 135, 2885-2890.

Crooks, GE, Hon, G, Chandonia, JM, and Brenner, SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14, 1188-1190.

de Castro, E, Sigrist, CJ, Gattiker, A, Bulliard, V, Langendijk-Genevaux, PS, Gasteiger, E, Bairoch, A, and Hulo, N (2006) ScanProsite: detection of PROSITE signature matches and

ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 34, W362-365.

El Hamidi, A, Tirsoaga, A, Novikov, A, Hussein, A, and Caroff, M (2005) Microextraction of bacterial lipid A: easy and rapid method for mass spectrometric characterization. *J Lipid Res* 46, 1773-1778.

Ernst, RK, Yi, EC, Guo, L, Lim, KB, Burns, JL, Hackett, M, and Miller, SI (1999) Specific lipopolysaccharide found in cystic fibrosis airway *Pseudomonas aeruginosa*. *Science* 286, 1561-1565.

Feldman, MF, Wacker, M, Hernandez, M, Hitchen, PG, Marolda, CL, Kowarik, M, Morris, HR, Dell, A, Valvano, MA, and Aebi, M (2005) Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc Natl Acad Sci U S A* 102, 3016-3021.

Figurski, DH, and Helinski, DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *Proc Natl Acad Sci U S A* 76, 1648-1652.

Franco, OL, and Rigden, DJ (2003) Fold recognition analysis of glycosyltransferase families: further members of structural superfamilies. *Glycobiology* 13, 707-712.

Gunn, JS, Lim, KB, Krueger, J, Kim, K, Guo, L, Hackett, M, and Miller, SI (1998) PmrA-PmrB-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. *Mol Microbiol* 27, 1171-1182.

Guo, L, Lim, KB, Gunn, JS, Bainbridge, B, Darveau, RP, Hackett, M, and Miller, SI (1997) Regulation of lipid A modifications by *Salmonella typhimurium* virulence genes phoP-phoQ. *Science* 276, 250-253.

Guzman, LM, Belin, D, Carson, MJ, and Beckwith, J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177, 4121-4130.

Hamad, MA, Di Lorenzo, F, Molinaro, A, and Valvano, MA (2012) Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. *Mol Microbiol* 85, 962-974.

Igura, M, Maita, N, Kamishikiryo, J, Yamada, M, Obita, T, Maenaka, K, and Kohda, D (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* 27, 234-243.

Impellitteri, NA, Merten, JA, Bretscher, LE, and Klug, CS (2010) Identification of a functionally important loop in *Salmonella typhimurium* ArnT. *Biochemistry* 49, 29-35.

Kaca, W, Radziejewska-Lebrecht, J, and Bhat, UR (1990) Effect of polymyxins on the lipopolysaccharide-defective mutants of *Proteus mirabilis*. *Microbios* 61, 23-32.

Kelleher, DJ, and Gilmore, R (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47R-62R.

Kim, SH, Jia, W, Parreira, VR, Bishop, RE, and Gyles, CL (2006) Phosphoethanolamine substitution in the lipid A of *Escherichia coli* O157 : H7 and its association with PmrC. *Microbiology* 152, 657-666.

Knauer, R, and Lehle, L (1999) The oligosaccharyltransferase complex from yeast. *Biochim Biophys Acta* 1426, 259-273.

Korkegian, A, Roberts, DM, Blair, R, and Parish, T (2014) Mutations in the essential arabinosyltransferase EmbC lead to alterations in *Mycobacterium tuberculosis* lipoarabinomannan. *J Biol Chem* 289, 35172-35181.

Korres, H, and Verma, NK (2006) Identification of essential loops and residues of glucosyltransferase V (GtrV) of *Shigella flexneri*. *Mol Membr Biol* 23, 407-419.

Kowarik, M, Young, NM, Numao, S, Schulz, BL, Hug, I, Callewaert, N, Mills, DC, Watson, DC, Hernandez, M, Kelly, JF, et al. (2006) Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* 25, 1957-1966.

Kus, JV, Kelly, J, Tessier, L, Harvey, H, Cvitkovitch, DG, and Burrows, LL (2008) Modification of *Pseudomonas aeruginosa* Pa5196 type IV Pilins at multiple sites with D-Araf by a novel GT-C family Arabinosyltransferase, TfpW. *J Bacteriol* 190, 7464-7478.

Lairson, LL, Henrissat, B, Davies, GJ, and Withers, SG (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* 77, 521-555.

Liu, J, and Mushegian, A (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci* 12, 1418-1431.

Lizak, C, Gerber, S, Numao, S, Aebi, M, and Locher, KP X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350-355.

Lizak, C, Gerber, S, Numao, S, Aebi, M, and Locher, KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350-355.

Lombard, V, Golaconda Ramulu, H, Drula, E, Coutinho, PM, and Henrissat, B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42, D490-495.

Loutet, SA, and Valvano, MA (2011) Extreme antimicrobial peptide and polymyxin B resistance in the genus *Burkholderia*. *Front Cell Infect Microbiol* 1, 6.

Maeda, Y, Watanabe, R, Harris, CL, Hong, Y, Ohishi, K, Kinoshita, K, and Kinoshita, T (2001) PIG-M transfers the first mannose to glycosylphosphatidylinositol on the lumenal side of the ER. *EMBO J* 20, 250-261.

Mahenthiralingam, E, Baldwin, A, and Dowson, CG (2008) *Burkholderia cepacia* complex bacteria: opportunistic pathogens with important natural biology. *J Appl Microbiol* 104, 1539-1551.

Mahenthiralingam, E, Coenye, T, Chung, JW, Speert, DP, Govan, JR, Taylor, P, and Vandamme, P (2000) Diagnostically and experimentally useful panel of strains from the *Burkholderia cepacia* complex. *J Clin Microbiol* 38, 910-913.

Maita, N, Nyirenda, J, Igura, M, Kamishikiryo, J, and Kohda, D (2010) Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J Biol Chem* 285, 4941-4950.

Marr, N, Tirsoaga, A, Blanot, D, Fernandez, R, and Caroff, M (2008) Glucosamine found as a substituent of both phosphate groups in *Bordetella* lipid A backbones: role of a BvgAS-activated ArnT ortholog. *J Bacteriol* 190, 4281-4290.

Matsumoto, S, Igura, M, Nyirenda, J, Matsumoto, M, Yuzawa, S, Noda, N, Inagaki, F, and Kohda, D (2012) Crystal structure of the C-terminal globular domain of oligosaccharyltransferase from *Archaeoglobus fulgidus* at 1.75 A resolution. *Biochemistry* 51, 4157-4166.

Matsumoto, S, Shimada, A, Nyirenda, J, Igura, M, Kawano, Y, and Kohda, D (2013) Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc Natl Acad Sci U S A* 110, 17868-17873.

Miller, SI, Ernst, RK, and Bader, MW (2005) LPS, TLR4 and infectious disease diversity. *Nat Rev Microbiol* 3, 36-46.

Mogensen, TH (2009) Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin Microbiol Rev* 22, 240-273.

Molinaro, A, Holst, O, Di Lorenzo, F, Callaghan, M, Nurisso, A, D'Errico, G, Zamyatina, A, Peri, F, Berisio, R, Jerala, R, et al. (2015) Chemistry of lipid A: at the heart of innate immunity. *Chemistry* 21, 500-519.

Morita, YS, Sena, CB, Waller, RF, Kurokawa, K, Sernee, MF, Nakatani, F, Haites, RE, Billman-Jacobe, H, McConville, MJ, Maeda, Y, et al. (2006) PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria. *J Biol Chem* 281, 25143-25155.

Nair, A, Korres, H, and Verma, NK (2011) Topological characterisation and identification of critical domains within glucosyltransferase IV (GtrIV) of *Shigella flexneri*. *BMC Biochem* 12, 67.

Needham, BD, and Trent, MS (2013) Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nat Rev Microbiol* 11, 467-481.

Nikaido, H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67, 593-656.

Ortega, XP, Cardona, ST, Brown, AR, Loutet, SA, Flannagan, RS, Campopiano, DJ, Govan, JR, and Valvano, MA (2007) A putative gene cluster for aminoarabinose biosynthesis is essential for *Burkholderia cenocepacia* viability. *J Bacteriol* 189, 3639-3644.

Perez, JM, McGarry, MA, Marolda, CL, and Valvano, MA (2008) Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. *Mol Microbiol* 70, 1424-1440.

Perrin-Tricaud, C, Rutschmann, C, and Hennet, T (2011) Identification of domains and amino acids essential to the collagen galactosyltransferase activity of GLT25D1. *PLoS One* 6, e29390.

Raetz, CR, Reynolds, CM, Trent, MS, and Bishop, RE (2007) Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* 76, 295-329.

Ramakrishnan, B, Boeggeman, E, Ramasamy, V, and Qasba, PK (2004) Structure and catalytic cycle of beta-1,4-galactosyltransferase. *Curr Opin Struct Biol* 14, 593-600.

Ramiscal, RR, Tang, SS, Korres, H, and Verma, NK (2010) Structural and functional divergence of the newly identified GtrIc from its Gtr family of conserved *Shigella flexneri* serotype-converting glucosyltransferases. *Mol Membr Biol* 27, 114-122.

Rolin, O, Muse, SJ, Safi, C, Elahi, S, Gerdts, V, Hittle, LE, Ernst, RK, Harvill, ET, and Preston, A (2014) Enzymatic modification of lipid A by ArnT protects *Bordetella bronchiseptica* against cationic peptides and is required for transmission. *Infect Immun* 82, 491-499.

Roy, A, Kucukural, A, and Zhang, Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5, 725-738.

Ruan, X, and Valvano, MA (2013) In vitro O-antigen ligase assay. *Methods Mol Biol* 1022, 185-197.

Sestito, SE, Sperandeo, P, Santambrogio, C, Ciaramelli, C, Calabrese, V, Rovati, GE, Zambelloni, L, Grandori, R, Polissi, A, and Peri, F (2014) Functional characterization of *E. coli* LptC: interaction with LPS and a synthetic ligand. *Chembiochem* 15, 734-742.

Silipo, A, Molinaro, A, Cescutti, P, Bedini, E, Rizzo, R, Parrilli, M, and Lanzetta, R (2005) Complete structural characterization of the lipid A fraction of a clinical strain of *B. cepacia* genomovar I lipopolysaccharide. *Glycobiology* 15, 561-570.

Strahl-Bolsinger, S, Immervoll, T, Deutzmann, R, and Tanner, W (1993) PMT1, the gene for a key enzyme of protein O-glycosylation in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 90, 8164-8168.

Tavares-Carreon, F, Patel, KB, and Valvano, MA (2015) *Burkholderia cenocepacia* and *Salmonella enterica* ArnT proteins that transfer 4-amino-4-deoxy-L-arabinose to lipopolysaccharide share membrane topology and functional amino acids. *Sci Rep* 5, 10773.

Telenti, A, Philipp, WJ, Sreevatsan, S, Bernasconi, C, Stockbauer, KE, Wieles, B, Musser, JM, and Jacobs, WR, Jr. (1997) The emb operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. *Nat Med* 3, 567-570.

Trent, MS, Ribeiro, AA, Doerrler, WT, Lin, S, Cotter, RJ, and Raetz, CRH (2001) Accumulation of a polyisoprene-linked amino sugar in polymyxin-resistant *Salmonella typhimurium* and *Escherichia coli*. Structural characterization and transfer to lipid A in the periplasm. *J. Biol. Chem.* 276, 43132-43144.

Trent, MS, Ribeiro, AA, Lin, S, Cotter, RJ, and Raetz, CR (2001) An inner membrane enzyme in *Salmonella* and *Escherichia coli* that transfers 4-amino-4-deoxy-L-arabinose to lipid A: induction on polymyxin-resistant mutants and role of a novel lipid-linked donor. *J Biol Chem* 276, 43122-43131.

Vandamme, P, and Dawyndt, P (2011) Classification and identification of the *Burkholderia cepacia* complex: Past, present and future. *Syst Appl Microbiol* 34, 87-95.

Wang, X, Ribeiro, AA, Guan, Z, and Raetz, CR (2009) Identification of undecaprenyl phosphate-beta-D-galactosamine in *Francisella novicida* and its function in lipid A modification. *Biochemistry* 48, 1162-1172.

Wiggins, CA, and Munro, S (1998) Activity of the yeast MNN1 α-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc Natl Acad Sci U S A* 95, 7945-7950.

Yan, Q, and Lennarz, WJ (2002) Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. *J Biol Chem* 277, 47692-47700.

Zhou, Z, Ribeiro, AA, and Raetz, CR (2000) High-resolution NMR spectroscopy of lipid A molecules containing 4-amino-4-deoxy-L-arabinose and phosphoethanolamine substituents. Different attachment sites on lipid A molecules from NH4VO3-treated *Escherichia coli* versus *kdsA* mutants of *Salmonella typhimurium*. *J Biol Chem* 275, 13542-13551.

TABLE I. Strains and plasmids used in this study.

Strain or Plasmids	Relevant Properties	Source or Reference
Strains		
E. coli		
DH5α	F$^-$ ϕ80*lacZ*M15 *endA recA hsdR*(r^-_K m^-_K) *supE thi gyrA relA* Δ(*lacZYA-argF*)*U169*	Laboratory stock
Δ*arnT*	BL21(DE3) Δ*arnT*	(Impellitteri, et al. 2010)
B. cenocepacia		
K56-2	Parental strain, clinical isolate of the ET12 clone	BCRRC[a] (Mahenthiralingam, et al. 2000)
MH55	Δ*arnT-arnBC*$^+$ $lptG_{D31H}$	(Hamad, et al. 2012)
Plasmids		
pBAD24	Expression vector inducible with arabinose, for C-terminal FLAG-10xHis fusions, Ap^R	(Guzman, et al. 1995)
pMH494	pSCrhaB2-$arnT_{Se}$	Hamad and Valvano, *in preparation.*
pRK2013	Helper plasmid used for bacterial conjugation; Km^R	(Figurski and Helinski 1979)
pSCrhaB2	Expression vector inducible with rhamnose. Broad host range replicative vector; Tp^R	(Cardona and Valvano 2005)
pFT1	pBAD expressing $ArnT_{FLAG\text{-}10xHis}$	(Tavares-Carreon, et al. 2015)
pFT3	pSCrhaB2, $ArnT_{FLAG\text{-}10xHis}$	(Tavares-Carreon, et al. 2015)
pFT9	pFT1, $ArnT_{\Delta CT}$	This study
pFT10	pFT1, $ArnT_{\Delta 80}$	This study
pFT11	pFT1, $ArnT_{\Delta 40}$	This study
pFT32	pFT1, $ArnT_{\Delta 10}$	This study
pFT17	pFT3, $ArnT_{\Delta CT}$	This study
pFT18	pFT3, $ArnT_{\Delta 80}$	This study
pFT19	pFT3, $ArnT_{\Delta 40}$	This study
pFT35	pFT3, $ArnT_{\Delta 10}$	This study
pFT124	pFT1, $ArnT_{E554A}$	This study
pFT129	pFT3, $ArnT_{E554A}$	This study
pFT143	pBAD, $ArnT_{Se\text{-}P547A\text{-}FLAG\text{-}10xHis}$	(Tavares-Carreon, et al. 2015)
pFT147	pFT1, $ArnT_{R549A}$	This study
pFT148	pFT1, $ArnT_{K555A}$	This study
pFT151	pFT3, $ArnT_{R549A}$	This study
pFT152	pFT3, $ArnT_{K555A}$	This study
pFT165	pFT1, $ArnT_{D39A}$	This study
pFT166	pFT1, $ArnT_{E40A}$	This study
pFT167	pFT3, $ArnT_{D39A}$	This study
pFT168	pFT3, $ArnT_{E40A}$	This study
pFT169	pFT143, $ArnT_{Se\text{-}D32A}$	This study
pFT170	pFT143, $ArnT_{Se\text{-}E33A}$	This study
pFT175	pBAD, $ArnT_{Se\Delta CT}$	This study

pFT176	pBAD, $ArnT_{Se\Delta 10}$	This study
pFT177	pFT1, $ArnT_{V551A\text{-}I552A}$	This study
pFT179	pFT3, $ArnT_{V551A\text{-}I552A}$	This study
pFT193	pFT1, $ArnT_{I544A}$	This study
pFT194	pFT3, $ArnT_{I544A}$	This study
pFT222	pFT1, $ArnT_{V551A\text{-}I552A\text{-}STOP}$	This study
pFT223	pFT3, $ArnT_{V551A\text{-}I552A\text{-}STOP}$	This study
pFT224	pFT38, $ArnT_{N62C\Delta CT}$	This study
pFT225	pFT77, $ArnT_{K69C\Delta CT}$	This study
pFT226	pFT79, $ArnT_{F247C\Delta CT}$	This study
pFT227	pFT55, $ArnT_{R254C\Delta CT}$	This study
pFT228	pFT38, $ArnT_{N62C\Delta 10}$	This study
pFT229	pFT77, $ArnT_{K69C\Delta 10}$	This study
pFT230	pFT79, $ArnT_{F247C\Delta 10}$	This study
pFT231	pFT55, $ArnT_{R254C\Delta 10}$	This study
pFT232	C-terminal domain of $ArnT_{FLAG\text{-}10xHis}$	This study

[a]*B. cepacia* Research and Referral Repository for Canadian CF Clinics.

TABLE II. Primers used in this study.

Primer	DNA Sequence	Restriction site	Purpose
252	5'-GATTAGCGGATCCTACCTGA	None	pBAD24 forward
258	5'-GACCGCTTCTGCGTTCTGAT	None	pBAD24 reverse
6269	5'-aaagaattcATGAACGATACGCCGTCGAGGC	*Eco*RI	$ArnT_{Bc}$ forward
6270	5'-aaagtcgactcCGATTGCGGTTTCTCGACGATC	*Sal*I	$ArnT_{Bc}$ reverse
6385	5'-CAGCGCAACCCCGAGTTCTTCAAC	None	$ArnT_{Bc}$, internal primer
6424	5'- aaagtcgactcCCCGGTGCCGCCGATCGTGCC	*Sal*I	$ArnT_{Bc\Delta CT}$ reverse
6425	5'- aaagtcgactcCGGGAACGTGTGGTCGAGCAT	*Sal*I	$ArnT_{Bc\Delta 80}$ reverse
6426	5'- aaagtcgactcCTGCTTCCAGCGCGTGATCCA	*Sal*I	$ArnT_{Bc\Delta 40}$ reverse
6475	5'-aaagaattcATGAAATCGATACGCTATTATC	*Eco*RI	$ArnT_{Se}$ forward
6739	5'- aaagtcgactcGTTGTCGCGCGCGATCACGCG	*Sal*I	$ArnT_{Bc\Delta 10}$ reverse
7179	5'- GTGGTTGTCCTTCCCTGGGCGATCG	None	$ArnT_{Se}$, internal primer
7295	5'- aaagtcgactcCCCCTGATAATCAATATTGTC	*Sal*I	$ArnT_{Se\Delta 10}$ reverse
7296	5'- aaagtcgactcGGGGATGGAAAAGCCGAACAG	*Sal*I	$ArnT_{Se\Delta CT}$ reverse
7380	5'-aaaaccatgggcGACGAGTTCGGCCGCTACAGC	*Nco*I	$ArnT_{Bc}$ forward

Restriction sites are underlined.

Figures

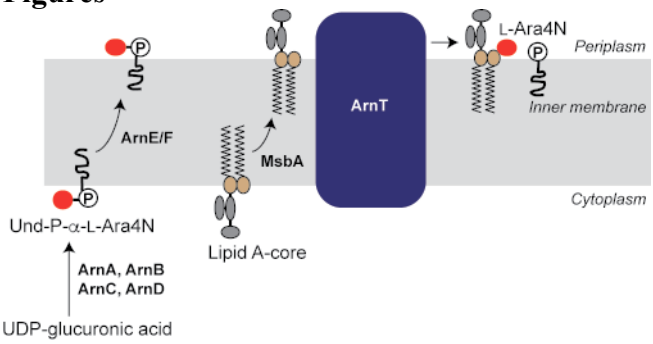

Fig 1. Diagram illustrating the synthesis and incorporation of 4-amino-4-*deoxy*-L-arabinose (L-Ara4N) into the lipid A. Soluble UDP-glucuronic acid is converted to Und-P-α-L-Ara4N by the cytoplasmic proteins ArnA, ArnB, ArnC, and ArnD on the cytoplasmic face on the inner membrane (Raetz, et al. 2007). The flippase ArnE/ArnF transports Und-P-α-L-Ara4N to the periplasmic face of the inner membrane. The lipid A-core is exported to the periplasmic face of the inner membrane by the MsbA flippase. The integral membrane protein ArnT transfers the L-Ara4N moiety to the lipid A-core.

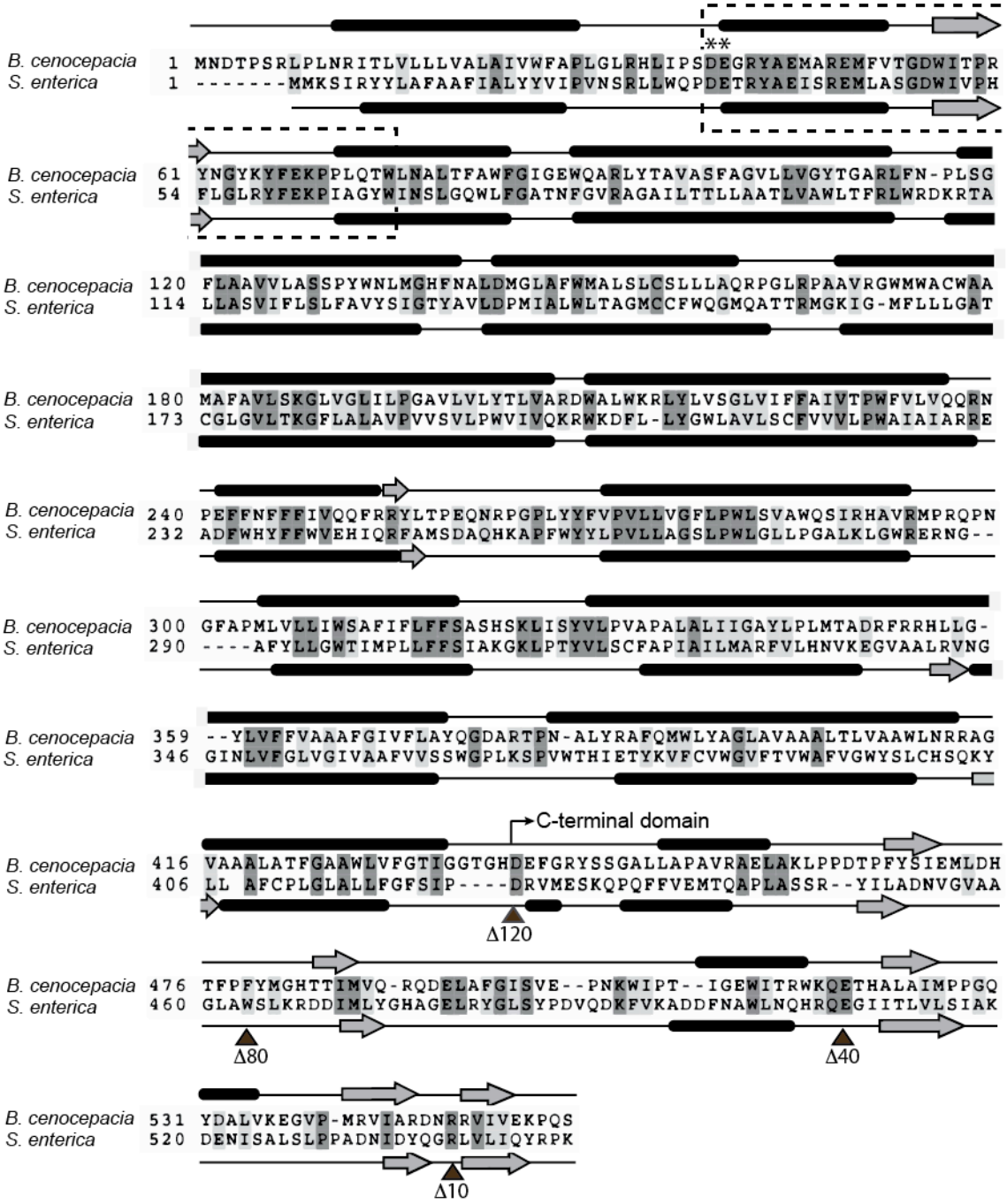

Fig 2. Comparison of the predicted secondary structure of *B. cenocepacia* and *S. enterica* ArnT proteins. Alignment showing identical amino acids in dark grey boxes. Asterisks indicate the DEX motif. Dashed-line box indicates the amino acids corresponding to the signature motif at the first periplasmic loop of ArnT family protein. The secondary structure was predicted with PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/); filled bars and arrows indicate α-helices and β-strands, respectively. The filled triangles below the sequences indicate the endpoint of truncations of the ArnT C-terminal domain.

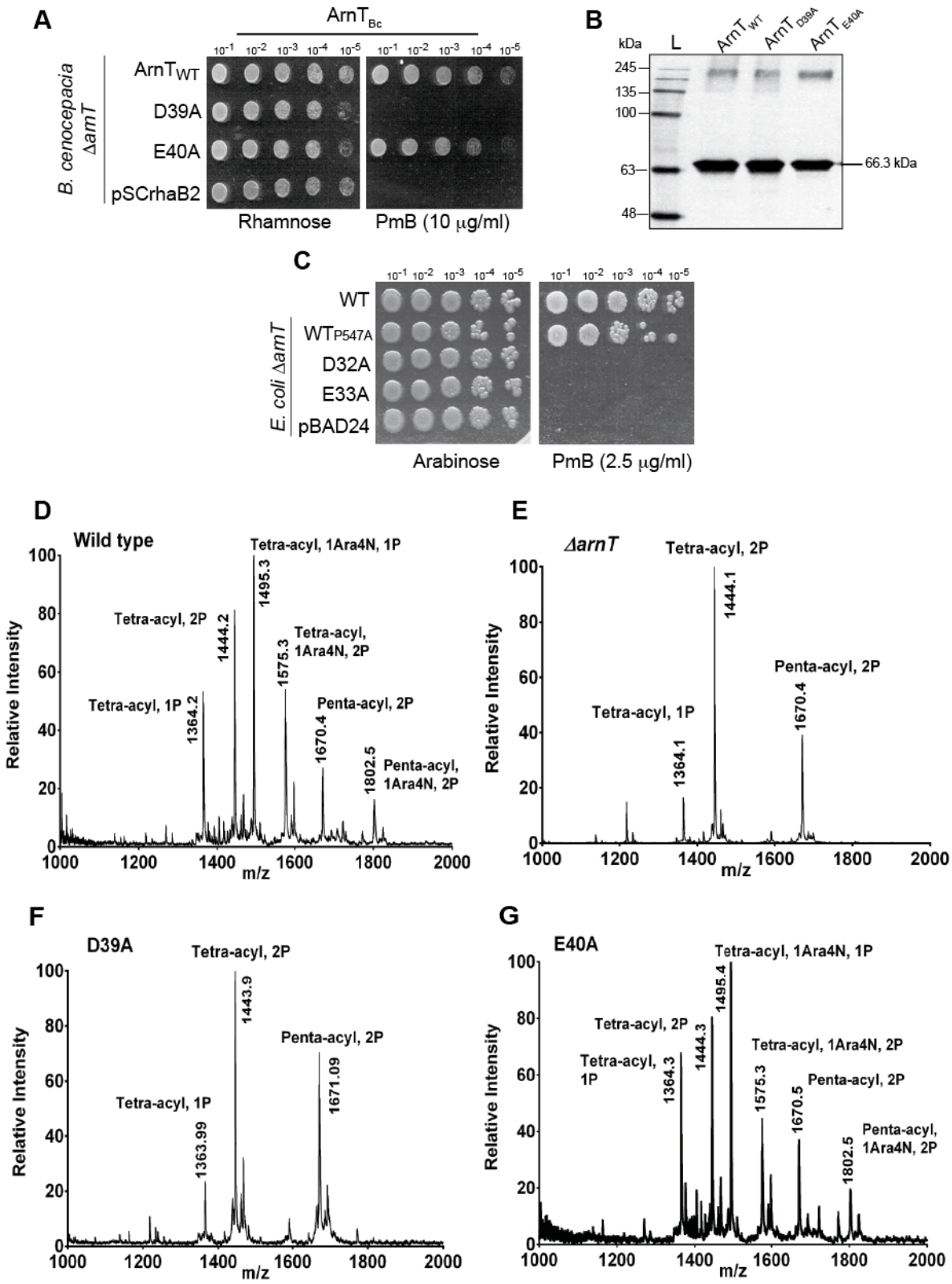

Fig 3. The DEX motif is relevant for ArnT function. (**A**) The functionality of ArnT$_{Bc}$ proteins with alanine replacements in the charged amino acid residues of the DEX motif was assessed using *B. cenocepacia* Δ*arnT-arnBC*$^{+}$ strain MH55 (Table 1) (Δ*arnT*). Bacteria containing pFT3

expressing $ArnT_{FLAG\text{-}10xHis}$ ($ArnT_{WT}$) or the cloning vector pSCrhaB2 were used as positive and negative controls, respectively. Bacteria were spotted on LB plates with 0.4% rhamnose (to allow for protein expression) or 0.4% rhamnose plus PmB 10 μg ml^{-1}. (**B**) Total membrane preparation from Δ*arnT*. Ten μg of protein was analysed by SDS-PAGE, and the immunoblot was probed with the antibody anti-FLAG. L, BLUeye prestained protein ladder (**C**) The functionality of $ArnT_{Se}$ proteins alanine replacements in the DEX motif was assessed in *E. coli* Δ*arnT*. Bacteria expressing $ArnT_{SeP547A\text{-}FLAG\text{-}10xHis}$ ($ArnT_{P547A}$) or the cloning vector pBAD24 were used as a positive control and negative controls, respectively. $ArnT_{P547A}$ is a modified $ArnT_{Se}$ that has an alanine instead of a proline at the C-terminal end of the protein, which does not affect function but makes the protein detectable by immunoblotting (Tavares-Carreon, et al. 2015). WT, indicates the parental *E. coli* BL21(DE3) containing pBAD24. Bacteria were spotted on LB plate with 0.2% arabinose (to allow for protein expression) or 0.2% arabinose plus PmB 2.5 μg ml^{-1}. (**D-G**) MALDI-TOF spectra obtained using the negative ion mode of purified lipid A produced by Δ*arnT* alone and Δ*arnT* expressing $ArnT_{Bc}$ (WT), $ArnT_{D39A}$, and $ArnT_{E40A}$. The profiles represented were tetra-acylated lipid A with 1 or 2 phosphates molecules, tetra-acylated lipid A with 1 L-Ara4N and 1 or 2 phosphate molecules, penta-acylated lipid A, and penta-acylated lipid A with L-Ara4N and 2 phosphates.

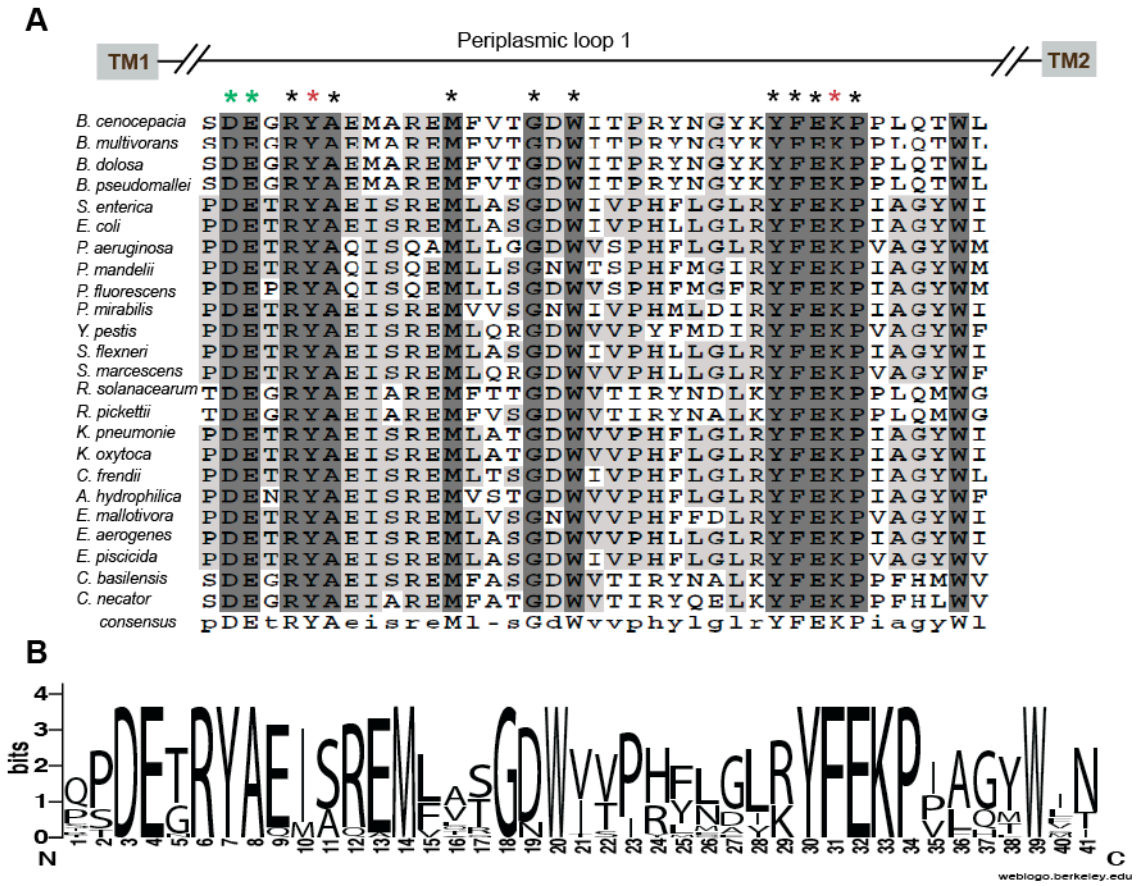

Fig 4. Conserved signature motif in ArnT family members. (**A**) CLUSTAL W multiple sequence alignment of a subset of bacterial ArnT homologues, which were identified from a scan prosite search using the *B. cenocepacia* ArnT protein as a query, as described in Results and in Materials and Methods. Alignments were determined with the program CLUSTAL W. The figure shows a partial alignment of the region spanning the periplasmic loop 1 located between transmembrane domains I and II (indicated with grey bars). Amino acids depicted with a dark grey background and asterisks denote identical residues across the proteins examined. Green and red asterisks indicate residues studied in this work and in a previous study, respectively (Tavares-Carreon, et al. 2015). Conserved residues are marked in light grey. The sequences from the following organisms were compared: *B. cenocepacia*, *B. multivorans*, *B. dolosa*, *B. pseudomallei*, *S. enterica*, *E. coli*, *Pseudomonas aeruginosa*, *P. mandelii*, *P. fluorescens*, *Proteus mirabilis*, *Yersinia pestis*, *Shigella flexneri*, *Serratia marcescens*, *Ralstonia solanacearum*, *R. picketii*, *Klebsiella pneumoniae*, *K. oxytoca*, *Citrobacter freundii*, *Aeromonas hydrophila*, *Erwinia mallotivora*, *Enterobacter aerogenes*, *Edwarsiella piscicida*, *Cupriavidus basilensis*, *C. necator*. (**B**) Weblogo illustrating the conserved DEXRYAX$_{(5)}$MX$_{(3)}$GXWX$_{(9)}$YFEKPX$_{(4)}$W signature motif in the ArnT family.

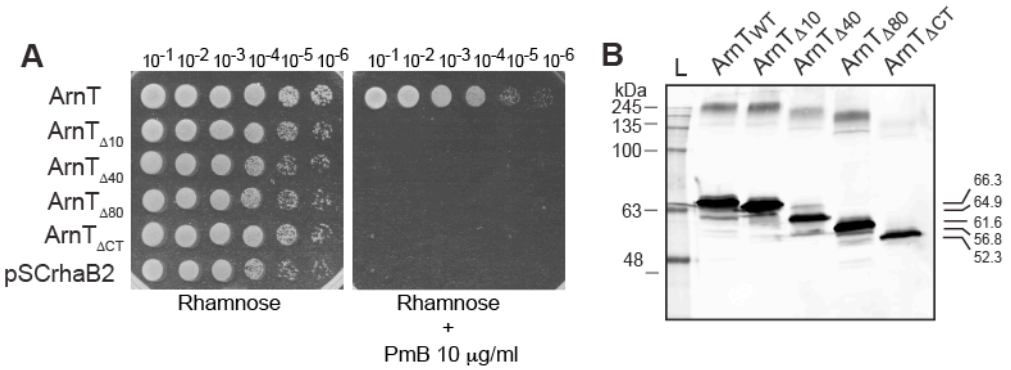

Fig 5. The C-terminal domain of *B. cenocepacia* ArnT is required for function. (**A**) Complementation of the Δ*arnT-arnBC*$^{+}$*lptG*$_{D31H}$ (Δ*arnT;* MH55, Table 1) suppressor strain transformed with pSCrhaB2 encoding, ArnT$_{\Delta 10}$, ArnT$_{\Delta 40}$, ArnT$_{\Delta 80}$, and ArnT$_{\Delta CT}$. Cells expressing ArnT$_{FLAG\text{-}10xHis}$ and the vector pSCrhaB2 were used as positive (WT) and negative controls, respectively. The transformants were spotted on LB supplemented with 0.4% of L-rhamnose and 10 μg ml^{-1} of PmB. (**B**) Immunoblot of total membranes from Δ*arnT* expressing ArnT$_{FLAG\text{-}10xHis}$, ArnT$_{\Delta 10}$, ArnT$_{\Delta 40}$, ArnT$_{\Delta 80}$, and ArnT$_{\Delta CT}$ performed with anti-FLAG. L, BLUeye prestained protein ladder.

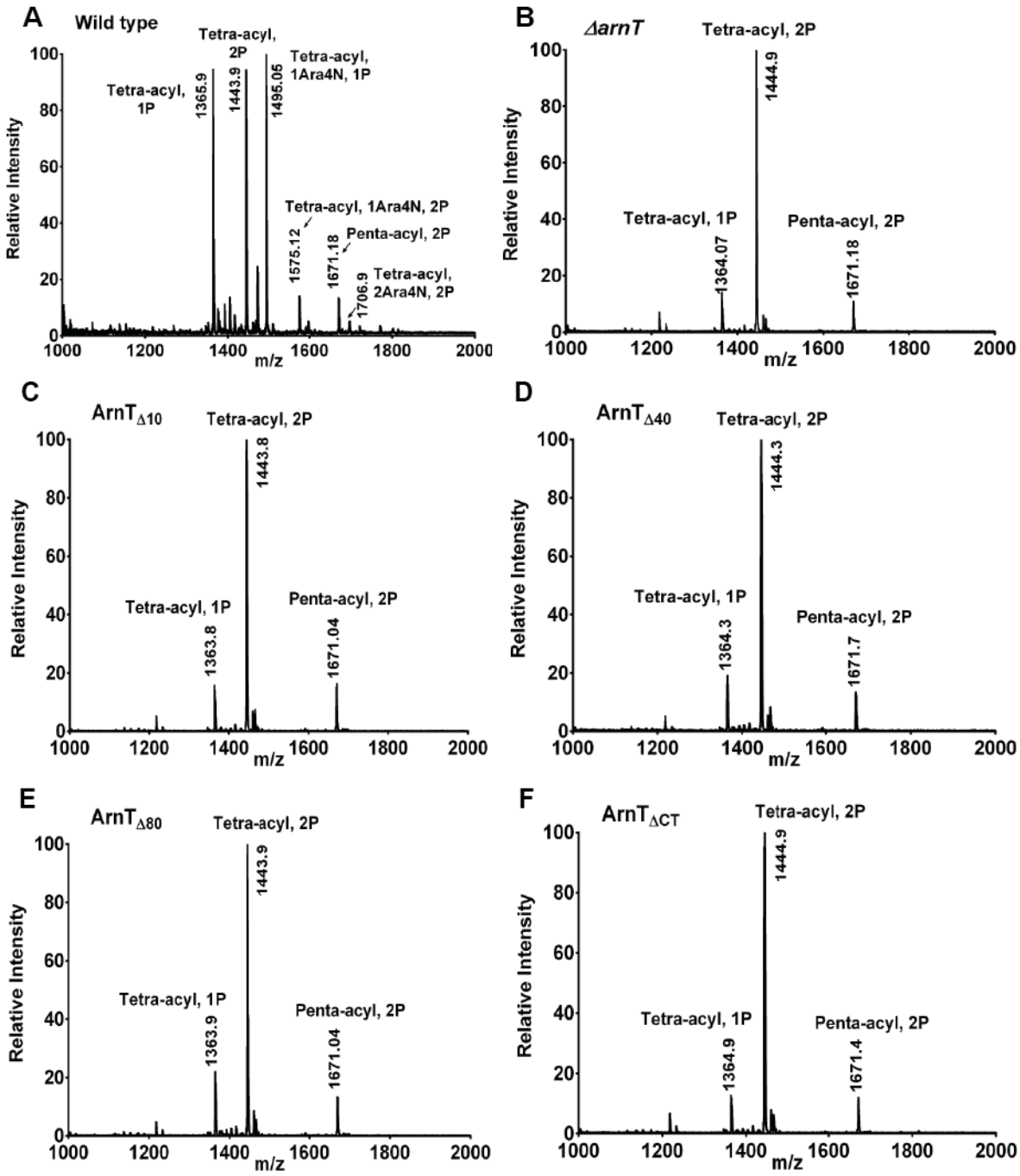

Fig 6. Structural analysis of purified lipid A. MALDI-TOF spectra of purified lipid A produced by the Δ*arnT* suppressor strain transformed with plasmids encoding ArnT wild-type, ArnT$_{\Delta 10}$, ArnT$_{\Delta 40}$, ArnT$_{\Delta 80}$, or ArnT$_{\Delta CT}$. The profiles represented were obtained using the negative ion mode.

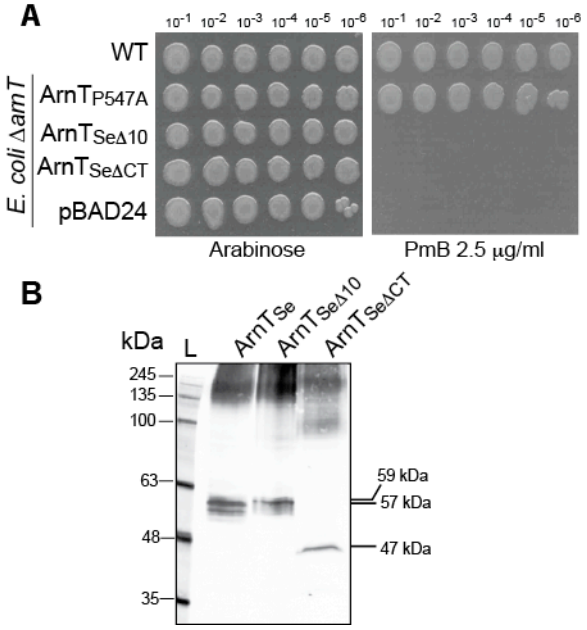

Fig 7. The C-terminal domain is also required for the function of *S. enterica* ArnT. (**A**) Complementation of *E. coli* Δ*arnT* with $ArnT_{SeΔ10}$ and $ArnT_{SeΔCT}$. pFT143 expressing $ArnT_{P547A}$ was used as a positive control. The cloning vector pBAD24 was used as a negative control. WT, indicates the parental *E. coli* BL21(DE3) containing pBAD24. Bacteria were spotted on LB plates with 0.2% arabinose or 0.2% arabinose plus PmB 2.5 μg ml^{-1}. (**B**) Total protein from cells expressing $ArnT_{SeΔ10}$ and $ArnT_{SeΔCT}$ proteins were separated by SDS-PAGE and detected with antibody anti-FLAG. L, BLUeye prestained protein ladder.

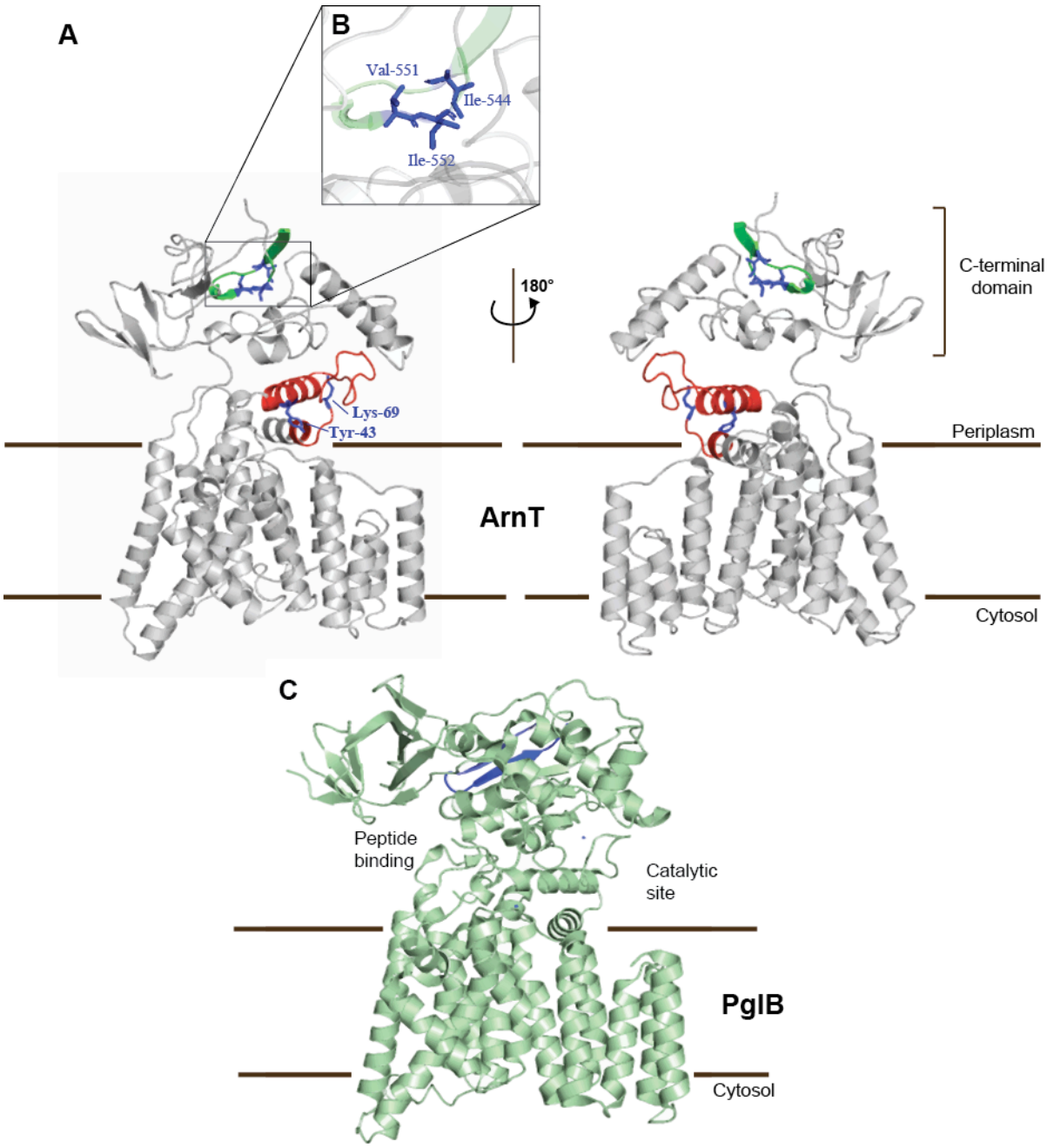

Fig 8. The ArnT predicted tertiary structure resembles the structure of the PglB protein. (**A**) *In silico* model of *B. cenocepacia* ArnT generated by the I-Tasser server. Two views of the structure related by 180° rotation are shown. The last β-strands of the C-terminal domain are indicated in green, the signature motif for the ArnT family in red, and the functional residues Y43, K69, I544, V551, and I552 in blue. (**B**) Close-up of residues I544, V551, and I552 located in the last β-strands of the C-terminal domain of ArnT. (**C**) Solved structure of PglB from *C. lari* (PDB accession number 3RCE) and cavities located in opposite sides of PglB indicated the catalytic side and the peptide binding.

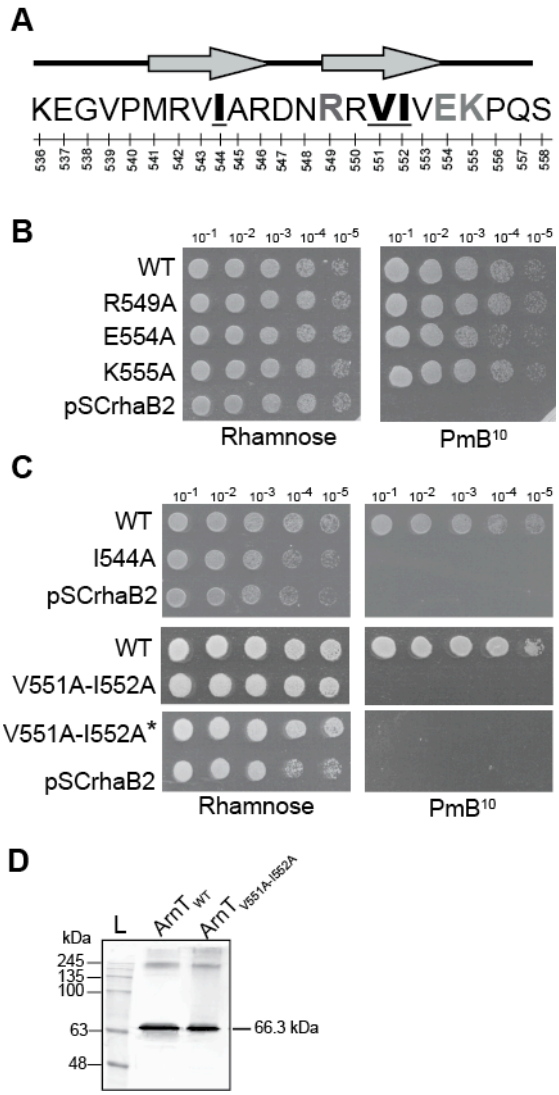

Fig 9. Hydrophobic residues located in the last β-sandwich structure contribute to the overall structure of ArnT. (**A**) Schematic representation of the primary and secondary structure of the last twenty-three amino acids of the C-terminal domain of ArnT. Arrows indicate β-strands and the residues replaced by alanine are underlined. (**B-C**) Functional analysis of alanine replacements in charged or hydrophobic amino acid residues by complementation of Δ*arnT* (strain MH55, Table 1). $ArnT_{FLAG\text{-}10xHis}$ and pSCrhaB2 were used as positive and negative control, respectively. Rham, 0.4% rhamnose; PmB, 0.4% rhamnose plus polymyxin B 10 μg ml^{-1}. V551A-I552A has a double epitope at the C-terminal whereas V551A-I552A* does not express the FLAG-10xHis double tag. (**D**) Immunoblot of total membranes preparation from Δ*arnT* containing plasmids expressing $ArnT_{WT}$ and $ArnT_{V551A\text{-}I552A}$. Ten μg of protein was separated by SDS-PAGE, and the immunoblot was probed with the antibody anti-FLAG. L, BLUeye prestained protein ladder.

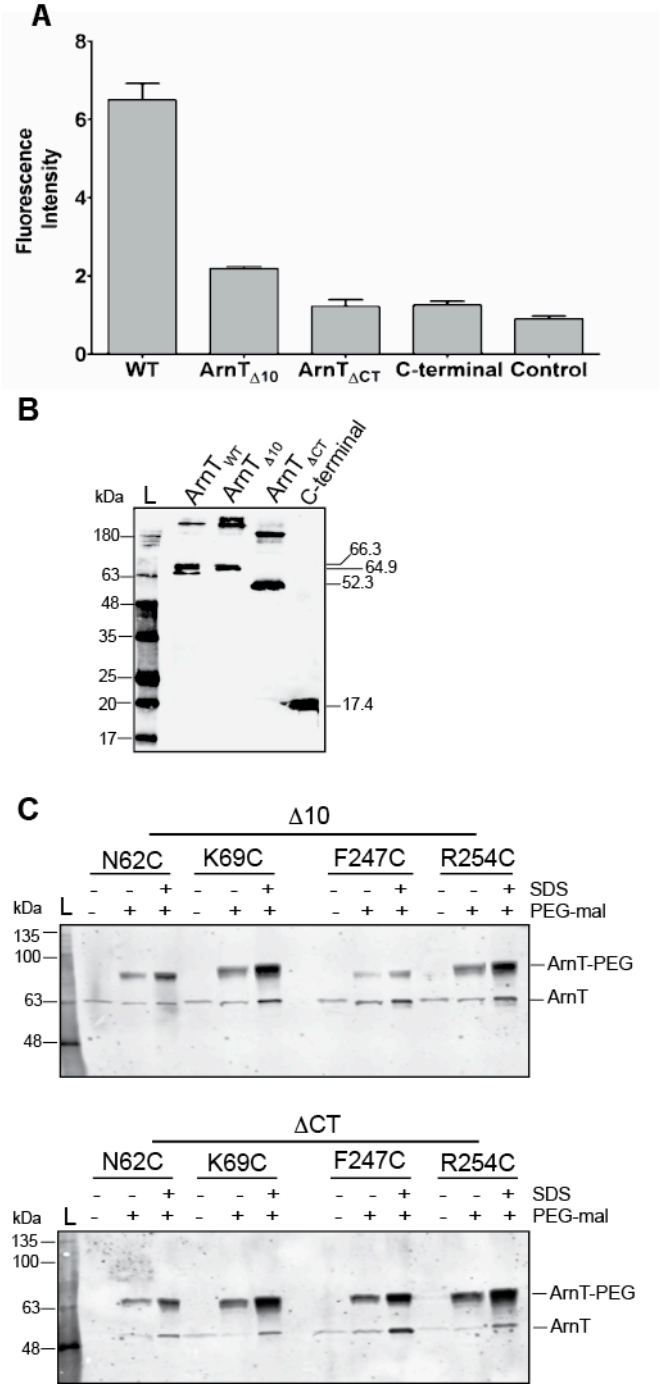

Fig 10. The C-terminal domain of ArnT is involved in the Lipid A binding. (**A**) Quantification of the LPS binding assay, 200 μg of each purified protein were incubated with 10 μg of Alexa-LPS at room temperature. Fluorescence signal was normalized to the buffer alone. Binding of LPS to the resin alone as a control. Values represent the mean of four independent experiments. (**B**) $ArnT_{WT}$, $ArnT_{\Delta 10}$, $ArnT_{\Delta CT}$ and C-terminal domain purified proteins were separated by SDS-PAGE 16% and detected with antibody anti-FLAG. The high-mass oligomeric forms of each result from the mild denaturation conditions used to visualize ArnT. L, BLUeye prestained protein ladder. (**C**) PEG-mal labelling of periplasmic amino acids in $ArnT_{\Delta 10}$ (Δ10) or $ArnT_{\Delta CT}$ (ΔCT). *E. coli* DH5α cells carrying plasmids encoding ArnT cysteine replacements with the

respectively C-terminal were harvested and resuspended in HEPES/$MgCl_2$ buffer. 0.1 ml of cell suspension was incubated with buffer alone or 1 mM PEG-mal with or without 2% SDS for 1 h at room temperature. Immunoblot was performed with anti-FLAG antibodies. L, BLUeye prestained protein ladder.