

Structural insights into terminal arabinosylation of mycobacterial cell wall arabinan

Received: 23 August 2024

Accepted: 14 March 2025

Published online: 29 April 2025

Check for updates

Yaqi Liu¹, Chelsea M. Brown^{2,3}, Satchal Erramilli⁴, Yi-Chia Su⁵, Shih-Yun Guu⁵, Po-Sen Tseng⁶, Yu-Jen Wang⁵, Nam Ha Duong^{5,7,8}, Piotr Tokarz⁴, Brian Kloss¹, Cheng-Ruei Han⁵, Hung-Yu Chen⁵, José Rodrigues⁹, Kay-Hooi Khoo^{5,10}, Margarida Archer⁹, Anthony A. Kossiakoff⁴, Todd L. Lowary^{5,6,10}✉, Phillip J. Stansfeld²✉, Rie Nygaard^{1,11}✉ & Filippo Mancía¹✉

The global challenge of tuberculosis, caused by *Mycobacterium tuberculosis* (*Mtb*), is compounded by the emergence of drug-resistant strains. A critical factor in *Mtb*'s pathogenicity is its intricate cell envelope, which acts as a formidable barrier against immune defences and pharmacological interventions. Central to this envelope are arabinogalactan (AG) and lipoarabinomannan (LAM), two complex polysaccharides containing arabinan domains essential for maintaining cell wall structure and function. The arabinofuranosyltransferase AftB plays a pivotal role in the biosynthesis of these arabinan domains by catalyzing the addition of β -(1 \rightarrow 2)-linked terminal arabinofuranose residues. Here, we present the cryo-EM structures of *Mycobacterium chubuense* AftB in both its apo form and bound to a donor substrate analog, resolved at 2.9 Å and 3.4 Å resolution, respectively. These structures reveal that AftB has a GT-C fold, with a transmembrane (TM) domain comprised of eleven TM helices and a periplasmic cap domain. AftB has a distinctive irregular, tube-shaped cavity that connects two proposed substrate binding sites. Through an integrated approach combining structural analysis, biochemical assays, and molecular dynamics simulations, we delineate the molecular basis of AftB's reaction mechanism and propose a model for its catalytic function.

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is the world's second deadliest infectious disease, surpassed only by COVID-19 during the recent pandemic¹. The resilience and virulence of *Mtb* are largely due to its unique cell envelope—a lipid-rich fortress that confers resistance to host immune mechanisms and provides an impenetrable barrier to a wide range of antibiotics^{2–5}. This complex cell envelope is comprised of three major components: long-chain mycolic acids that

form a protective lipid layer, peptidoglycan that provides cell shape and structural rigidity, and arabinogalactan (AG), which links these two elements (Fig. 1a). AG is a polysaccharide with a linear galactan backbone and a highly branched arabinan domain. Together, these three components constitute the mycolyl–arabinogalactan–peptidoglycan (mAGP) complex, an impermeable shield critical to *Mtb*'s survival and pathogenicity^{6–9} (Fig. 1a). In addition to the mAGP complex, the cell

A full list of affiliations appears at the end of the paper. ✉e-mail: tlowary@as.edu.tw; Phillip.Stansfeld@warwick.ac.uk; rin7007@med.cornell.edu; fm123@cumc.columbia.edu

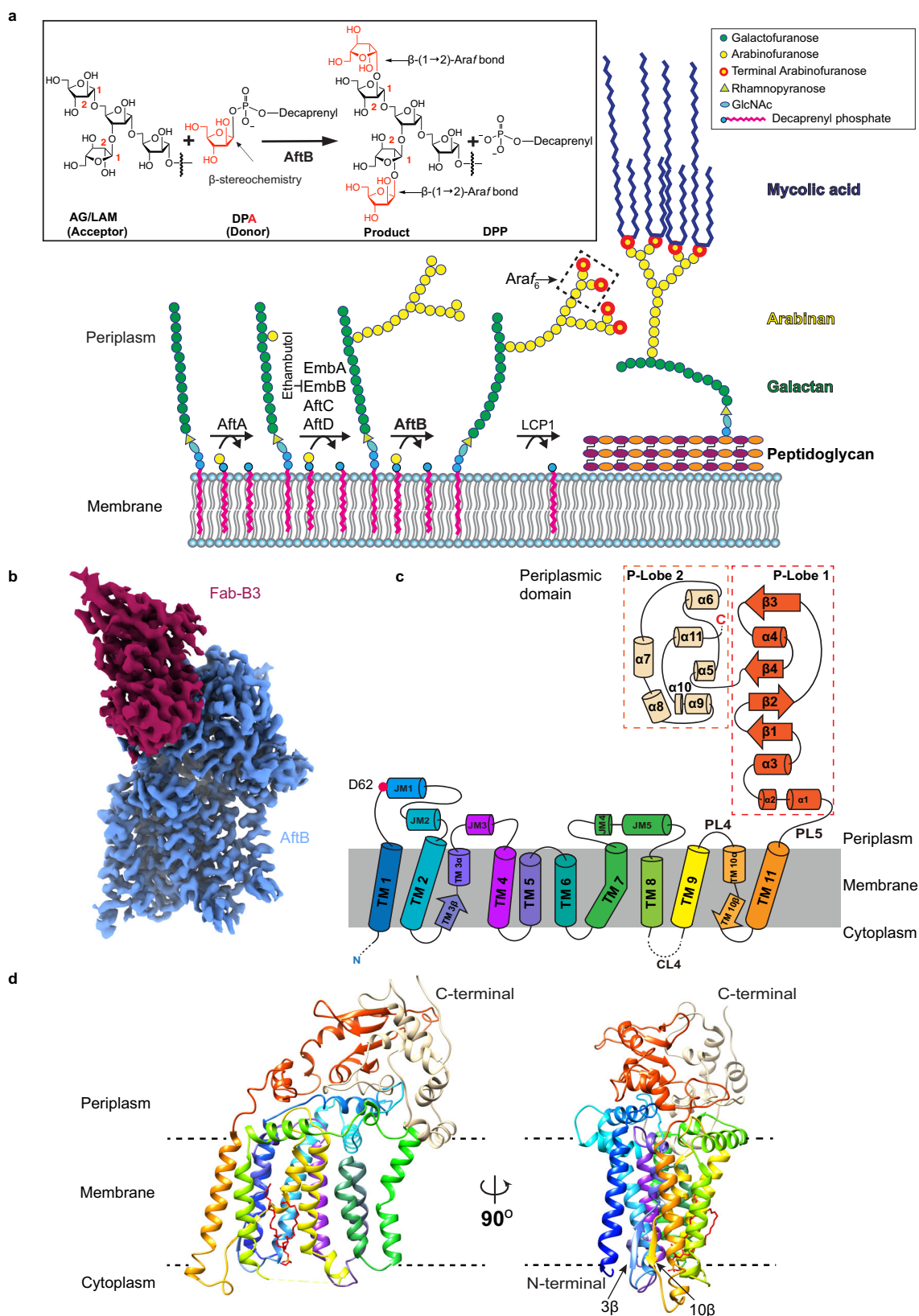


Fig. 1 | Biosynthetic pathway and structural architecture of AftB. **a** Schematic of the arabinogalactan biosynthetic pathway in mycobacteria, featuring the enzymatic reactions carried out by AraTs. Enzyme names are indicated above the reaction arrows. The inset shows the β -(1 \rightarrow 2) arabinosyl transfer reaction catalyzed by AftB. The arabinan chain is shown linked to the 10th position of the galactan chain through a β -(1 \rightarrow 2) bond. While this linkage also occurs at the 8th or 12th position, the 10th position is depicted here for clarity. **b** Cryo-EM density

map of AftB in complex with Fab-B3. AftB is shown in blue, while Fab-B3 is shown in maroon. **c** Topological diagram of AftB showing the arrangement of secondary elements in the TM region and in the PD, annotated and colored in correspondence with **(d)**. The catalytic residue D62 is denoted as a red dot. **d** Structure of AftB shown in ribbon. The unmodeled region of AftB is marked as dashed lines, and membrane boundaries are marked by dotted horizontal bars.

envelope contains an array of glycolipids, including phosphatidylinositol mannosides (PIMs) and their structurally related derivatives: lipomannan (LM), lipoarabinomannan (LAM), and its delipidated form, arabinomannan (AM). These glycolipids play key roles in *Mtb*'s growth, virulence, and interaction with the host immune system^{9,10}. Among them, LAM/AM stands out due to its intricate structure, comprising a mannan core with a highly branched arabinan domain^{10–14} (Supplementary Fig. 1a).

The unique characteristics of the mycobacterial cell envelope have made it a pivotal target for anti-tuberculosis (TB) therapeutics^{5,6,15–17}. Front-line drugs, such as isoniazid and ethambutol, specifically target *Mtb* cell envelope biosynthesis^{18,19}. Isoniazid disrupts the synthesis of mycolic acids, a key component of the mAGP^{18–20} complex, while ethambutol inhibits enzymes essential for arabinan biosynthesis, thereby impairing AG and LAM maturation^{21–24}. Despite their efficacy, the rise of drug-resistant and multidrug-resistant *Mtb* strains has exposed the limitations of existing TB treatments and underscore the urgent need for innovative therapeutic targets and strategies^{5,25,26}.

The biosynthesis of the arabinan domains of AG and lipoarabinomannan LAM is catalyzed by a series of arabinofuranosyltransferases (AraTs)^{5,6}. These membrane-embedded glycosyltransferases are responsible for transferring arabinofuranose (Araf) units from β -D-arabinofuranosyl-1-monophosphoryldecaprenol (DPA), the only known D-arabinose donor substrate, onto growing polysaccharide chains^{27–29}. For AG, the process begins with the α -(1 \rightarrow 5)-arabinosyltransferase AftA, which attaches the initial α -D-Araf residue to the mature galactan backbone³⁰ (Fig. 1a). The chain is then extended via α -(1 \rightarrow 3)²⁴ linkages by the enzymes EmbA and EmbB. (Fig. 1a). For LAM, the priming enzyme responsible for initiating arabinan synthesis remains unidentified. Nonetheless, EmbC is involved in the extension of arabinan-primed LM with α -(1 \rightarrow 5)-Araf residues³¹. AftC introduces α -(1 \rightarrow 3)-branching within both AG and LAM, adding structural complexity to these components of the cell envelope³² (Fig. 1a). Recent studies suggest that AftD, previously thought to function like AftC³³, is instead an α -(1 \rightarrow 5) AraT that extends α -(1 \rightarrow 3) branched motifs in the arabinan domain³⁴. Finally, AftB, an integral membrane enzyme encoded by the gene Rv3805c in *Mtb*, adds terminal β -(1 \rightarrow 2)-linked D-Araf residues to both AG and LAM/AM^{35,36} (Fig. 1a and Supplementary Fig. 1a). In AG, this results in a terminal hexaarabinofuranoside motif (Araf₆) at the non-reducing end, which provides the anchoring point for mycolic acid^{6,37} (Fig. 1a). In LAM/AM, AftB adds β -(1 \rightarrow 2)-linked terminal-Araf residues to both branched and linear arabinan motifs, with the latter resulting in linear Araf₄ structures with terminal β -linked Araf motifs³⁵ (Supplementary Fig. 1a). The AG is further ligated to peptidoglycan by the phosphotransferase Lcp1, completing the assembly³⁸ (Fig. 1a).

AftB plays a critical role in maintaining the integrity of the mycobacterial cell envelope and is a promising drug target^{35,36,39} due to its essentiality^{35,36} in *Mtb*. Despite its importance, the structural and molecular basis of AftB's function remains poorly understood.

Here, we present single-particle cryogenic-electron microscopy (cryo-EM) structures of *Mycobacterium chubuense* AftB in both apo and substrate-analog-bound states. By integrating structural data, biochemical assays, and molecular dynamics (MD) simulations, we identify key substrate-binding sites and propose a mechanism for AftB's enzymatic activity.

Results

Structure determination of AftB

We screened AftB orthologs from 45 mycobacterial species for expression in *E. coli* and identified *Mycobacterium chubuense* AftB (McAftB) as the most promising target for structural studies. To verify the activity of recombinant AftB, we adapted a previously reported arabinosyltransferase assay⁴⁰. This assay uses a synthetic tetrasaccharide acceptor substrate (Supplementary Fig. 2a) corresponding to the non-reducing terminus of AftB's natural substrate in AG and LAM/AM, and farnesyl phosphoarabinose (FPA) as a donor

surrogate (Supplementary Fig. 2b). FPA, known to be an effective donor for mycobacterial AraTs, was chosen for its ease of preparation and handling compared to the native substrate, DPA^{40,41}. Mass spectrometry analysis of the acetylated reaction products confirmed that AftB catalyzes the stepwise transformation of the tetrasaccharide into a pentasaccharide and, ultimately, a hexasaccharide, thus demonstrating the enzymatic activity of recombinant McAftB (Supplementary Fig. 2a–c and Supplementary Fig. 2e).

A previous study demonstrated AftB activity using a disaccharide acceptor³⁶. To expand upon this, we opted for a larger, branched tetrasaccharide acceptor to investigate whether AftB exhibits a preference in the sequential addition of the two β -(1 \rightarrow 2)-linked Araf residues during the formation of Araf₆. The resulting pentasaccharide and hexasaccharide products were analyzed using mass spectrometry (MS/MS) (Supplementary Fig. 3).

The permethylated Araf₅ and Araf₆ oligomers with an azide group were detected by MALDI-MS as [M+Na]⁺, accompanied by prominent in source loss of an N₂ group (–28 Da). When the sodiated molecular ions were selected for Collision-Induced Dissociation (CID) MS/MS, all reducing end fragment ions detected, including the base peaks, similarly showed a loss of N₂. Of note, the fragmentation pattern afforded is identical to that observed previously when permethylated Ara oligomers were analyzed by MALDI TOF/TOF MS/MS^{42,43}. This allows unambiguous assignment of the expected branching pattern of Ara₅ and Ara₆, as illustrated in Supplementary Fig. 3.

For the Araf₅ product (Supplementary Fig. 3a), the loss of the third Araf from the non-reducing end could not occur without creating another free OH (*m/z* 472 instead of 486), whereas the loss of one (*m/z* 806), two (*m/z* 646) and four (*m/z* 326) Araf residues could result from a single glycosidic cleavage. This is consistent with a branched Araf₅ structure extended by one and two Araf residues from the branch point. The same applies to the Araf₆ product (Supplementary Fig. 3b), whereby a single glycosidic cleavage could only lead to the loss of one (*m/z* 966) and two (*m/z* 806) but not three Araf residues, which would require a second cleavage (*m/z* 632 instead of 646). The branch point is further corroborated in both cases by the prominent sodiated B ion of Araf₂ (*m/z* 375) but not Araf₃.

The sodiated B ion for Araf₄ (*m/z* 695) could additionally be detected for the Araf₅ product, accompanied by the cross-ring cleavage ions, ^{2,4}A and ^{0,3}A ions at *m/z* 765 and 737, respectively. The corresponding ^{2,4}A ion for Araf₃ detected at *m/z* 605 defines a O3,O5-di-Araf-substitution. Importantly, the unique ^{0,3}A ions at *m/z* 257 and 417 indicate that both isomeric Araf₅ products were made. This is further supported by the ions at *m/z* 297 and 457 resulting from further loss of the O3-substituent from the E ion (see the illustrated cleavage drawing in Supplementary Fig. 3a). Based on the similar intensity for the pairs of *m/z* 257/297 versus *m/z* 417/457, it can be further inferred that both isomeric products were approximately of equal abundance.

Finally, the ^{0,2}X ion, which would retain the O2-substituent, was detected at *m/z* 862 for the Araf₅ product. Because it could also be derived from the non- β -(1 \rightarrow 2)-extended terminal Ara of either of the two isomeric Ara₅ products, it is by itself not critical evidence for terminal β -(1 \rightarrow 2)-Araf-substitution. However, in the case of the Araf₆ product, the corresponding ion at *m/z* 1022 unambiguously defines the O2-substitution (Supplementary Fig. 3b).

Insufficient quantities of the products were available for ¹H nuclear magnetic resonance (NMR) spectroscopy to confirm the β -stereochemistry of the added Araf residues. However, the absence of reports describing α -Araf-(1 \rightarrow 2)- α -Araf motifs in mycobacterial arabinan supports the conclusion that the products contained β -(1 \rightarrow 2)-linked Araf residues. Taken together, these data confirm that McAftB converts the tetrasaccharide acceptor into the native Araf₆ motif present at the non-reducing terminus of AG and LAM/AM, and that the enzyme does not have a preference for which arm of the acceptor is first arabinofuranosylated.

The 78 kDa *McAftB*, was purified in detergent and reconstituted into lipid-filled nanodiscs for cryo-EM structure determination (Supplementary Fig. 1b, c). To provide fiducials for particle alignment and increase particle size to facilitate structure determination, we screened a synthetic phage display library to select recombinant antigen-binding fragments (Fabs) against *McAftB*^{44,45}. Seven high-affinity Fabs were identified, with Fab-B3 selected for its strong binding affinity (Supplementary Fig. 1d, e).

Using 7164 micrographs of nanodisc-reconstituted apo *McAftB* in complex with Fab-B3, iterative 2D classification yielded high-quality class averages with distinct features for the transmembrane (TM) domain and bound Fab. Three-class *ab initio* particle sorting produced a 2.9 Å resolution map, enabling nearly complete atomic modeling of *McAftB*. (Supplementary Fig. 4a–c). This allowed us to build an almost complete atomic model of *McAftB* except for disordered regions comprising 28 residues at the N-terminus, 11 at the C-terminus, and a segment (323–334) in a cytoplasmic loop connecting two TM helices (8 and 9) (Fig. 1c, d and Supplementary Fig. 5). The Fab variable region was well resolved, allowing precise modeling of the *McAftB*–Fab interaction interface, while the disordered constant domain was excluded from the model (Fig. 1b and Supplementary Fig. 1f, g).

Additionally, we observed a distinct diacyl-glycerophospholipid-like density between TM helices 8 and 9. We tentatively attributed this to phosphatidylethanolamine (PE) (Supplementary Fig. 5), given its abundance in *E. coli* membranes⁴⁶.

Overall structure of AftB

AftB comprises two domains: a TM domain with 11 α -helices and a C-terminal periplasmic domain (PD) composed of a combination of α -helices and β -strands (Fig. 1c, d). The TM helices of AftB are connected by five cytoplasmic loops (CL1–CL5) and four periplasmic loops (PL1–PL4). While most cytoplasmic loops are well resolved in the structure, CL4 (residues 323–334), linking TM helices 8 and 9, appears disordered in the structure (Fig. 1c, d).

Among the periplasmic loops, three loops (PL1, PL2, and PL3) have a more complex structure. PL1, between TM helices 1 and 2, includes two juxtamembrane (JM) helices, JM1 and JM2. PL2, connecting TM helices 3 and 4, contains JM3, while PL3, spanning between TM helices 7 and 8, contains JM4 and JM5 (Fig. 1c, d). The C-terminal PD of AftB extends from the end of TM helix 11, forming a dome-like structure that caps the TM helical bundle in the periplasmic space (Fig. 1b, d). This domain consists of two lobes: the first consists of four β -strands (β 1– β 4) and four α -helices (α 1– α 4), along with inter-linking loops. The second lobe contains seven inter-connected α -helices (α 5– α 11). PL5 serves as a linker, connecting TM11 to α 1, and anchoring the PD to the TM bundle (Fig. 1d).

Fab-B3 binds specifically to the PD of AftB, with both its light and heavy chains contributing to the interaction interface (Supplementary Fig. 1f). The binding interface is characterized by a network of hydrogen bonds, predominantly mediated by the heavy chain of the Fab fragment (Supplementary Fig. 1g).

Conserved GTC fold of AftB

Glycosyltransferases, a structurally diverse enzyme family, can be classified into three primary fold types: GT-A, GT-B, and GT-C. GT-A and GT-B folds typically feature Rossmann-like domains, while GT-C fold enzymes are characterized by multiple TM helices. GT-C enzymes are further divided into GT-C_A and GT-C_B subclasses^{47–49}. Using the DALI server⁵⁰, we identified AftB as a member of the GT-C_A superfamily of glycosyltransferases⁴⁷. The first seven transmembrane helices of AftB adopt the GT-C glycosyltransferase fold^{47,51}. Enzymes with the highest structural homology to AftB include bacterial oligosaccharyltransferase PglB^{52,53}, bacterial aminoarabinose-transferase ArnT⁵⁴, and yeast glucosyltransferase ALG6⁵⁵ (Supplementary Fig. 7). These

enzymes all belong to the GT-C_A class⁴⁷ and share a conserved N-terminal module with AftB. This includes TM helices 1–7 and the two JM helices (JM1 and JM2) between TM helices 1 and 2.

While adopting the core GT-C_A fold, AftB exhibits some notable differences. Specifically, both TM helices 3 and 10 in AftB are split into a short α -helix and a β -strand, forming a small anti-parallel β -sheet within the TM domain (Fig. 1c, d and Supplementary Fig. 5b). This feature is distinct to AftB and contrasts with other GT-C_A enzymes, where TM3 typically exhibits a kink that splits it into two helices. The functional significance of these structural differences in AftB remains to be elucidated but could represent adaptations to its specific role in mycobacterial arabinan biosynthesis.

The putative substrate cavity of AftB

A tube-shaped cavity was observed in AftB, spanning between the PD and TM domains. One end of the cavity opens towards the lipid membrane, and the other extends into the periplasmic space (Fig. 2a, b). This cavity appears to serve as the pathway for substrate entry. All GT-CA enzymes contain a conserved catalytically essential aspartate residue at the tip of JM helix 1^{53–56}. In *McAftB*, this conserved aspartate corresponds to D62, which is part of a DD motif and is buried inside the cavity (Fig. 2a and Supplementary Fig. 8). A D62A mutant exhibited complete loss of enzymatic activity while maintaining comparable expression levels and proper folding compared to wild-type *McAftB* (Supplementary Fig. 9 and Supplementary Fig. 10), confirming its essential catalytic role.

The cavity is lined by JM helices 1, 2, and 4, as well as PL4, and connecting loops between TM and JM helices (Fig. 2a). On the periplasmic side, the cavity is defined by loops connecting TM helix 11 and α 1, β 3 and α 4, β 4 and α 5, and α 5 and α 6. This architecture creates an entry point for the donor substrate, DPA, through the membrane-facing opening. This would permit the polar head of DPA to be oriented towards the catalytic core of the enzyme, with the lipid tail extending out into the lipid bilayer. The wider opening of the cavity towards the periplasm (Fig. 2a) presents a predominantly hydrophilic environment (Supplementary Fig. 11f), which could accommodate arabinan acceptor substrates with different end structures, enabling AftB to generate both the branched *Araf*₆ terminal motif in AG and LAM/AM and the linear *Araf*₄ terminal structures in LAM/AM.

To further substantiate our hypothesis for entry of the acceptor and donor substrates, we performed coarse-grained molecular dynamics (CG-MD) simulations to identify preferential binding sites and entry pathways for the substrates. We used the native donor substrate DPA (Supplementary Fig. 11a, left) and a hexasaccharide acceptor fragment, referred to as “terminal-*Araf*₄” (Supplementary Fig. 11a, right). This acceptor species includes the branched motif present in the acceptor substrate used in the functional assay with two additional α -(1 \rightarrow 5)-*Araf* residues at the non-reducing end (Fig. 1a, Supplementary Fig. 11a). DPA, initially randomly distributed in the membrane’s upper leaflet, preferentially bound within the cavity defined by TM helices 5, 6, and 7, and JM helices 1, 2 and 4 (Fig. 2c), consistent with the cavity proposed above (Fig. 2a). Terminal-*Araf*₄, starting in bulk solvent, rapidly associated with the cavity from the periplasmic side, binding near loops connecting β 3 and α 4, β 4 and α 5, and α 5 and α 6 (Fig. 2d). This corresponds to the predicted entryway for the acceptor substrate (Fig. 2a, b).

The presence of Fab-B3, used during structural determination, did not significantly alter substrate interactions, as shown by the CG-MD simulations (Supplementary Fig. 13a, right). Substrate binding was observed both with and without Fab. Furthermore, RMSF analyses of atomistic simulations revealed similar protein dynamics in apo, Fab-bound, and substrate-bound states (Supplementary Fig. 13a, left). These findings support that the Fab-bound structure represents a catalytically relevant conformation.

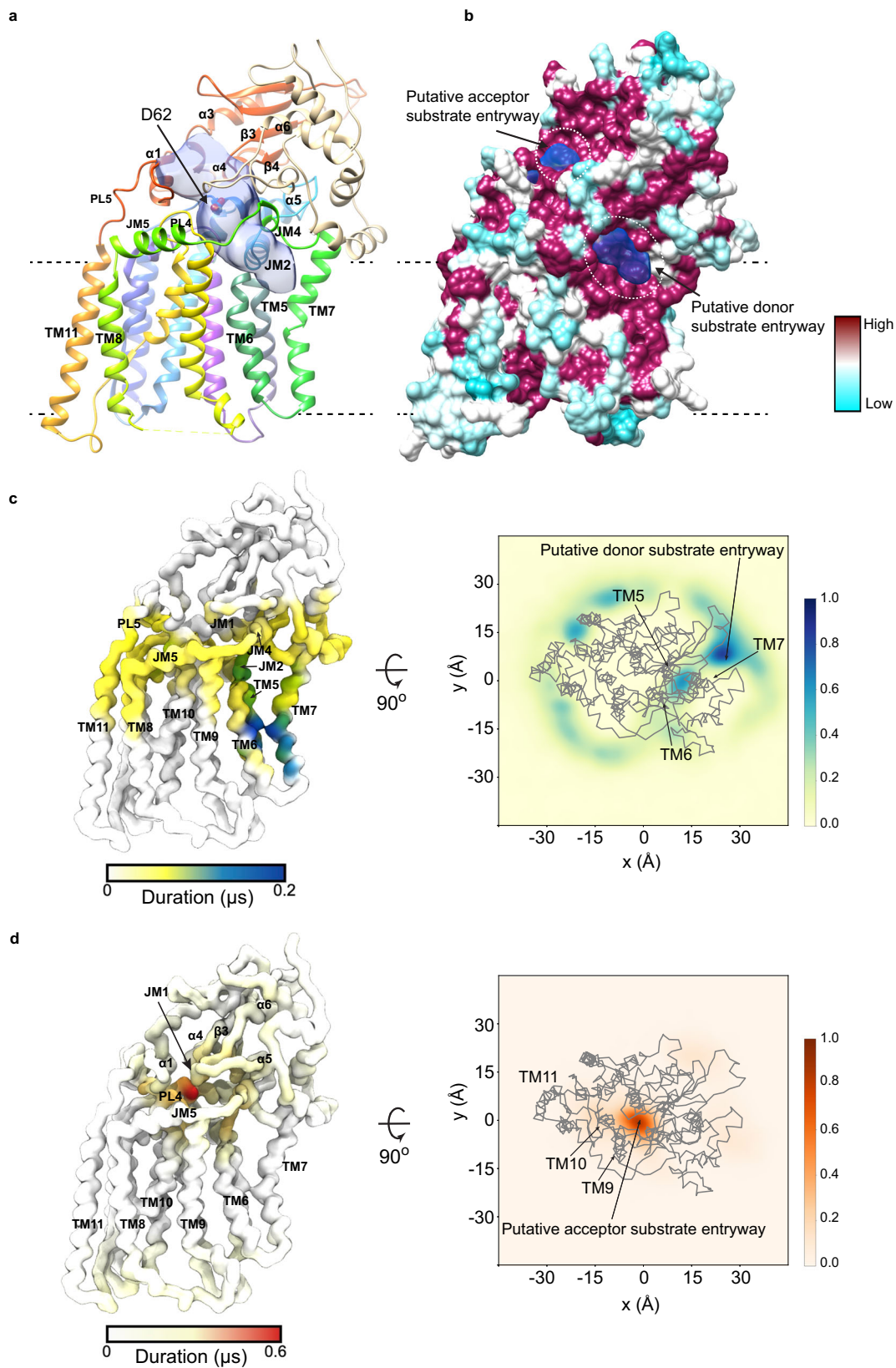


Fig. 2 | Putative substrate-binding cavity. **a** AftB structure shown as ribbon, and colored as in Fig. 1c, with the cavity shown in semitransparent blue. Volumes were calculated using the Voss Volume Voxelator server³⁴ with probe sizes of 2 Å and 8 Å. **b** Surface of AftB by conservation, with a gradient from magenta (absolutely conserved) to cyan (least conserved), cavity shown in semitransparent blue. Predicted entryways

for substrates are marked with dotted circles. **c** Left shows the AftB backbone as a surface, colored by the duration of interaction with DPA in CG-MD simulations. The darker the color, the longer the duration of interaction. The right plot shows the density in the x and y dimensions of DPA relative to the protein shown in gray, where the darker regions show higher density. **d** Shows the same as (c) but for terminal-Araf₄.

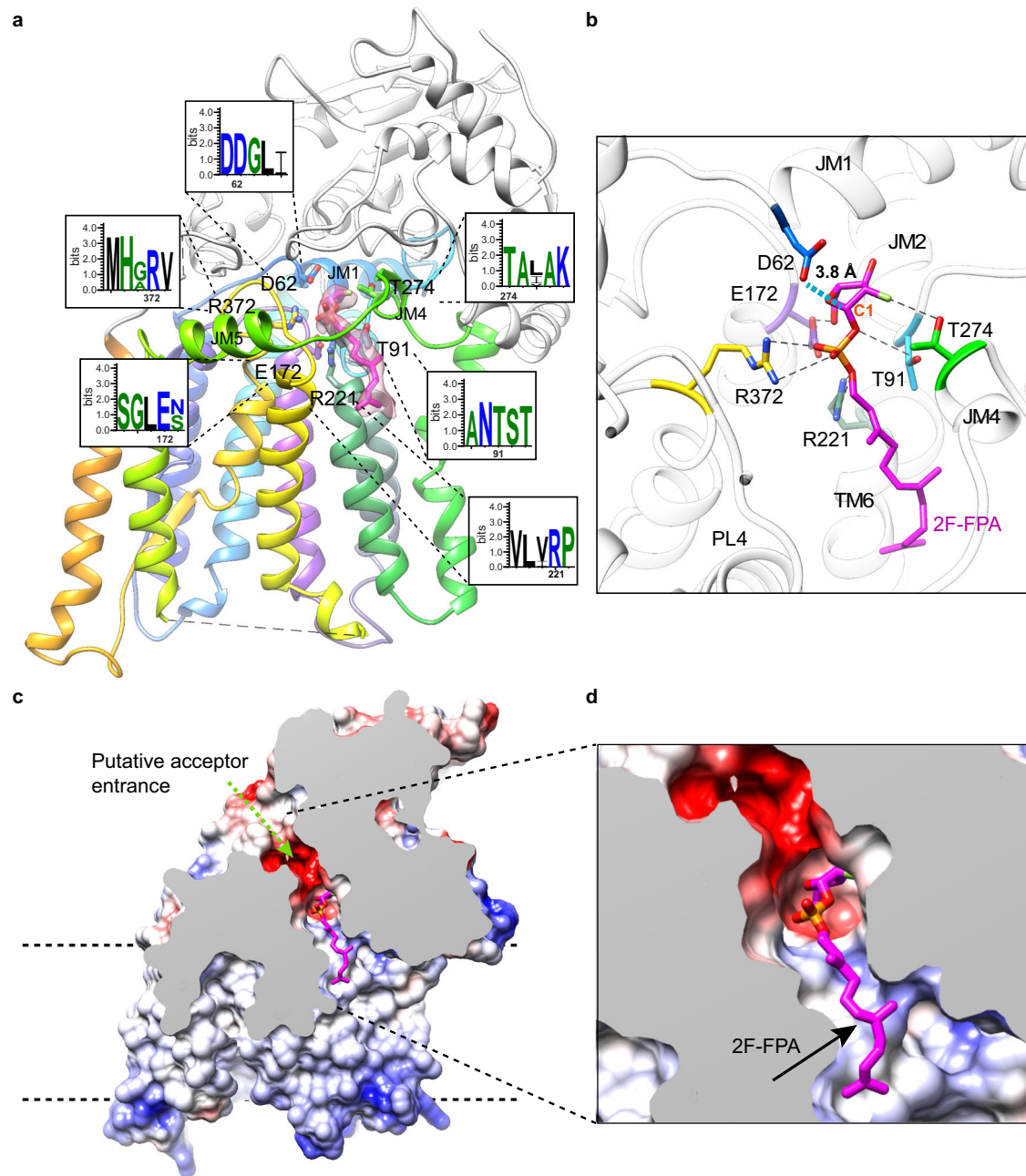


Fig. 3 | Interaction between AftB and bound 2F-FPA. **a** Structure of AftB complexed with 2F-FPA shown in ribbon. The TM domain of AftB is colored in rainbow as in Fig. 1d, while the PD is colored in white. 2F-FPA is shown in sticks and colored in magenta. *Cryo-EM density* of 2F-FPA colored with semi-transparent magenta. WebLogo plots¹⁰² show the amino acids interacting with 2F-FPA. **b** Close-up view of the interaction between AftB and 2F-FPA. Polar contacts are indicated as black

dashed lines. Residues involved in the interaction are shown as sticks. The distance between the catalytic residue D62 and C1 of 2F-FPA is indicated with blue dashed lines. **c, d** Electrostatic surface depiction of the AftB-2F-FPA complex, a partial cross-sectional perspective is applied to highlight the active site cavity with the bound donor substrate-analog 2F-FPA.

The cryo-EM structure of 2F-FPA-bound AftB

To elucidate the molecular basis of substrate recognition and catalysis, a cryo-EM structure of AftB was determined in complex with 2-F-farnesyl phosphoarabinose (2F-FPA), a donor analog for mycobacterial arabinosyltransferases (AraTs)⁵⁷. 2F-FPA substitutes the decaprenyl chain of DPA with a shorter, less hydrophobic farnesyl group, and incorporates a fluorine atom into the AraF ring to stabilize the glycosylphosphate bond (Supplementary Fig. 3d). This modification enhances its suitability for structural studies.

The structure was resolved to 3.4 Å using nanodisc-reconstituted AftB bound to 2F-FPA and Fab-B3 (Fig. 3 and Supplementary Fig. 4d–f). The entire protein was modeled except for 28 residues at the N-

terminus, 11 residues at the C-terminus, a disordered cytoplasmic loop region (319–335) connecting TM helices 8 and 9, and a small periplasmic loop region (432–437) between TM helix 11 and periplasmic helix α 1 (Supplementary Fig. 6). The Fab-B3 binding conformation is consistent with the apo structure (Supplementary Fig. 4d). The density for 2F-FPA was well resolved, revealing a partially curved conformation with the sugar ring projecting deep into the substrate cavity (Fig. 3a). The hydrophobic acyl chain extends along TM helix 6, while the phosphate moiety is coordinated by R221 and R372, with an additional hydrogen bond formed with T91, in the loop between JM helices 1 and 2 (Fig. 3b). T274 interacts with the fluorine atom on the arabinofuranose ring, and E172 forms a hydrogen bond with the C5 hydroxyl group of

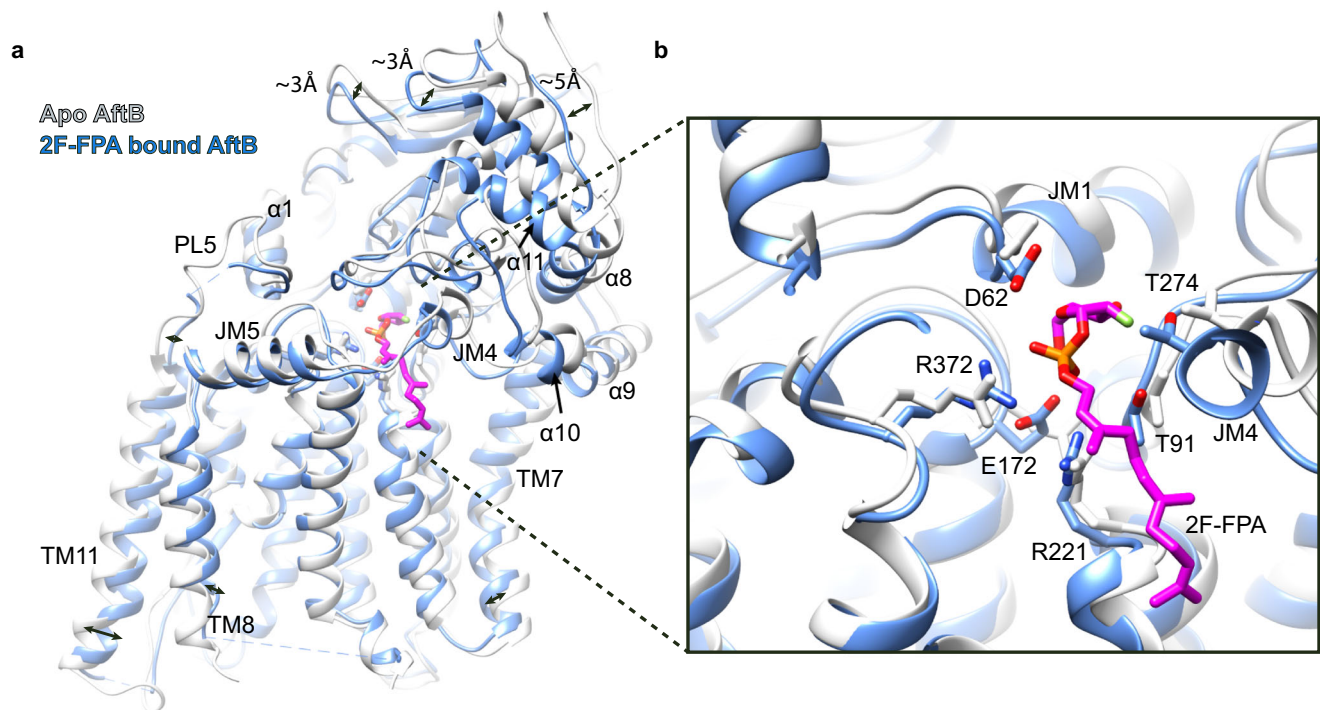


Fig. 4 | Conformational rearrangement of AftB upon 2F-FPA binding. **a** Superimposition of apo (white) and 2F-FPA bound AftB (blue). PD undergoes a clamp-like movement while TM helices exhibit minor inward movement.

b Magnified view of the substrate pocket in the apo structure (white) and the 2F-FPA-bound structure (blue), superimposed, showing interacting residues shifted toward 2F-FPA.

the sugar (Fig. 3b). All five of the interacting residues noted above are conserved across mycobacterial homologs of AftB (Supplementary Fig. 8). Furthermore, mutants R221A, R221E, R372A, and R372E were catalytically inactive in functional assays (Supplementary Fig. 10), despite being expressed at levels comparable to wild-type *McAftB* (Supplementary Fig. 9), underscoring the important role of these residues in AftB activity. This structure provides key insights into substrate binding and catalysis, highlighting the conserved nature of substrate coordination in mycobacterial AraTs.

Structural rearrangements upon substrate binding

Comparison of the apo and 2F-FPA-bound AftB structures shows conformational changes primarily within the PD. In the TM domain, TM helices 7 and 11 exhibit inward shifts, most pronounced in TM11. This is accompanied by a noticeable displacement of PL5, the loop connecting TM11 to $\alpha 1$ (Fig. 4). The shift in PL5 appears to act as a trigger for a conformational rearrangement in the PD, inducing a clamp-like movement of the domain towards the membrane (Fig. 4). This movement highlights the dynamic coupling between the TM domain and the PD during substrate binding.

Residues within the substrate-binding cavity—including D62, E172, R372, R221, T91, and T274—exhibit positional shifts towards the bound 2F-FPA, indicating a coordinated adaptation to facilitate substrate interaction (Fig. 4). These movements contribute to an inward pivoting of TM helices 1 and 7. The interplay between these structural rearrangements results in a more compact conformation of AftB.

Insights into DPA and acceptor binding mode with docking, co-folding, and MD simulations

To advance our understanding of the interaction between the enzyme and its substrates, we used a combination of ligand docking, protein-ligand co-folding, and MD simulations. Ligand docking studies, coupled with Chai-1⁵⁸ and RoseTTAFold⁵⁹ analyses of the donor substrate propose that the native DPA substrate binds similarly to the 2F-FPA in the AftB-2F-FPA complex. Specifically, the polar head of DPA is

oriented towards D62 within the active site (Supplementary Fig. 11b, d). The phosphate group of DPA is coordinated by Arg372 and Arg221, along with interactions involving T91 and T243. This mirrors the coordination seen with 2F-FPA (Supplementary Fig. 11d).

To build on the Ara_f binding site proposed by the earlier CG-MD simulations, we used a combination of ligand docking, with Autodock Vina⁶⁰, and protein-ligand co-folding, with both RoseTTAFold All-Atom⁵⁹ and Chai-1⁵⁸, to propose how the terminal-Ara_f and the subsequent products interact with AftB. The docking model suggests that two conserved residues, F448 and Y449, from the first periplasmic $\alpha 1$, are positioned near the branch point of terminal-Ara_f. Y449 may form hydrogen bonds with the substrate, while F448 could contribute to substrate coordination through non-polar interactions (Supplementary Fig. 11d). Additional coordinating residues include T524 in the loop between periplasmic $\beta 3$ and $\alpha 4$, D541 in the loop connecting periplasmic $\alpha 4$ to $\beta 4$, and R560 and H563 in the loops linking $\beta 4$ to $\alpha 5$ and $\alpha 5$ to $\alpha 6$ (Supplementary Fig. 11d). These residues appear to help position terminal-Ara_f close to the active site near DPA. The predicted importance of F448 and Y449 in acceptor substrate engagement is supported by the complete loss of enzymatic activity upon their mutation to alanine (Supplementary Fig. 10).

To further investigate the interaction of the substrates and products of this reaction with AftB, atomistic resolution MD simulations were performed. In these simulations, we used the native donor substrate DPA, its by-product decaprenyl-phosphate DP, and two polysaccharide mimics: terminal-Ara_f, as described above (representing the acceptor substrate) (Supplementary Fig. 11a), and a “terminal-Ara_f” (Supplementary Fig. 11c). The terminal-Ara_f represents the arabinofuranosylated product (Fig. 1a), characterized by the Ara_f motif at the end of the arabinan domain, including two additional α -(1→5)-Ara_f residues from the linear backbone (Fig. 1a and Supplementary Fig. 11c).

We conducted MD simulations with either the substrates (DPA and terminal-Ara_f) or the products (DP and terminal-Ara_f) present. In addition, we modeled the state where the sidechain of D62 was

arabinofuranosylated, representing an intermediate state. In all simulation set-ups, the terminal-Araf₄ and terminal-Araf₆ dissociated from their bound poses (Supplementary Fig. 12), likely due to the lack of extended repeating sugar units and lipid membrane anchors. Conversely, the lipid donor (DPA), the sidechain of the arabinofuranosylated D62, and the by-product (DP) remained stably docked within the binding pocket (Supplementary Fig. 12). The addition of either substrates (DPA with terminal-Araf₄) or products (DP and terminal-Araf₆) had no significant effect on the overall dynamics of the protein compared to the apo state (Supplementary Fig. 13a, b).

Discussion

AG and LAM/AM are complex polysaccharides that constitute major components of the mycobacterial cell envelope. We have determined the structures of AftB, an AraT responsible for adding terminal-Araf residues to both AG and LAM/AM, using cryo-EM in both its apo and 2F-FPA-bound state at resolutions of 2.9 Å and 3.4 Å, respectively.

These structures reveal that AftB belongs to the GT-C_A subclass of GT-C glycosyltransferases, characterized by a conserved structural core of seven TM helices. Notably, an irregular tube-shaped cavity is observed, spanning the two proposed substrate-binding sites. The structural characteristics of the cavity, combined with experimental evidence, provide insights into substrate binding and enable us to propose a mechanism of action for AftB.

Catalytic mechanism of AftB

Glycosyltransferases are classified as inverting or retaining based on the stereochemical outcome of the glycosyl transfer^{51,61}. The former inverts the configuration of the anomeric carbon of the donor substrate in the resulting product^{48,62}, while the latter maintains the configuration^{48,62,63}. AftB catalyzes the transfer of D-Araf from the donor DPA to the acceptor, forming a β-(1 → 2) glycosidic bond^{35,36,64} (Fig. 1a). This reaction preserves the β-configuration of the anomeric carbon, identifying AftB as a retaining glycosyltransferase^{28,35,36,64}.

The catalytic mechanism of retaining glycosyltransferases is a subject of debate, with two main mechanisms proposed: the front-face S_Ni (substitution nucleophilic internal) mechanism^{62,63} or a double-displacement mechanism^{51,62,65}. While most structurally characterized retaining glycosyltransferases lack a suitably positioned catalytic nucleophile, favoring the S_Ni mechanism^{62,63,65}, recent structural and biochemical studies of^{66,67} and KpsC⁶⁸, two β-Kdo transferases, have provided clear evidence for double-displacement mechanisms. These findings suggest this catalytic strategy may be more widespread among retaining glycosyltransferases than previously.

Our structural analysis reveals that D62, located within the conserved DD motif, is positioned ~3.8 Å away from the anomeric carbon (C1) of 2F-FPA, making it well-placed to function as a catalytic nucleophile. This distance is identical to that between the analogous amino acid and the donor anomeric center in the donor-bound complex of WbbB^{61,66}. In addition, structural and computational studies on retaining glycoside hydrolases, which proceed via a double-displacement mechanism, have shown that the analogous nucleophile is typically positioned at distances of 3.2–3.9 Å from the anomeric center^{69–72}, consistent with our observations in the 2F-FPA bound structure. Supporting the critical role of D62, mutating it to alanine abolishes enzymatic activity. Furthermore, MD simulations indicate that the pK_a of D62 decreases substantially in the presence of substrates or products (Supplementary Fig. 13c). This effect is not observed for other substrate-interacting residues (Supplementary Fig. 13c), which further emphasizes the enzymatic importance of D62. Based on these findings, we propose that at physiological pH, D62 would predominantly be in its deprotonated form, ready to perform a nucleophilic attack on the anomeric carbon (C1) of the DPA substrate.

Similar to what we observed in our 2F-FPA-bound structure, the MD contact analysis (Supplementary Fig. 12d) revealed that R221 forms

significant interactions with DPA, and both R221 and R372 show strong contacts with DP during simulations. Unlike D62, the pK_a values of R221 and R372 remain relatively stable across different simulation conditions (Supplementary Fig. 13c), suggesting these residues maintain their protonated, positively charged states. Mutations of these conserved residues lead to a complete loss of AftB function (Supplementary Fig. 10), further underscoring their essential roles in the catalytic process.

Given these insights and AftB's retaining nature, we propose a double-displacement mechanism. The catalytic cycle begins with D62 initiating nucleophilic attack on the anomeric carbon of DPA (Fig. 5b). This results in the formation of a covalent α-Araf-enzyme intermediate. This step inverts the stereochemistry of the Araf residue to be transferred. Concurrently, R221 and R372 stabilize the departing phosphate group of DPA which then deprotonates the C2 hydroxyl group (or polarizes the O–H bond) of the acceptor substrate (Fig. 5b), priming it for a nucleophilic attack.

The second step involves the C2 hydroxyl group of the acceptor attacking the anomeric carbon of the covalent α-Araf-enzyme intermediate, breaking the arabinose-enzyme bond and forming a new β-(1 → 2) glycosidic bond. This second nucleophilic attack restores the original β-configuration of the donor Araf moiety, consistent with the retaining nature of AftB (Fig. 5b).

Although attempts to capture a ternary structure with donor and acceptor substrates were unsuccessful, structural and functional evidence supports a double-displacement mechanism for AftB. This mechanism accommodates AftB's bulky polysaccharide substrates, which may present steric constraints incompatible with an S_Ni mechanism.

AftB represents a structurally characterized example of a retaining GT-C fold enzyme. This structure provides insights into retaining GT-C fold enzymes, though the absence of other structural examples in this class currently precludes more extensive comparative analysis.

Finally, no metal ions were observed in the active site in our cryo-EM structures of AftB. This observation aligns with findings from several other GT-C_A enzymes; for instance, studies on both AftA and ALG6 have indicated that their activities do not require metal ions^{30,55}. Similarly, our recent study on PimE⁷³, a mycobacterial mannosyltransferase also belonging to the GT-C_A superfamily, demonstrated that PimE is metal-independent⁷³. Nonetheless, future studies are needed to definitively determine if metal ions play any role in AftB function.

Structural insights into the function of AftB

The donor substrate of AftB, decaprenyl-phosphoryl-arabinose (DPA), is a lipid-linked arabinose (Araf) anchored to the membrane via a decaprenyl-phosphate moiety. Structural analysis reveals that the lipid tail of DPA extends out of the catalytic cavity towards TM helices 6 and 7, embedding into the lipid bilayer, while orienting the reactive Araf-phosphate moiety towards the catalytic core.

The acceptor substrates for AftB are the polysaccharides AG and LAM/AM polysaccharides, both anchored to the membrane via a galactan or mannan core, respectively. These polysaccharides share conserved motifs at the non-reducing end of their arabinan domains, which serve as the acceptor sites for AftB. In AG, AftB features branched motifs at its arabinan termini, where AftB generates the characteristic Araf₆ structure required for mycolic acid attachment. In contrast, LAM/AM biosynthesis involves both branched and linear arabinan regions, with AftB generating a terminal β-linked Araf₆ structure, as well as linear Araf₄ structures (Supplementary Fig. 1a).

The irregular tube-shaped catalytic cavity of AftB features a relatively spacious periplasmic-facing opening. This appears to be adapted to accommodate both branched and linear terminal arabinan acceptor motifs, enabling AftB to complete the final steps of both AG and LAM/AM biosynthesis. Further structural studies of AftB in complex with its

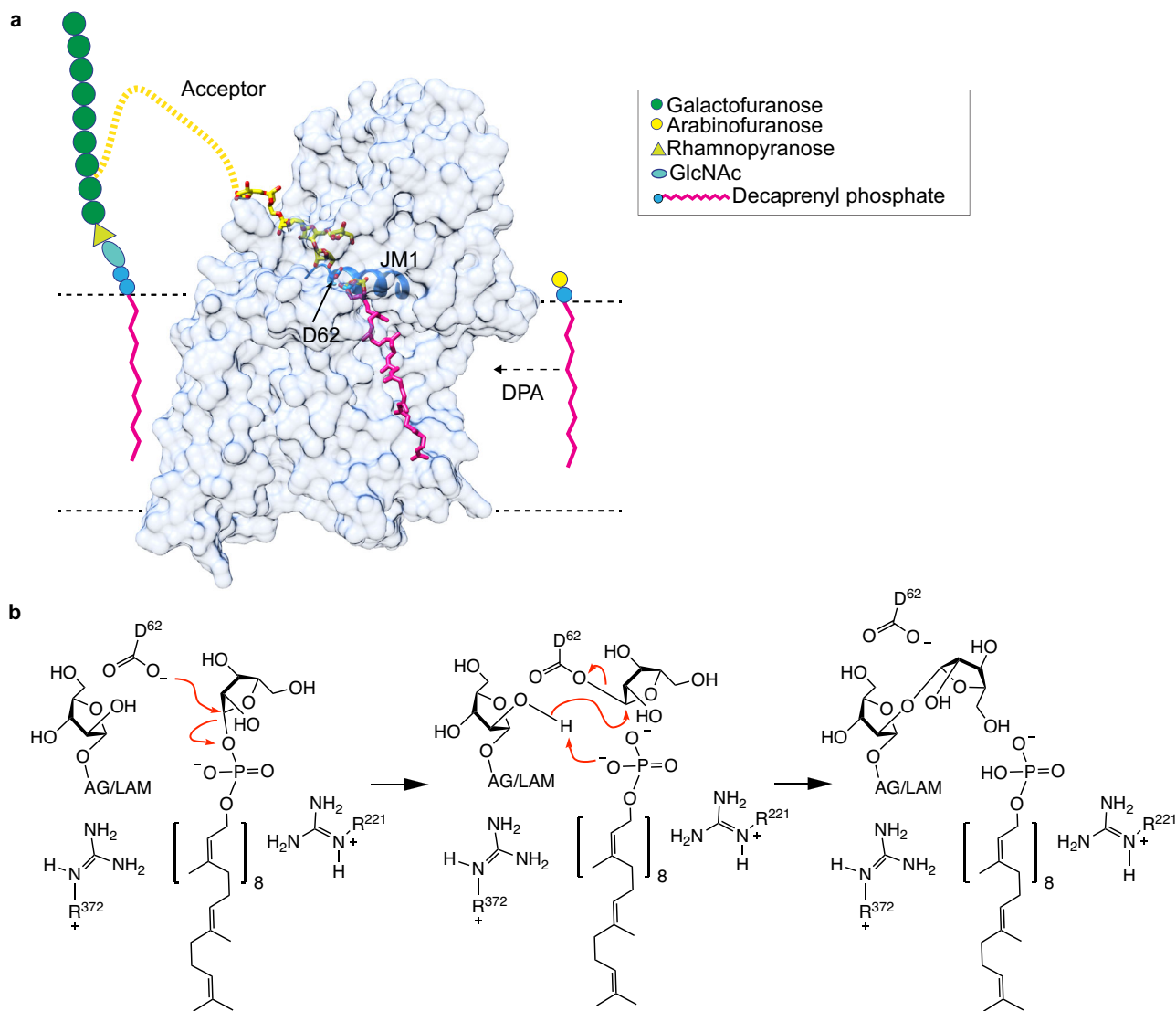


Fig. 5 | Proposed mechanism for AftB-catalyzed retaining arabinosyl transfer. **a** Schematic representation of AftB's role in the biosynthesis of arabinogalactan. AftB directs the addition of arabinofuranosyl units from the donor molecule DPA

onto the growing arabinan chain, maintaining the glycosidic bond stereochemistry. **b** Proposed catalytic model of AftB.

diverse acceptor substrates will enhance our understanding of substrate recognition and catalysis.

Structural conservation across mycobacterial species

To evaluate the broader relevance of our *M. chubuense* AftB structure, structural comparisons were conducted with AlphaFold2 predictions of AftB orthologues from *M. tuberculosis*, the causative agent of tuberculosis, and *M. smegmatis*, a well-established model organism in mycobacterial research. The analysis revealed significant structural conservation across the species, with root-mean-square-deviation (RMSD) values of 2.8 Å (across 606 pairs) for *M. tuberculosis* AftB and 2.4 Å (across 617 pairs) for *M. smegmatis* AftB (Supplementary Fig. 14a, b). Sequence identity was also notably high, with 71% identity between *M. chubuense* and *M. tuberculosis* AftB, and 74% between *M. chubuense* and *M. smegmatis* AftB (Supplementary Fig. 14c). Key catalytic and substrate-binding residues identified in the *M. chubuense* AftB structure are identically positioned across the orthologues. This high degree of conservation strongly suggests that the structural and functional insights gained from the *M. chubuense* AftB structure can be extended to homologs from other mycobacterial species.

Structural comparison with other mycobacterial arabinosyltransferases

The biosynthesis of mycobacterial arabinan involves a cascade of AraTs, each with distinct roles, with AftB adding the terminal β -(1 \rightarrow 2) linkages via a retaining mechanism. Despite sharing a conserved GT-C_A fold, these enzymes exhibit architectural variations that may be tailored to their specific functions. All these AraTs feature a conserved catalytic aspartate residue, but its position differs (Supplementary Fig. 3). In AftB, this residue is predicted to act as a nucleophile in a retaining double-displacement mechanism, whereas in other AraTs it serves as a general base to facilitate an inverting reaction.

AftB is characterized by an irregular, tube-shaped. The architecture of this cavity may be to accommodate AftB's dual role in binding both AG and LAM/AM acceptor substrates. AftA uses the galactan as its acceptor substrate, initiating arabinan synthesis, with structural features that appear to be adapted for selective galactan recognition³⁰ (Supplementary Fig. 3d). AftD presents a prominent cleft spanning its periplasmic and transmembrane domains, accompanied by an additional groove on its periplasmic face⁷⁴ (Supplementary Fig. 3f), in contrast to the more enclosed cavities observed in other AraTs.

For EmbA/B/C, the DPA donor substrate binds with its hydrophobic decaprenyl tail situated within a groove formed by trans-membrane helices, while the sugar head group resides near the active site cavity^{24,75} (Supplementary Fig. 3h). This binding mode is aligned with the observed DPA binding mode for AftB. However, the Emb enzymes exhibit a relatively constricted connection between their donor and acceptor binding sites.

AftD and EmbA/B/C possess a highly positively charged surface on the cytoplasmic side that facilitates ACP binding and a hydrophobic tunnel for accommodating the 4'-phosphopantetheine prosthetic group of ACP^{24,74,75}. These motifs suggest an evolutionary adaptation for specialized functions, including co-functioning with ACP partner proteins³⁰. In contrast, ACP is absent in the structures of AftB and AftA, although they both feature positively charged surfaces on their cytoplasmic domains (Supplementary Fig. 4). These structural distinctions likely reflect the specialized catalytic roles and regulatory requirements of AraTs at different stages of arabinan biosynthesis. Further studies are required to elucidate the precise mechanisms underlying these variations.

Therapeutic potential of targeting AftB

The mycobacterial cell wall, containing the key structural components AG and LAM/AM, is critical for pathogen resilience and virulence. Inhibiting the biosynthesis of these components, as demonstrated by the front-line anti-TB drug ethambutol targeting EmbB and EmbC, is a proven therapeutic strategy^{21,76}. AftB catalyzes the final steps in AG and LAM/AM biosynthesis and plays an essential role in mycobacterial viability³⁵. Depleting AftB has been shown to sensitize *M. tuberculosis* to rifampicin, a front-line TB drug targeting bacterial RNA polymerase, particularly under infection-relevant conditions³⁹. This highlights AftB as a potential target for anti-TB drug development^{35,36,39}. Additionally, conservation of AftB across pathogenic mycobacterial species, including non-tuberculous mycobacteria (NTM) such as *M. chubuense* (studied herein), *M. abscesses*, and *M. chelonae*⁷⁷⁻⁷⁹ (Supplementary Fig. 8 and Supplementary Fig. 14), suggests its potential for broad-spectrum therapeutic applications. Targeting AftB could thus provide avenues to combat both drug-resistant and drug-sensitive mycobacterial pathogens^{26,78}.

Our structural studies provide insights into the substrate-binding cavity and active site architecture of AftB. Functional assays and computational analyses presented here provide further insight into how substrates interact with conserved residues within the cavity. This cavity could serve as a target for competitive inhibitors that either directly inhibit AftB or enhance the effectiveness of existing antibiotics like rifampicin. While future studies will be needed to fully explore the druggability of this cavity, our integrated characterization of AftB provides a foundation for further investigation of this essential enzyme as a therapeutic target⁷⁷⁻⁷⁹.

Methods

Overexpression and purification of *Mycobacterium chubuense* AftB

Wild-type *Mycobacterium chubuense* AftB was cloned into pNY-COMPSC23 plasmid⁸⁰. The procedure for design and synthesis of pNYCOMPSC23 plasmid was described in detail in a previous protocol⁸⁰. The plasmids were transformed into BL21 (DE3) pLys *E. coli* for protein expression. The following day, a single colony was selected to inoculate a starter culture consisting of 50 mL Terrific Broth (TB) medium (Fisher), supplemented with 50 µg/mL ampicillin and 35 µg/mL chloramphenicol. The starter culture was grown overnight at 37 °C with shaking at 220 rpm in an incubator shaker (New Brunswick Scientific). The next day, four flasks, each containing 1 L of TB medium with 50 µg/mL ampicillin and 35 µg/mL chloramphenicol, were inoculated with 10 mL of the starter culture and grown at 37 °C with shaking at 220 rpm until the OD₆₀₀ reached 0.6–0.8. The

cultures were then cooled to 22 °C, and protein expression was induced by adding 0.5 mM IPTG for 16 h. Cells were harvested by centrifugation at 4000 × *g* in H6000A/HBB6 rotor (Sorvall) for 30 min at 4 °C, and the pellet was resuspended with lysis buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 20 mM MgSO₄, 10 mg/mL DNase I (Roche), 8 mg/mL RNase A (Roche), 1 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP), 1 mM PMSF, 1 tablet/1.5 L buffer EDTA-free cOmplete protease inhibitor cocktail (Roche). Cells in lysis buffer were lysed by passing through a chilled Emulsiflex C3 homogenizer (Avestin) three times. The lysate was centrifuged at 3000 × *g* in a Centrifuge 5810 R (Eppendorf) at 4 °C for 5 min to remove cell debris and non-lysed cells. To isolate the cell membrane, the supernatant was ultracentrifuged in a Type 45 Ti Rotor (Beckman Coulter) at 185,600 × *g* for 1 h. For 4 L of culture, the yield was ~4 g of membrane pellet, which was resuspended in the lysis buffer up to 80 mL and homogenized using a handheld glass homogenizer (Kontes) on ice. The membrane fraction was stored at –80 °C until further use. The membrane fraction was thawed and solubilized by adding *n*-dodecyl-β-D-maltopyranoside (DDM) to a final concentration of 1% (w/v) and was rotated gently at 4 °C for 2 h. Insoluble material was removed by ultracentrifugation at 185,600 × *g* in a Type 45 Ti Rotor at 4 °C for 40 min. The supernatant was added to preequilibrated Ni²⁺-NTA resin (QIAGEN) in the presence of 40 mM imidazole and incubated with gentle rotation at 4 °C for 2 h.

The resin was then washed with 10 column volumes of wash buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 65 mM imidazole, 0.1% DDM and eluted with 3 column volumes of elution buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 300 mM imidazole, 0.03% DDM. The eluted protein was exchanged into a buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 0.03% DDM using a PD-10 desalting column (GE).

To characterize the expression and folding properties of *McaftB* mutants relative to wild-type protein, both wild-type and mutant *McaftB* were expressed from 250 mL cultures following identical procedures. All purification steps remained the same as described above. After buffer exchange using the PD-10 desalting column, both wild-type and mutant proteins were analyzed by size-exclusion chromatography using a Superose™ 6 Increase 5/150 GL column to assess their oligomeric state and folding properties.

Preparation of mycobacterial membranes for arabinosyltransferase assay

Wild-type *M. smegmatis* mc²155 cells were grown in 50 mL of Luria-Bertani (LB) broth for 72 h at 37 °C and 200 rpm. An Erlenmeyer flask with 2 L of LB broth was inoculated with 10 mL of the saturated starter culture and incubated for 48 h at 37 °C and 130 rpm. Cells were harvested by centrifugation at 2600 × *g* for 15 min at 4 °C, resuspended in ice-cold buffer A [50 mM MOPS (pH 7.9), 5 mM 2-mercaptoethanol and 10 mM MgCl₂], and disrupted by passing the suspension through a high-pressure homogenizer at 20,000 psi. The lysate was centrifuged at 27,000 × *g* for 30 min at 4 °C. The cell wall pellet was removed, and the supernatant was recentrifuged at 100,000 × *g* for 2 h at 4 °C. The supernatant was discarded, and the pellet of enzymatically active membrane was gently resuspended in 1 mL of buffer A with gentle homogenization using a Dounce homogenizer. The protein concentration of the membrane fractions was determined using the Pierce BCA protein assay kit and was typically between 15 and 20 mg/mL.

Arabinosyltransferase assay

The arabinosyltransferase activity was assessed using wild-type *M. smegmatis* membranes or *E. coli* membranes expressing *McaftB* as described above. Untransformed *E. coli* membranes were used as a negative control. A typical reaction mixture contained FPA (1 mM), synthetic acceptor (0.2 mM), ATP (1 mM), DMSO (2% v/v), buffer A [50 mM MOPS (pH 7.9), 5 mM 2-mercaptoethanol and 10 mM MgCl₂]

and the respective membranes (2 mg) in a total volume of 200 μL . The reaction mixture was incubated at 37 °C overnight and then terminated by adding 200 μL of ethanol. The resulting mixture was centrifuged at 20,000 $\times g$ for 15 min, and the supernatant was loaded onto a preequilibrated strong anion exchanger (SAX) cartridge. The cartridge was eluted with 5 mL of 50% ethanol–water. The eluate was evaporated to dryness, and the dried enzymatic products were treated with Ac_2O (100 μL) in pyridine (100 μL). The reaction mixture was stirred overnight at room temperature before being concentrated and then partitioned between the two phases (1:1) of chloroform and water. The organic layer was concentrated before being analyzed by MALDI mass spectrometry. Samples for MALDI analysis were prepared by 10-fold dilution in methanol or water. The MALDI matrix was prepared by dissolving 2,5-dihydroxybenzoic acid (DHB, 10 mg) in 1 mL of 50% ACN. Next, spotting 1 μL of DHB matrix and then 1 μL of the diluted sample onto the MALDI target plate was followed by mixing by gentle pipetting on the plate. Samples were allowed to dry completely (15–20 min) before analysis. MALDI-MS measurements were carried out on a Bruker ULTRAFLEXTRME instrument in reflector mode. The mass range investigated was 600–3500 m/z in positive ion mode using 500 Da matrix suppression. The laser was operated at a frequency of 50 Hz using the “large” predefined shot pattern and 4000 laser shots were collected per spectrum. The following settings were used: Ion Source 1: 20 kV, Ion Source 2: 17.75 kV, Lens: 7 kV, Reflector: 20.8 kV, Reflector 2: 10.95 kV. Data analysis was done using the following software: Compass for flexSeries 1.4, flexControl Version 3.4, and flexAnalysis Version 3.4.

MS/MS on Araf₅ and Araf₆ products from enzymatic reactions

Enzymatic synthesized products were permethylated and analyzed by MALDI-MS/MS in reflector positive ion mode as described previously^{42,43}, but using instead an AB SCIEX MALDI TOF/TOF 5800 system. MS/MS spectra were acquired with a laser intensity at 5000 and 1 kV CID energy. Five MS/MS spectra (1000 laser shots/spectrum) were accumulated to yield the final spectra (total laser shots = 5000) that were assigned manually according to the fragmentation pattern established in previous analyses of permethylated arabinan oligomers^{42,43}.

Nanodisc reconstitution of AftB

The protein was incorporated into lipid nanodiscs with a molar ratio 1:5:250 of membrane scaffold protein IE3D1 (MSPIE3D1): (POPG) 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (Avanti), and incubated for 2 h with gentle agitation at 4 °C. For the substrate-bound structure, 2F-FPA was added to the reconstitution mixture at a molar ratio of 1:3 (AftB: 2F-FPA). The POPG lipid was prepared by adding the solid extract to deionized water to a final concentration of 20 mM. The mix was put on ice and then gently sonicated with a tip sonicator (Fisher Scientific) to dissolve the lipids until the mixture became semitransparent.

To initiate nanodisc reconstitution, 100 mg Biobeads (Bio-Rad) per mL of protein solution was added to the mixture and incubated by gentle rotation at 4 °C overnight. Biobeads were removed the next day by filtering the reconstitution mixture through an Ultrafree centrifugal filter unit (Fisher) at 16,100 $\times g$ in a Centrifuge 5415 R (Eppendorf) at 4 °C for 1 min. The reconstitution mixture was rebound to Ni^{2+} -NTA resin in the presence of 25 mM imidazole for 2 h at 4 °C in order to remove free nanodisc. The resin was washed with 10 column volumes of wash buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, and 50 mM imidazole, followed by 4 column volumes of elution buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, and 300 mM Imidazole. The eluted protein was subsequently purified by size-exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 GL column in buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl.

Cryo-EM sample preparation

Fractions containing AftB incorporated into nanodiscs were pooled and incubated with the Fab at 4 °C for 2 h in a 1:3 molar ratio of protein to Fab. The protein mixture was further purified using a Superdex 200 Increase 10/300 GL SEC column in SEC buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl.

For the apo structure, peak fractions were pooled and concentrated to 7 mg/mL using a 50 kDa cut-off filter concentrator (Amicon). The sample was frozen using a Vitrobot (Thermo Fisher) by applying 3 μL to a plasma cleaned (Gatan Solarus) 0.6/1-mm holey gold grid (Quantifoil UltrAuFoil). After a 30 s incubation, the grids were blotted using 595 filter paper (Ted Pella, Inc) for 8 s before being immediately plunged into liquid ethane for vitrification. The plunger was operated at 4 °C with greater than 90% humidity to minimize evaporation and sample degradation. For the substrate-bound structure, peak fractions were pooled and concentrated to concentration of 5 mg/mL. The sample was incubated with 2F-FPA at a 1:2 molar ratio of protein to substrate for 1 h before being applied onto the grids. The grids were prepared in the same way as for the apo sample, but with a blotting time of 6 s.

Data collection

For the apo structure, images were recorded at the Columbia University Cryo-Electron Microscopy Center on a Titan Krios electron microscope (FEI), equipped with an energy filter and a K3 direct electron detection filter camera (Gatan K3-BioQuantum) using a 0.87 Å pixel size. An energy filter slit width of 20 eV was used during the collection and was aligned automatically every hour using the Leginon software package⁸¹. Data collection was performed using a dose of around 58 e^- per \AA^2 across 50 frames (50 ms per frame) at a dose rate of $-16 e^-$ per pixel per second, using a set defocus range of $-1.2 \mu\text{m}$ to $-2.2 \mu\text{m}$. A total of 7164 micrographs were collected over a single 2-day session.

For the substrate-bound structure, images were recorded at the New York Structural Biology Center (NYSBC) on a Titan Krios (FEI; NYSBC Krios 2) operating at 300 kV equipped with a spherical aberration corrector, an energy filter (Gatan GIF Quantum), and a post-GIF K2 Summit direct electron detector, using a 0.8460 Å pixel size. An energy filter slit width of 20 eV was used during the collection and was aligned automatically every hour using the Leginon software package⁸¹. Data collection was performed using a dose of around 50.3 e^- per \AA^2 across 24 frames (50 ms per frame) at a dose rate of $-41 e^-$ per pixel per second, using a set defocus range of $-0.8 \mu\text{m}$ to $-2.5 \mu\text{m}$. A total of 16,290 micrographs were collected over a single session of 3 days and 20 h (92 h total).

Identification of McAftB-specific Fab using phage display

AftB was reconstituted into chemically biotinylated MSPIE3D1 as previously described^{44,82}. Selection for Fabs was performed starting with Fab Library E^{45,83}. Targets and the library were first diluted in selection buffer (20 mM HEPES, pH 7.4, 150 mM NaCl, and 1% BSA). Five rounds of sorting were performed using a protocol adapted from published protocols^{84,85}. In the first round, bio-panning was performed manually using 400 nM of AftB, which was first immobilized onto magnetic beads and washed three times with selection buffer. The library was incubated for 1 h with the immobilized target, beads were subsequently washed three times with selection buffer, and then beads were used to directly infect log-phase *E. coli* XL-1 Blue cells. Phage were amplified overnight in 2XYT media supplemented with ampicillin (100 $\mu\text{g}/\text{mL}$) and M13-K07 helper phage (10^9 pfu/mL). To increase the stringency of selection pressure, four additional rounds of sorting were performed by stepwise reduction of the target concentration: 200 nM in the 2nd round, 100 nM in the 3rd round, and 50 nM in the 4th and 5th rounds. These rounds were performed semi-automatically using a KingFisher magnetic beads handler (Thermo Fisher Scientific).

For each round, the amplified phage population from each preceding round was used as the input pool. Additionally, amplified phage was precleared prior to each round using 100 μ L of streptavidin paramagnetic particles, and 2.0 μ M of empty MSP1E3D1 nanodiscs were used throughout the selection as competitors in solution. For rounds 2–5, prior to infection of log-phage cells, bound phage particles were eluted from streptavidin beads by 15 min incubation with 1% Foscholine-12 (Anatrace).

Single-point phage ELISA to validate Fab binding to AftB

96-well plates (Nunc) were coated with 2 μ g/mL Neutravidin and blocked with selection buffer. Colonies of *E. coli* XL-1 Blue cells harboring phagemids from the 4th and 5th rounds were used to inoculate 400 μ L 2XYT media supplemented with 100 μ g/mL ampicillin and 10⁹ pfu/mL M13-KO7 helper phage, and phage were subsequently amplified overnight in 96-well deep blocks with shaking at 280 rpm. Amplifications were cleared of cells with a centrifuge step and then diluted 10-fold into ELISA (selection buffer with 2% BSA). All phage were tested against wells with immobilized biotinylated MSP1E3D1-reconstituted AftB (30 nM), empty biotinylated-MSP1E3D1 nanodiscs (50 nM), or buffer alone to determine specific target binding. Phage ELISA was subsequently performed as previously described^{82,84} where the amount of bound phage was detected by colorimetric assay using an anti-M13 HRP-conjugated monoclonal antibody (GE Healthcare; Code: 27-9421-01; 1:5000 dilution). Binders with high target and low non-specific signal were chosen for subsequent experiments.

Fab cloning, expression, and purification

Specific binders based on phage ELISA results were sequenced at the University of Chicago Comprehensive Cancer Center DNA Sequencing facility and unique clones were then sub-cloned into the Fab expression vector RH2.2 (kind gift of S. Sidhu) using the In-Fusion Cloning kit (Takara). Successful cloning was verified by DNA sequencing. Fabs were then expressed and purified as previously described⁸². Following purification, Fab samples were verified for purity by SDS-PAGE and subsequently dialyzed overnight in 20 mM HEPES, pH 7.4, 150 mM NaCl.

Assessment of Fab binding affinity to AftB

To measure the apparent binding affinity, multi-point ELISAs using each purified Fab were performed in triplicate. Briefly, AftB (30 nM) or empty biotinylated-MSP1E3D1 nanodiscs (50 nM) were immobilized onto 96-well plates coated with Neutravidin (2 μ g/mL). Fabs were diluted serially 3-fold into ELISA buffer using a starting concentration of 3 μ M, and each dilution series was tested for binding to wells containing either AftB, empty nanodiscs, or no target at all. The Fab ELISA was subsequently performed as previously described⁸⁴, where the amount of bound Fab was measured by a colorimetric assay using an HRP-conjugated anti-Fab monoclonal antibody (Jackson ImmunoResearch). Measured A_{450} values were plotted against the log Fab concentration, and EC_{50} values were determined in GraphPad Prism version 8.4.3 using a variable slope model assuming a sigmoidal dose response.

Single-particle cryo-EM data processing and map refinement

All data sets were corrected for beam-induced motion with Patch Motion Correction implemented in cryoSPARC v.2.15⁸⁶ and the contrast transfer function (CTF) was estimated with Patch CTF. For the apo structure, 5.03 million particles were automatically picked using a blob-picker job and subjected to multiple rounds of 2D classification. Representative 2D classes of 118,434 particles clearly showing the Fab-bound complex in the side views were selected as input for Topaz training. The resulting model was used to pick particles using Topaz Extract. Initially, 2,790,710 particles were extracted with a box size of 320 pixels and binned four times. Multiple rounds of 2D classification were performed to clean up the particles using a batchsize per class of

400 and “Force Max over poses/shifts” turned off with 40 online-EM iterations. Classes of total 335,357 particles which display clear features of a Fab-bound nanodisc-embedded membrane protein were reextracted using a 360-pixel box size without binning. Ab initio reconstruction was performed in cryoSPARC v.2.15 using three classes and a class similarity parameter of 0.1. One good class comprised of 241,095 particles was subjected to heterogeneous refinement and a final class comprising 192,312 particles was subjected to Nonuniform refinement and yielded a reconstruction with a resolution of 3.0 Å (FSC = 0.143). A subsequent local refinement was performed using a mask covering AftB and the variable region of the Fab and resulted in a density map at 2.85 Å resolution.

For the substrate-bound structure, 5.29 million particles were automatically picked using a blob-picker job and subjected to multiple rounds of 2D classification. Representative 2D classes of 10,186 particles clearly showing the Fab-bound complex in the side views were selected as input for Topaz training. The resulting model was used to pick particles using Topaz Extract. Initially, 2.01 million particles were extracted in a box of 320 pixels and Fourier cropped to 80 pixels for initial cleanup. Multiple rounds of 2D classification were performed to clean up the particles using a batchsize per class of 400 and “Force Max over poses/shifts” turned off with 40 online-EM iterations. Classes of total 279,138 particles which display clear features of a Fab-bound nanodisc-embedded membrane protein were reextracted using a 360-pixel box size without binning. Ab initio reconstruction was performed in cryoSPARC v.2.15 using two classes and a class similarity parameter of 0.1. One good class comprised of 183,657 particles was subjected to heterogeneous refinement and a final class comprised of 146,894 particles was subjected to Nonuniform refinement and yielded a reconstruction with a resolution of 3.8 Å (FSC = 0.143). A subsequent local refinement was performed using a mask covering AftB and the variable region of the Fab and resulted in a density map at 3.40 Å resolution.

Model building

All model building was performed in Coot⁸⁷. The apo AftB model was constructed de novo from the globally sharpened 2.85 Å map using the Phenix Map to Model tool⁸⁸. Manual segment joining and residue assignment were guided by Xtalpred’s secondary structure prediction⁸⁹. Ramachandran outliers were fixed manually in Coot before the structure was refined with Phenix real space⁸⁸ refine with secondary structure and Ramachandran restraints

For substrate-bound AftB, the ligand files were generated in Coot, and similarly to the apo structure, manual model building, and refinement were conducted iteratively with Coot⁸⁷ and Phenix³⁹. The structural quality of both apo and substrate-bound forms was assessed using Molprobit⁹⁰. For most of the Fab region, with the exception of the binding interface, modeling was based on an existing high-resolution structure (PDB ID 5UCB [<https://doi.org/10.2210/pdb5UCB/pdb>]).

Molecular dynamics simulations

The CG parameters for DPA and terminal-Araf₄ were generated based on previously published lipid⁹¹ and Martini 3 parameters^{92–94}. The values for bond lengths and angles were generated from AT simulation data, where the molecules of interest were parameterized using CGenFF⁹⁵, and PyCGTOOL⁹⁶ was used to get exact values.

For CG simulations, the apo structure (with or without the Fab bound) was converted to the Martini 3 forcefield using martinize2⁹⁷ including a 1000 kJ mol⁻¹ nm⁻² elastic network and default protonation states. The structure was then embedded in a PE:PG (80:20) membrane using the insane script⁹⁸. Each simulation box was then solvated and neutralized with 150 M NaCl. For DPA simulations, 2% DPA was added to the upper leaflet of the membrane. For the simulations including the terminal-Araf₄, two copies of coarse-grained substrate were added to

the solvent in a random orientation and positioning. In each simulation, the system was comprised of ~40,000 particles (~30,000 of these being water molecules) in a $14 \times 18 \times 18 \text{ nm}^3$ box. Systems were energy minimized using the steepest descents method. Production simulations used a timestep of 20 fs, and five simulations of 3 μs were performed for both DPA and terminal-Araf₄ containing systems. The Parrinello-Rahman barostat⁹⁹ was set at 1 bar and the velocity-rescaling thermostat¹⁰⁰ was used at 310 K. The Reaction-Field algorithm was used for electrostatic interactions with a cut-off of 1.1 nm. A single cut-off of 1.1 nm was used for van der Waals interactions. All simulations were performed using GROMACS 2021.4^{101,102}.

Coordinates for the atomistic systems were generated from the end snapshots of the CG simulations using CG2AT2¹⁰³. Coordinates for both the substrates and products were built using the densities from the ligand-bound cryo-EM structure. For the substrate state the bound DPA analog was modified to DPA and energy minimized. RoseTTAFold All-atom⁵⁹ was used to fold and dock both substrates and products for comparison. Atomistic molecular simulations were performed with the CHARMM36m forcefield¹⁰⁴, with parameters for the terminal-Araf₄ and terminal-Araf₆ (substrates and products respectively) created using CHARMM-GUI¹⁰⁵. Parameters for the polyprenyl phosphate and DPA were created using CHARMM-GUI and combined with previous parameters from the past studies^{106,107}. Energy minimizations of the systems were performed using the steepest descents method. A timestep of 2 fs was used for production simulations, with three independent simulations of 500 ns for substrate, product, Fab bound, and apo states of AftB. The C-rescale barostat¹⁰⁸ was set at 1 bar and the velocity-rescaling thermostat¹⁰⁰ was used at 310 K. The PME algorithm was used for electrostatic interactions with a cut-off of 1.2 nm. A single cut-off of 1.2 nm was used for van der Waals interactions. All simulations were performed using GROMACS 2021.4^{101,102}. In each simulation, the system was comprised of ~222,000 particles (~50,000 of these being water molecules) in a $13 \times 15 \times 10 \text{ nm}^3$ box.

For analysis, PyLipID¹⁰⁹ was used to measure interactions between the substrates and the protein. PLUMED¹¹⁰ was used in conjunction with Matplotlib¹¹¹ to generate the density plots. The pK_a of the residues were measured over the course of the simulations with PROPKA3¹¹² and propkatraj (version 1.1.0). Visualization of the duration was performed with Visual Molecular Dynamics (VMD)¹¹³.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The MD data can be found at: <https://doi.org/10.5281/zenodo.14916580>. The cryo-EM maps and atomic coordinates have been deposited through the wwPDB OneDep system, with map accession numbers EMD-46976 (apo AftB) and EMD-46978 (2F-FPA-bound AftB). The atomic coordinate accession numbers are 9DLF (apo AftB) and 9DLH (2F-FPA-bound AftB). The PDB code of the previously published structure used in this study is 5UCB. Source data are provided as a Source Data file. Source data are provided with this paper.

References

- World Health Organization. *Global Tuberculosis Report* (WHO, 2023).
- McNeil, M. R. & Brennan, P. J. Structure, function and biogenesis of the cell envelope of mycobacteria in relation to bacterial physiology, pathogenesis and drug resistance; some thoughts and possibilities arising from recent structural information. *Res. Microbiol.* **142**, 451–463 (1991).
- Kalscheuer, R. et al. The Mycobacterium tuberculosis capsule: a cell structure with key implications in pathogenesis. *Biochem. J.* **476**, 1995–2016 (2019).
- Batt, S. M., Burke, C. E., Moorey, A. R. & Besra, G. S. Antibiotics and resistance: the two-sided coin of the mycobacterial cell wall. *Cell Surf.* **6**, 100044 (2020).
- Abrahams, K. A. & Besra, G. S. Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology* **145**, 116–133 (2018).
- Jankute, M., Cox, J. A. G., Harrison, J. & Besra, G. S. Assembly of the mycobacterial cell wall. *Ann. Rev. Microbiol.* **69**, 405–423. <https://doi.org/10.1146/annurev-micro-091014-104121> (2015).
- Alderwick, L. J., Harrison, J., Lloyd, G. S. & Birch, H. L. The mycobacterial cell wall—peptidoglycan and arabinogalactan. *Cold Spring Harb. Perspect. Med.* **5**, 1–16 (2015).
- Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H. & Jackson, M. The cell envelope glycoconjugates of Mycobacterium tuberculosis. *Crit. Rev. Biochem. Mol. Biol.* **49**, 361–399. <https://doi.org/10.3109/10409238.2014.925420> (2014).
- Grzegorzewicz, A. E. et al. Assembling of the Mycobacterium tuberculosis cell wall core. *J. Biol. Chem.* **291**, 18867–18879 (2016).
- Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H. & Jackson, M. The cell envelope glycoconjugates of mycobacterium tuberculosis. *Crit. Rev. Biochem. Mol. Biol.* **49**, 361–399 (2014).
- Dulberger, C. L., Rubin, E. J. & Boutte, C. C. The mycobacterial cell envelope—a moving target. *Nat. Rev. Microbiol.* **18**, 47–59. <https://doi.org/10.1038/s41579-019-0273-7> (2020).
- Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M. & Engelhardt, H. Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc. Natl. Acad. Sci. USA* **105**, 3963–3967 (2008).
- Crick, D. C., Mahapatra, S. & Brennan, P. J. Biosynthesis of the arabinogalactan-peptidoglycan complex of Mycobacterium tuberculosis. *Glycobiology* **11**, 107–118 (2001).
- Fukuda, T. et al. Critical roles for lipomannan and lipoarabinomannan in cell wall integrity of mycobacteria and pathogenesis of tuberculosis. *mBio* **4**, 8–10 (2013).
- Alderwick, L. J., Birch, H. L., Mishra, A. K., Eggeling, L. & Besra, G. S. Structure, function and biosynthesis of the Mycobacterium tuberculosis cell wall: arabinogalactan and lipoarabinomannan assembly with a view to discovering new drug targets. *Biochem. Soc. Trans.* **35**, 1325–1328 (2007).
- Islam, M. M. et al. Drug resistance mechanisms and novel drug targets for tuberculosis therapy. *J. Genet. Genom.* **44**, 21–37 (2017).
- Jankute, M., Grover, S., Rana, A. K. & Besra, G. S. Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiol.* **7**, 129–147. <https://doi.org/10.2217/fmb.11.123> (2012).
- Takayama, K., Wang, L. & David, H. L. Effect of isoniazid on the in vivo mycolic acid synthesis, cell growth, and viability of Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **2**, 29–35 (1972).
- Vilchèze, C. & Jacobs, W. R. The isoniazid paradigm of killing, resistance, and persistence in mycobacterium tuberculosis. *J. Mol. Biol.* **431**, 3450–3461 (2019).
- Banerjee, A. et al. inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. *Science* **263**, 227–230 (1994).
- Goude, R., Amin, A. G., Chatterjee, D. & Parish, T. The arabinosyl-transferase EmbC is inhibited by ethambutol in Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **53**, 4138–4146 (2009).
- Telenti, A. et al. The emb operon, a gene cluster of Mycobacterium tuberculosis involved in resistance to ethambutol. *Nat. Med.* **3**, 567–570 (1997).
- Belanger, A. E. et al. 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. USA* **93**, 11919–11924 (1996).
24. Zhang, L. et al. Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. *Science* **368**, 1211–1219 (2020).
 25. Seung, K. J., Keshavjee, S. & Rich, M. L. Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harb. Perspect. Med.* **5**, 1–20 (2015).
 26. Dean, A. S. et al. 25 years of surveillance of drug-resistant tuberculosis: achievements, challenges, and way forward. *Lancet Infect. Dis.* **22**, e191–e196 (2022).
 27. Lee, R. E., Mikušová, K., Brennan, P. J. & Besra, G. S. Synthesis of the mycobacterial arabinose donor β -D-arabinofuranosyl-1-monophosphoryldecaprenol, development of a basic arabinosyltransferase assay, and identification of ethambutol as an arabinosyl transferase inhibitor. *J. Am. Chem. Soc.* **117**, 11829–11832 (1995).
 28. Wolucka, B. A., McNeil, M. R., De Hoffmann, E., Chojnacki, T. & Brennan, P. J. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J. Biol. Chem.* **269**, 23328–23335 (1994).
 29. Lee, R. E., Brennan, P. J. & Besra, G. S. Mycobacterial arabinan biosynthesis: the use of synthetic arabinoside acceptors in the development of an arabinosyl transfer assay. *Glycobiology* **7**, 1121–1128 (1997).
 30. Gong, Yicheng. Structure of the priming arabinosyltransferase AftA required for AG biosynthesis of Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. USA* **120**, 2017 (2023).
 31. Berg, S. et al. Roles of conserved proline and glycosyltransferase motifs of EmbC in biosynthesis of lipoarabinomannan. *J. Biol. Chem.* **280**, 5651–5663 (2005).
 32. Birch, H. L. et al. Biosynthesis of mycobacterial arabinogalactan: Identification of a novel α (1 \rightarrow 3) arabinofuranosyltransferase. *Mol. Microbiol.* **69**, 1191–1206 (2008).
 33. Škovierová, H. et al. AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology* **19**, 1235–1247 (2009).
 34. Alderwick, L. J. et al. AftD functions as an α 1 \rightarrow 5 arabinofuranosyltransferase involved in the biosynthesis of the mycobacterial cell wall core. *Cell Surf.* **1**, 2–14 (2018).
 35. Jankute, M. et al. Disruption of mycobacterial aftB results in complete loss of terminal β (1 \rightarrow 2) arabinofuranose residues of Lipoarabinomannan. *ACS Chem. Biol.* **12**, 183–190 (2017).
 36. Seidel, M. et al. Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in Corynebacteriaceae, such as Corynebacterium glutamicum and Mycobacterium tuberculosis. *J. Biol. Chem.* **282**, 14729–14740 (2007).
 37. McNeil, M., Daffe, M. & Brennan, P. J. Location of the mycolyl ester substituents in the cell walls of mycobacteria. *J. Biol. Chem.* **266**, 13217–13223 (1991).
 38. Harrison, J. et al. Lcp1 is a phosphotransferase responsible for ligating arabinogalactan to peptidoglycan in mycobacterium tuberculosis. *mBio* **7**, 10–1128 (2016).
 39. Koh, E. et al. Chemical–genetic interaction mapping links carbon metabolism and cell wall structure to tuberculosis drug efficacy. *Proc. Natl. Acad. Sci. USA* **119**, e2201632119 (2022).
 40. Angala, S. K. et al. Use of synthetic glycolipids to probe the number and position of arabinan chains on mycobacterial arabinogalactan. *ACS Chem. Biol.* **16**, 20–26 (2021).
 41. Zhang, Jian. et al. Reconstitution of functional mycobacterial arabinosyltransferase AftC proteoliposome and assessment of decaprenylphosphorylarabinose analogues as arabinofuranosyl donors. *ACS Chem. Biol.* **6**, 819–828 (2011).
 42. Lee, A. et al. Sequencing of oligoarabinosyl units released from mycobacterial arabinogalactan by endogenous arabinanase: Identification of distinctive and novel structural motifs. *Biochemistry* **45**, 15817–15828 (2006).
 43. Zhang, J., Khoo, K. H., Wu, S. W. & Chatterjee, D. Characterization of a distinct arabinofuranosyltransferase in Mycobacterium smegmatis. *J. Am. Chem. Soc.* **129**, 9650–9662 (2007).
 44. Dominik, P. K. & Kossiakoff, A. A. *Phage Display Selections for Affinity Reagents to Membrane Proteins in Nanodiscs. Methods in Enzymology* Vol. 557 (Elsevier Inc., 2015).
 45. Rizk, S. S. et al. Allosteric control of ligand-binding affinity using engineered conformation-specific effector proteins. *Nat. Struct. Mol. Biol.* **18**, 437–444 (2011).
 46. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**, 133–159 (2015).
 47. Alexander, J. A. N. & Locher, K. P. Emerging structural insights into C-type glycosyltransferases. *Curr. Opin. Struct. Biol.* **79**, 102547 (2023).
 48. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
 49. Zinkle, A. P., Morgan, R. T., Nygaard, R. & Mancina, F. Structural insights into polyisoprenyl-binding glycosyltransferases. *Structure* 1–13. <https://doi.org/10.1016/j.str.2025.01.003> (2025).
 50. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44**, W351–W355 (2016).
 51. Moremen, K. W. & Haltiwanger, R. S. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. *Nat. Chem. Biol.* **15**, 853–864 (2019).
 52. Napiórkowska, M., Boilevin, J., Darbre, T., Reymond, J. L. & Locher, K. P. Structure of bacterial oligosaccharyltransferase PglB bound to a reactive LLO and an inhibitory peptide. *Sci. Rep.* **8**, 1–9 (2018).
 53. Napiórkowska, M. et al. Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase. *Nat. Struct. Mol. Biol.* **24**, 1100–1106 (2017).
 54. Petrou, V. I. et al. Structural biology: structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation. *Science* **351**, 608–612 (2016).
 55. Bloch, J. S. et al. Structure and mechanism of the ER-based glucosyltransferase ALG6. *Nature* **579**, 443–447 (2020).
 56. Ramírez, A. S., Kowal, J. & Locher, K. P. Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science* **366**, 1372–1375 (2019).
 57. Joe, M. & Lowary, T. L. Synthesis of 2-deoxy-2-fluoro analogs of polyprenyl β -D-arabinofuranosyl phosphates. *Carbohydr. Res.* **341**, 2723–2730 (2006).
 58. Discovery, C. Chai-1: decoding the molecular interactions of life. 1–18 (2024).
 59. Krishna, R. et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **384**, ead12528 (2024).
 60. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* **61**, 3891–3898 (2021).
 61. Liang, D. M. et al. Glycosyltransferases: mechanisms and applications in natural product development. *Chem. Soc. Rev.* **44**, 8350–8374. <https://doi.org/10.1039/c5cs00600g> (2015).
 62. Breton, C., Šnajdrová, L., Jeanneau, C., Koča, J. & Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29–37 (2006).
 63. Schuman, B., Evans, S. V. & Fyles, T. M. Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. *PLoS ONE* **8**, e71077 (2013).
 64. Liav, A., Huang, H., Ciepichal, E., Brennan, P. J. & McNeil, M. R. Stereoselective synthesis of decaprenylphosphoryl β -D-arabinofuranose. *Tetrahedron Lett.* **47**, 545–547 (2006).

65. Lee, S. S. et al. Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* **7**, 631–638 (2011).
66. Forrester, T. J. B. et al. The retaining β-Kdo glycosyltransferase WbbB uses a double-displacement mechanism with an intermediate adduct rearrangement step. *Nat. Commun.* **13**, 1–13 (2022).
67. Rao, D., Zhu, L., Liu, W. & Guo, Z. Molecular mechanism of double-displacement retaining β-Kdo glycosyltransferase WbbB. *J. Phys. Chem. B* <https://doi.org/10.1021/acs.jpcc.4c02073> (2024).
68. Doyle, L. et al. Mechanism and linkage specificities of the dual retaining β-Kdo glycosyltransferase modules of KpsC from bacterial capsule biosynthesis. *J. Biol. Chem.* **299**, 1–18 (2023).
69. Hövel, K. et al. Crystal structure and snapshots along the reaction pathway of a family 51 α-L-arabinofuranosidase. *EMBO J.* **22**, 4922–4932 (2003).
70. Gandhi, N. S., Freeman, C., Parish, C. R. & Mancera, R. L. Computational analyses of the catalytic and heparin-binding sites and their interactions with glycosaminoglycans in glycoside hydrolase family 79 endo-d-glucuronidase (heparanase). *Glycobiology* **22**, 35–55 (2012).
71. Pengthaisong, S., Piniello, B., Davies, G. J., Rovira, C. & Ketudat Cairns, J. R. Reaction mechanism of glycoside hydrolase family 116 utilizes perpendicular protonation. *ACS Catal.* **13**, 5850–5863 (2023).
72. Romero-Téllez, S., Lluch, J. M., González-Lafont, À. & Masgrau, L. Comparing hydrolysis and transglycosylation reactions catalyzed by *Thermus thermophilus* β-glycosidase. A combined MD and QM/MM study. *Front. Chem.* **7**, 1–19 (2019).
73. Liu, Y. et al. Mechanistic studies of mycobacterial glycolipid biosynthesis by the mannosyltransferase PimE. *Nat. Commun.* Accepted, Preprint at <https://doi.org/10.1101/2024.09.17.613550> (2025).
74. Tan, Y. Z. et al. Cryo-EM structures and regulation of arabinofuranosyltransferase AftD from mycobacteria. *Mol. Cell* **78**, 683–699.e11 (2020).
75. Zhang, L. et al. Cryo-EM snapshots of mycobacterial arabinosyltransferase complex EmbB2-AcpM2. *Protein Cell* **11**, 505–517 (2020).
76. Takayama, K. & Kilburn, J. O. Inhibition of synthesis of arabinogalactan by ethambutol in *Mycobacterium smegmatis*. *Antimicrob. Agents Chemother.* **33**, 1493–1499 (1989).
77. Ratnatunga, C. N. et al. The rise of non-tuberculosis mycobacterial lung disease. *Front. Immunol.* **11**, 1–12 (2020).
78. Ruan, J., Li, X. Y. & Chen, H. *Mycobacterium chubuense* hand infection. *IDCases* **20**, e00742 (2020).
79. Abdelaal, H. F. M., Chan, E. D., Young, L., Baldwin, S. L. & Coler, R. N. *Mycobacterium abscessus*: it's complex. *Microorganisms* **10**, 1–24 (2022).
80. Bruni, R. & Kloss, B. High-throughput cloning and expression of integral membrane proteins in *Escherichia Coli*. *Curr. Protoc. Protein Sci.* **74**, 29.6.1–29.6.34 (2013).
81. Suloway, C. et al. Automated molecular microscopy: the new Legion system. *J. Struct. Biol.* **151**, 41–60 (2005).
82. Kim, J. et al. Structure and drug resistance of the *Plasmodium falciparum* transporter PfCRT. *Nature* **576**, 315–320 (2019).
83. Miller, K. R. et al. T cell receptor-like recognition of tumor in vivo by synthetic antibody fragment. *PLoS ONE* **7**, 1–14 (2012).
84. Dominik, P. K. et al. Conformational chaperones for structural studies of membrane proteins using antibody phage display with nanodiscs. *Structure* **24**, 300–309 (2016).
85. Paduch, M. & Kossiakoff, A. A. Generating conformation and complex-specific synthetic antibodies. *Methods Mol. Biol.* **1575**, 93–119 (2017).
86. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. Cryo-SPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
87. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
88. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).
89. Slabinski, L. et al. XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403–3405 (2007).
90. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
91. Oluwole, A. O. et al. Peptidoglycan biosynthesis is driven by lipid transfer along enzyme-substrate affinity gradients. *Nat. Commun.* **13**, 1–12 (2022).
92. Souza, P. C. T. et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **18**, 382–388 (2021).
93. Alessandri, R. et al. Martini 3 Coarse-grained force field: small molecules. *Adv. Theory Simul.* **5**, 2100391 (2022).
94. Borges-Araújo, L. et al. Martini 3 coarse-grained force field for carbohydrates. *J. Chem. Theory Comput.* **19**, 7387–7404 (2023).
95. Vanommeslaeghe, K. et al. CHARMM General Force Field (CGenFF): a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
96. Graham, J. A., Essex, J. W. & Khalid, S. PyCGTOOL: automated generation of coarse-grained molecular dynamics models from atomistic trajectories. *J. Chem. Inf. Model.* **57**, 650–656 (2017).
97. Kroon, P. C. et al. Martinize2 and Vermouth: unified framework for topology generation. *eLife* **12**, 1–39 (2023).
98. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).
99. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
100. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
101. Abraham, M. J. et al. Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
102. Lindahl, E., Abraham, M. J., Hess, B. & van der Spoel. GROMACS 2021.4 Source code. Version 2021.4 <https://zenodo.org/records/5636567> (2021).
103. Vickery, O. N. & Stansfeld, P. J. CG2AT2: an enhanced fragment-based approach for serial multi-scale molecular dynamics simulations. *J. Chem. Theory Comput.* **17**, 6472–6482 (2021).
104. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
105. Kim, S. et al. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J. Comput. Chem.* **38**, 1879–1886 (2017).
106. Ashraf, K. U. et al. Structural basis of lipopolysaccharide maturation by the O-antigen ligase. *Nature* **604**, 371–376 (2022).
107. Nygaard, R. et al. Structural basis of peptidoglycan synthesis by *E. coli* RodA-PBP2 complex. *Nat. Commun.* **14**, 5151 (2023).
108. Bernetti, M. & Bussi, G. Pressure control using stochastic cell rescaling. *J. Chem. Phys.* **153**, 1–11 (2020).
109. Song, W. et al. PyLipID: a Python package for analysis of protein-lipid interactions from molecular dynamics simulations. *J. Chem. Theory Comput.* **18**, 1188–1201 (2022).
110. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

111. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
112. Olsson, M. H. M., S ndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
113. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
114. Voss, N. R. & Gerstein, M. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**, 555–562 (2010).

Acknowledgements

This work was supported by R35GM132120 (to F.M.). P.J.S. acknowledges R01AI174416 (PI: M. Stephen Trent), Wellcome (208361/Z/17/Z), MRC, BBSRC (BB/P01948X/1), and the Howard Dalton Centre for funding. P.J.S. and C.M.B. acknowledge Sulis at HPC Midlands+, which was funded by the EPSRC on grant EP/T022108/1, and the University of Warwick Scientific Computing Research Technology Platform for computational access. C.M.B. was supported by an MRC studentship (MR/NO14294/1). T.L.L. thanks Academia Sinica intramural funding and the Canadian Glycomics Network (SD-1) for support. K.H.K. thanks Academia Sinica intramural funding. This project made use of time on ARCHER2 and JADE2 granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (<http://www.hecbiosim.ac.uk>), supported by EPSRC (grant no. EP/R029407/1). M.A. acknowledges support by FCT—Fundação para a Ciência e a Tecnologia, I.P., through PTDC/BIA-BQM/4056/2020, MOSTMICRO-ITQB R&D Unit (DOI 10.54499/UIDB/04612/2020; DOI 10.54499/UIBP/04612/2020) and LS4FUTURE Associated Laboratory (DOI 10.54499/LA/P/0087/2020); and EU H2020 grant No 823780. Mass spectrometry data for products from the functional assay were acquired at the Medicinal Chemistry and Analytical Core Facilities in the Biomedical Translation Research Center located at National Biotechnology Research Park, supported by Academia Sinica Core Facility and Innovative Instrument (AS-NBRPCF-111-201). A.A.K. acknowledges support from NIH grant R01GM117372.

Author contributions

Y.L. and B.K. conducted the initial high-throughput screening and J.R. contributed to these early stages of the project. Y.L. performed molecular cloning, protein expression, purification, cryo-EM sample preparation, data processing, model building, and structural refinement. Y.C.S. and Y.J.W. conducted the AftB functional assay. P.S.T. and N.H.D. synthesized FPA for the functional assay. C.R.H. and H.Y.C. synthesized the Araf₄ acceptor for the functional assay. Contributions related to the functional assay were made under the guidance of T.L.L. The MS/MS experiments used to support the structure of the products from the

functional assays were carried out by S.Y.G. under the supervision of K.H.K. C.M.B. and P.J.S. carried out all molecular dynamics simulations and docking studies. S.E. and P.T. identified, characterized, and purified the Fabs under the supervision of A.A.K. Y.L. wrote the manuscript with critical input from R.N., F.M., P.J.S., C.M.B., M.A., T.L.L., and S.E. T.L.L., F.M., R.N., and P.J.S. supervised the project. All authors reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-58196-5>.

Correspondence and requests for materials should be addressed to Todd L. Lowary, Phillip J. Stansfeld, Rie Nygaard or Filippo Mancia.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

  The Author(s) 2025

¹Department of Physiology and Cellular Biophysics, Columbia University Irving Medical Center, New York, NY, USA. ²School of Life Sciences & Department of Chemistry, University of Warwick, Coventry, UK. ³Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands. ⁴Department of Biochemistry and Molecular Biophysics, University of Chicago, Chicago, IL, USA. ⁵Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan, ROC. ⁶Department of Chemistry, University of Alberta, Edmonton, AB, Canada. ⁷Chemical Biology and Molecular Biophysics, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan, ROC. ⁸Department of Chemistry, National Tsing Hua University, Hsinchu, Taiwan, ROC. ⁹Instituto de Tecnologia Qu mica e Biol gica Ant nio Xavier, Universidade Nova de Lisboa (ITQB-UNL), Oeiras, Portugal. ¹⁰Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan, ROC. ¹¹Department of Radiation Oncology, Weill Cornell Medicine, New York, NY, USA.   e-mail: tlowary@as.edu.tw; Phillip.Stansfeld@warwick.ac.uk; rin7007@med.cornell.edu; fm123@cumc.columbia.edu