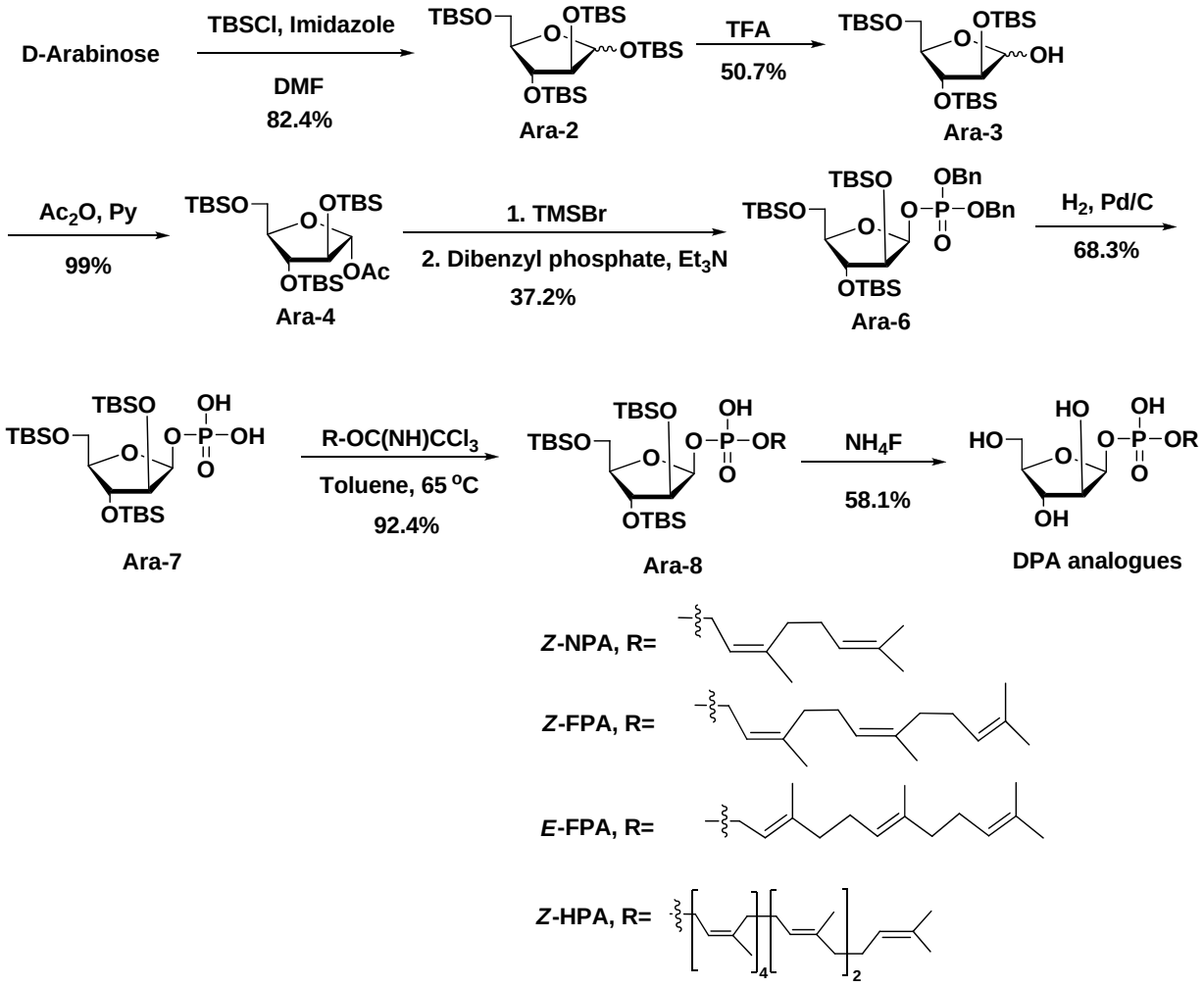
Supporting information for Revised Manuscript ID cb-2011-00091m

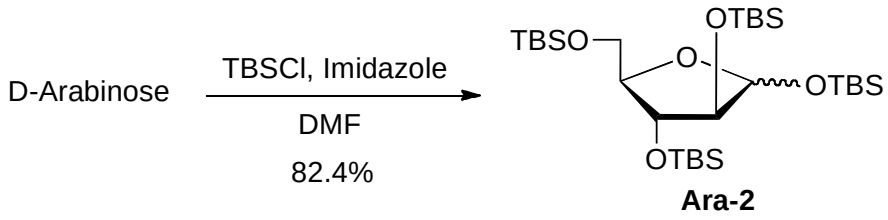
Reconstitution of Functional Mycobacterial Arabinosyltransferase AftC Proteoliposome and Assessment of Decaprenylphosphorylarabinose Analogues as Arabinofuranosyl Donors

Jian Zhang[1, †], Shiva K. Angala[2], Pradeep K. Pramanik[2], Kai Li[2], Dean C. Crick[2], Abraham Liav[2], Adam Jozwiak[3], Ewa Swiezewska[3] Mary Jackson[2] and Delphi Chatterjee[2,*]

1. Synthesis of *Z*-NPA, *E*-FPA and *Z*-HPA



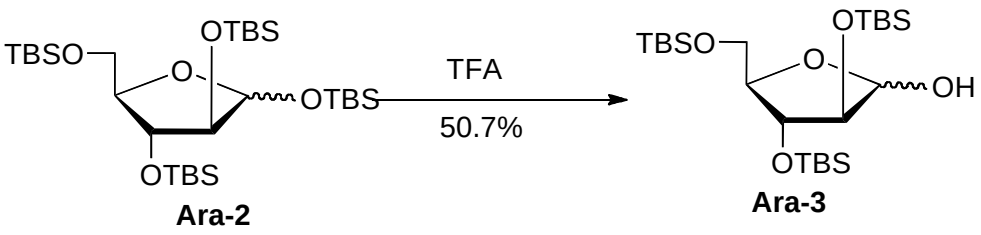
Scheme S1. Synthesis of DPA analogues



Ara-2

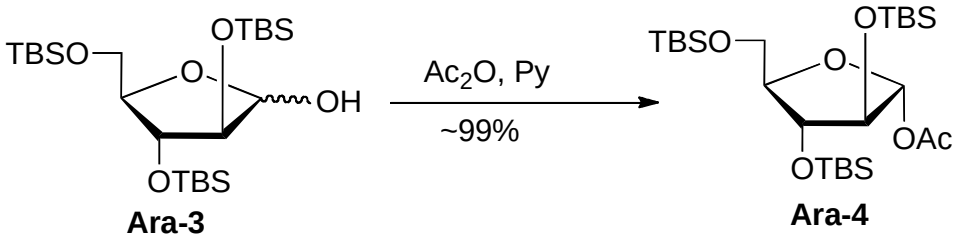
1,2,3,5-(Tetra *tert*-butyldimethylsilyl)-D-arabinose (Ara-2). D-Arabinose (2 g, 13.3 mmol) and imidazole (4.5 g, 66 mmol) were dissolved in DMF (100 mL) and to the stirred solution TBSCl (12 g, 80 mmol) was slowly added. The mixture was stirred at room temperature over night. Ice water was added to the reaction until no more precipitate was produced. The reaction was filtered and the pellet was washed with water, dissolved in acetone and concentrated by

rotatory evaporator. The residue was purified by silica-gel column chromatography at EtOAc/Hexane, 1:40 to give **Ara-2** (6.631 g, 84.2%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.13 (s, 1 H), 4.01 (dd, J = 9.6, 4.8 Hz), 3.93 (m, 2 H), 3.68 (m, 2 H), 0.90 (m, 36 H), 0.10 (m, 24 H).



Ara-2 **Ara-3**

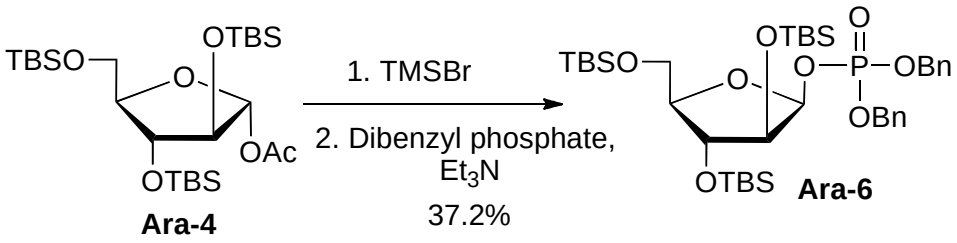
2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinose (Ara-3). To a solution of Ara-2 (6.631 g) in dichloromethane (200 ml), trifluoroacetic acid (43 ml) was added. The mixture was stirred for two minutes before being poured into a stirred solution of NH_4OH (16 ml) in MeOH at -20^{o}C. The mixture was then allowed to warm to room temperature before partitioning between CH_2Cl_2 (50 mL) and water (50 mL). The organic phase was taken, washed with NaCl, dried over Na_2SO_4 and concentrated by rotatory evaporator followed by column chromatography (EtOAc/Hexane, 1:15) to give Ara-3 (2.8g, 50.7%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.56 (dd, J = 6.3, 3.5 Hz, 1 H), 5.34 (dt, J = 6.6, 1.1 Hz, 1 H), 5.13-5.06 (m, 2 H), 4.21 (m, 2 H), 4.09 (m, 1 H), 4.00 (m, 1 H), 3.94-3.56 (m, 3 H), 0.90 (m, 27 H), 0.11-0.08 (m, 18 H). ^{13}C NMR (400 MHz, $CDCl_3$): δ 104.46, 96.77, 88.36, 85.19, 81.19, 78.10, 77.85, 63.85, 26.16, 25.92, 18.58, 18.06, -4.677, -5.195.



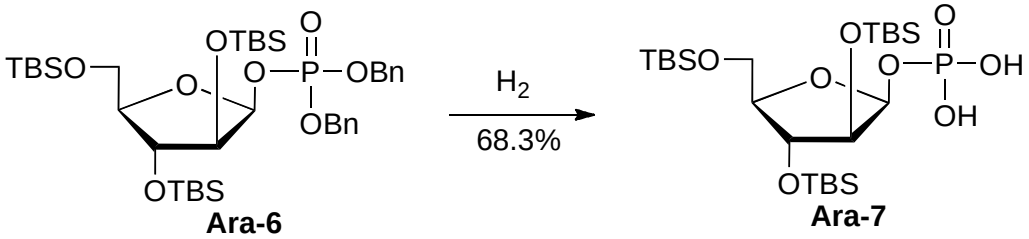
Ara-3 **Ara-4**

Acetyl 2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinose (Ara-4). To a solution of Ara-3 (2.8 g, 1 mmol) in dry pyridine (5 mL) at 0 ^{o}C, Ac_2O (0.78 mL, 1.5 mmol) was added drop-wise with a syringe. The solution was stirred for 4 h at room temperature before being diluted with CH_2Cl_2 and washed with 5% HCl and water. The CH_2Cl_2 layer was dried over Na_2SO_4, and concentrated to give rough product. The crude product was then purified by chromatography on silica gel (hexane/EtOAc, 60:1 $\rightarrow$ 40:1) to give Ara-4 (2.7 g, 99%). TLC performed on hexane/EtOAc, 40:1. ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.92 (s, 1 H), 4.16-4.05 (m, 8 H), 3.80-3.61 (m, 2 H), 2.05

(s, 3 H), 0.90 (m, 27 H), 0.11-0.08 (m, 18 H). ^{13}C NMR (400 MHz, $CDCl_3$): δ 170.45, 103.40, 89.21, 82.56, 78.30, 63.44, 21.37, 18.55, 18.02, 17.94, -3.97, -4.62, -4.82, -5.17, -5.20.

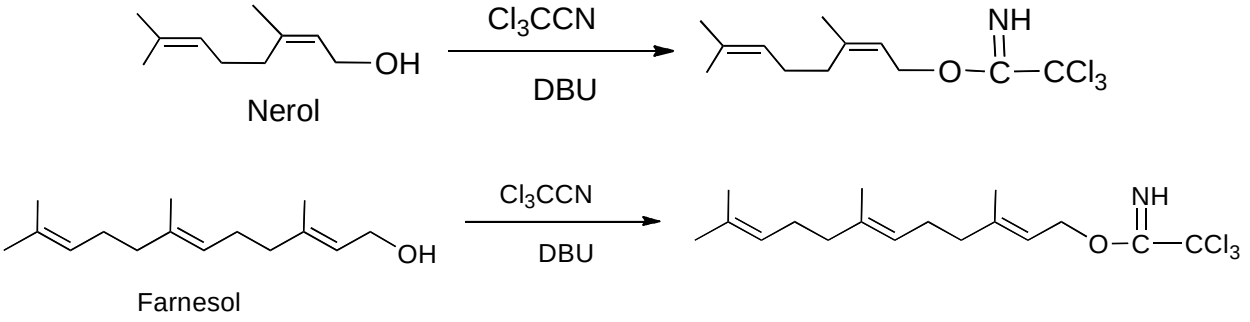


Dibenzyl 2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinosyl phosphate (Ara-6). To a solution of Ara-4 (450 mg, 0.8 mmol) in CH_2Cl_2 (5 mL) was added TMSBr (0.51 ml, 4.0 mmol). The reaction mixture was stirred at room temperature for 2 h and dried in vacuo for 2 h. The resulting Ara-5 (345 mg, 1.1 mmol) and dibenzyl phosphate (460 mg, 1.65 mmol) were dried under vacuum for an additional hour and then dry toluene (5 mL) was added to the reaction mixture. Dry triethylamine (0.34 mL, 1.82 mmol) was added to the mixture. A precipitate was observed within 1–2 minutes and the reaction mixture was allowed to stir for 16 hours at room temperature under an atmosphere of nitrogen. The solvent was removed *in vacuo.* The product was purified by chromatography on silica gel (hexane/EtOAc, 8:1 $\rightarrow$ 4:1) to yield Ara-6 (0.225 g, 37.2% 2 step yield). TLC was performed with hexane/EtOAc 8:1. ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.72 (dd, 0.75 H, J = 3.2, 5.2 Hz, H-1β), 5.65 (d, 0.25 H, J = 5.2 Hz, H-1α), 5.08-5.01 (m, 4 H), 4.04 (m, 2 H), 3.84-3.56 (m, 3 H), 0.90-0.82 (m, 27 H), 0.111- -0.026 (m, 18 H). HRMS (ES^-) calculated $C_{37}H_{65}O_8PSi_3$ 775.3617 [M-H]. Found 775.3614.



2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinosyl phosphate (Ara-7). Ara-6 (149 mg, 0.06 mmol) was dissolved in a mixture of ethanol (5 mL) and triethylamine (0.1 mL, 3.0 mmol). Then 10% Pd/C (15 mg, 30% w/w) was added. The reaction vessel was evacuated and flushed three times with hydrogen gas before allowing the reaction mixture to stir for 48 hours at room temperature under a balloon of hydrogen. TLC tracked the reaction using $CHCl_3$/MeOH/NH_4OH,

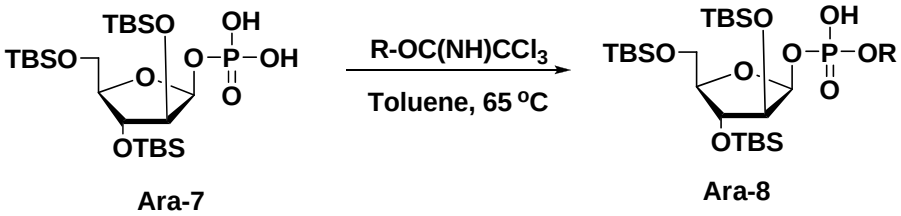
50:10:0.6. The catalyst was filtered and washed with ethanol. The filtrate was concentrated and the residue was dried under vacuum to afford the triethlammonium salt **Ara-7** as an oil (0.102 g, 90%). HRMS (ES^-) calculated $C_{23}H_{52}O_8PSi_3$ 571.2713 [M-H]. Found 571.2716.



To an ice-cold solution of Nerol or Farnesol (88 μL, 0.5 mmol) in DCM (5 mL) was added trichloroacetonitrile (0.1 mL, 1.0 mmol) and DBU (10 μL). The mixture was stirred at room temperature for 90 mins and dried. The mixture was purified by column chromatography (EtOAc/Hexane, 1:8) to give the trichloroacetimidate products.

***Z*-Neroyl trichloroacetimidate**. ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.45 (t, 1 H, J = 6.4 Hz), 5.08 (m, 1 H), 4.79 (d, 2 H, J = 9.6 Hz), 2.14 (m, 4 H), 1.80 (s, 3 H), 1.69 (s, 1 H), 1.61 (s, 3 H). ^{13}C NMR (400 MHz, $CDCl_3$): δ 163.09, 143.80, 132.43, 123.82, 118.71, 66.30, 32.54, 26.92, 25.92, 23.81, 17.92.

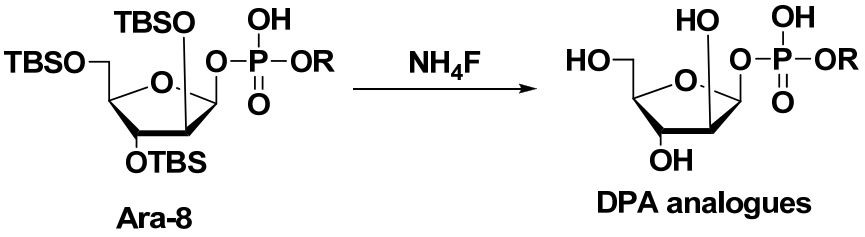
***E*-Farnesyl trichloroacetimidate.** ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.45 (t, 1 H, J = 6.4 Hz), 5.08 (m, 2 H), 4.79 (d, 2 H, J = 7.2 Hz), 2.11-1.93 (m, 8 H), 1.71 (s, 3 H), 1.66 (s, 1 H), 1.59 (s, 3 H), 1.53 (s, 3 H).



***Z*-Neroyl-2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinosyl phosphate (Ara-8a).** A solution of trichloroacetimidate (35 mg, 0.06 mmol) in toluene (2.5 mL) was added to triethyl amine salt of Ara-7 (30 mg, 0.19 mmol) and mixture was stirred at 65 ^{o}C for 3 h. It was then dried and the residue was purified by chromatography with silica gel ($CHCl_3$:MeOH, 5:1) to give Ara-8a (40

mg, 92.4%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.85 (m, 0.75 H, H-1β), 5.45 (m, 0.25 H, H-1α), 5.38 (t, J = 6.0 Hz, 0.25 H, H-3α), 5.05 (m, 1 H), 4.40 (m, 1 H), 4.00 (m, 2 H), 3.62 (m, 2 H), 2.11-2.00 (m, 6 H), 1.67, 1.65, 1.56 (3 × Me), 0.91-0.85 (m, 18 H), 0.13-0.03 (m, 27 H). HRMS (ES^-) calculated $C_{33}H_{68}O_8PSi_3$ 707.3965 [M-H]. Found 707.3966.

***E*-Farnesyl-2,3,5-(Tri *tert*-butyldimethylsilyl)-D-arabinosyl phosphate (Ara-8b).** A solution of trichloroacetimidate (25 mg, 0.3 mmol) in toluene (2.5 mL) was added to triethyl amine salt of Ara-7 (30 mg, 0.19 mmol) and mixture was stirred at 65 ^{o}C for 3 h. It was then dried and the residue was purified by chromatography with silica gel ($CHCl_3$:MeOH, 5:1) to give Ara-8b (65 mg). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.73 (dd, J = 10.8, 20.0 Hz, 0.75 H, H-1β), 5.46 (d, J = 6.0 Hz, 0.25 H, H-1α), 5.39 (t, J = 6.0 Hz, 0.25 H, H-3α), 5.32 (t, J = 6.4 Hz, 0.75 H-3β), 5.16-5.06 (m, 5 H), 4.14-3.98 (m, 2 H), 3.91 (d, 1 H, J =6.8 Hz), 3.81-3.59 (m, 1 H), 3.30 (s, 1 H), 3.13 (s, 1 H), 2.30-2.00 (m, 10 H), 1.65 (s, 6 H), 1.57 (s 6 H), 0.90-0.86 (m, 18 H), 0.11-0.03 (m, 27 H).



Ara-8 DPA analogues

***Z*-Nerylphosphoryl D-arabinose (*Z*-NPA).** Ara-8a (30 mg, 0.042 mmol) was deprotected by treatment with ammouium fluoride (100 mg) and 15% methanolic ammonium hydroxide (2 mL) in MeOH (5 mL) at 65 ^{o}C for 22 h. The mixture was cooled and diluted with DCM. The precipitate was filtered off and washed with DCM/MeOH, 5:1. The filtrate was dried and residue was chromatographed on silica gel. The product was eluted with mixture of 65:125:4, DCM/MeOH/NH_4OH to give ***Z*-NPA** (9 mg, yield: 58.1%). HRMS (ES^-) calculated $C_{15}H_{26}O_8P$ 365.1371 [M-H]. Found 365.1369.

***E,E*-Farnesylphosphoryl D-arabinose (*E*-FPA).** *E*-FPA was synthesized following the procedure of preparing *Z*-NPA. HRMS (ES^-) calculated $C_{20}H_{34}O_8P$ 433.1997 [M-H]. Found 433.2000.

***Z,Z*-Farnesylphosphoryl D-arabinose (*Z*-FPA).** Z-FPA was synthesized following the procedure of preparing *Z*-NPA. ^{1} H NMR (300 MHz, CD_3OD): d 5.49 (m, 1H, partially

overlapped with the H-1 signal), 5.48 (d, *J* =4.5 Hz,1H, H-1), 5.44 (t, *J* =7.2 Hz), 4.44 (t, *J* =6.6 Hz, 2H), 4.08 (t, *J* =7.5 Hz, 1H), 3.98 (m, 1H), 3.75 (dd, *J* =3.0, 11.8 Hz), 3.63 (dd, *J* =5.7, 12.0 Hz, 1H), 2.20-1.92 (m, 8H), 1.75 (s, 3H), 1.67 (s, 3H), 1.61 (s, 3H), 1.60 (s, 3H). HRMS (ES^-) calculated $C_{20}H_{34}O_8P$ 433.1997 [M-H]. Found 433.2050.

***Z,Z,Z,Z,E,E*-heptaprenylphosphoryl D-arabinose (*Z*-HPA).** *Z*-HPA was synthesized according to an established procedure[1]. ^1H NMR (500 MHz, CD3OD): ḍ5.48 (d, *J* =4.8 Hz, 1H, H-1), 5.42 (t, *J* =6.8 Hz, 1H), 5.16–5.09 (m, 4 H), 4.41 (t, *J* =6.6 Hz, 2H), 4.30 (dd, *J* =4.4, 7.3 Hz, 1H), 3.99 (d, *J* =4.4 Hz, 1H), 3.98–3.95 (m, 1H), 3.82 (dd, *J* =2.9, 12.2 Hz, 1H), 3.63 (dd, *J* =5.37, 12.2 Hz 1H), 2.14–1.97 (m, 20H), 1.76 (s, 3H), 1.70 (s, 9H), 1.69 (s, 3H), 1.63 (s, 3H), 1.61 (s, 6H), 1.37–1.31 (m, 6H). HRMS (ES^-) calculated $C_{40}H_{66}O_8P$ 705.4495 [M-H]. Found 705.4480.

---

[1] Liav, A.; Ciepichal, E.; Swiezewska, E.; Bobovská, A.; Dianišková, P.; Blaško, J., Stereoselective syntheses of heptaprenylphosphoryl β-D-arabino-and β-D-ribo-furanoses. *Tetrahedron Letters* **2009,** 50, (19), 2242-2244.

2. Protein Identification for AftC (Rv2673)

(1) SDS-PAGE of AftC stained with Coomasie Blue

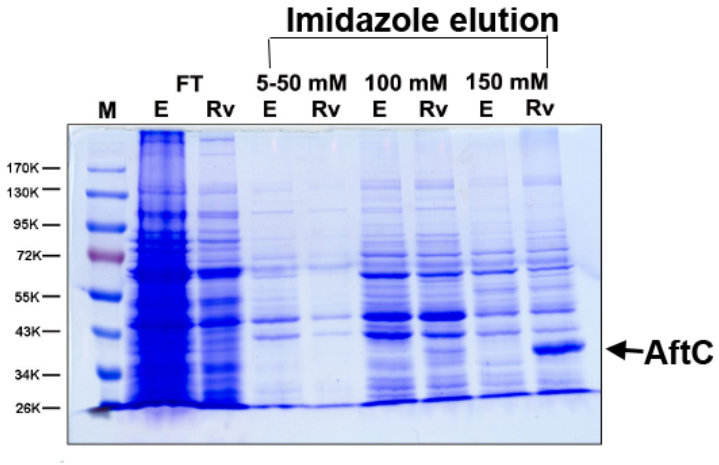


Figure S1. SDS-PAGE analysis of AftC (Rv2673). The cell lysate was applied to affinity chromatography and the column was washed with 20 mL of Tris-HCl buffer with 0.1% Igepal CA-630 (pH 8.0) and eluted with Tris-HCl buffer containing 5 mM to 300 mM of imidazole. The partially purified AftC was eluted in the buffer containing 150mM of imidazole. "E" represents the *M. smegmatis* control strain carried the empty plasmid, pJAM2 and "Rv" represents the pJAMRv2673 strain with expression of AftC.

(2) Tryptic digestion.

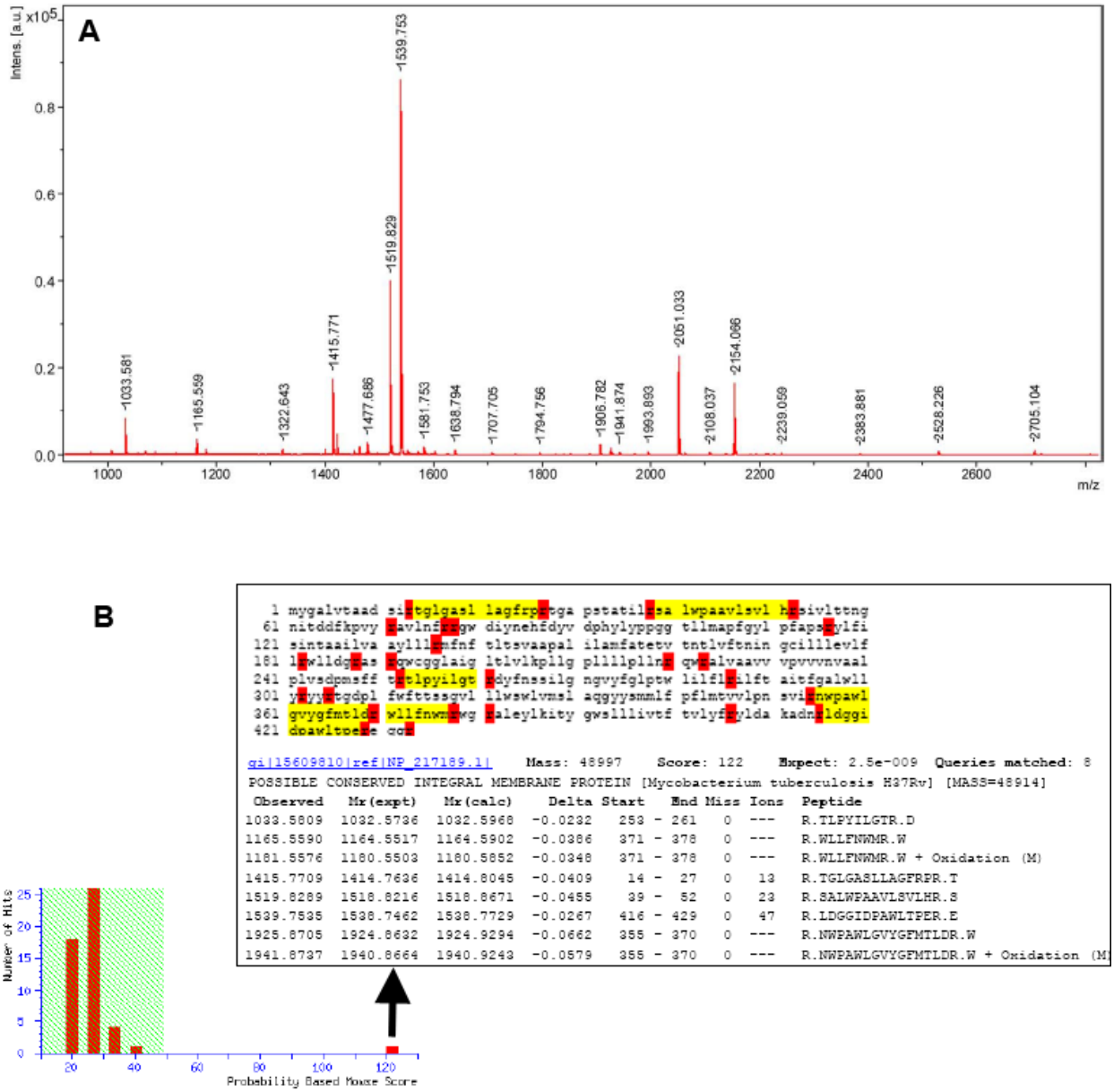


Figure S2. MALDI analysis of purified AftC (Rv2673). To prepare MALDI sample, AftC was subjected to SDS-PAGE (15%), stained with Coomassie Brilliant Blue, and AftC band was cut from the gel and subjected to in-gel tryptic digestion. (A) MALDI spectrum of peptides (assigned in the frame in panel B) after tryptic digestion. (B) Results on the MASCOT searching engineer. The score of 122 in MASCOT (left) represented AftC whose peptide fragments and coverage map are shown in the frame.

3. AraT assays using AftC proteoliposome and *Z*-FPA

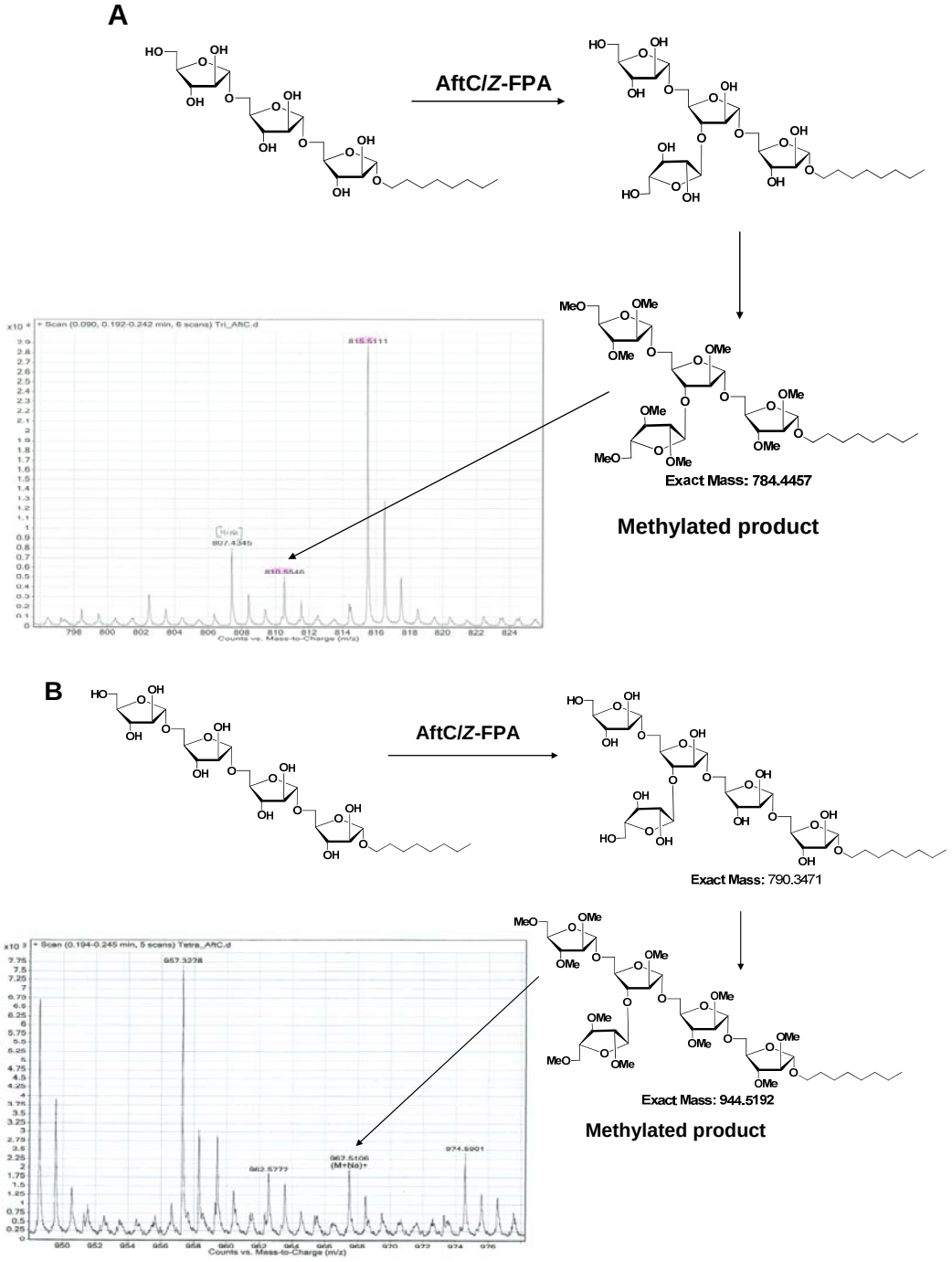


Figure S3. AraT assays with AftC proteoliposome, a linear trisaccharide (A) and a linear tetrasaccharide (B) as acceptor. The enzymatic products were immediately subjected to methylation and ESI-MS analysis after completing the assays. The linear disaccharide did not yield any enzymatic product with AftC proteoliposome.

4. Thin layer chromatography of lipid extracts from *M.tuberculosis*, H37Rv

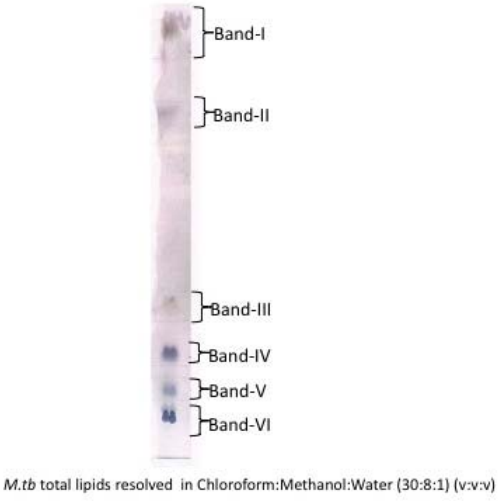


Figure S4A. Thin Layer Chromatography plate of lipid extract from *M. tuberculosis* H37Rv cells. Bands were visualized after dipping in α-naphthol. The total lipid extracts where subjected to preparative TLC, bands were excised as labeled on the TLC plate and extracted from the silica gel with $CHCl_3$:CH_3OH (2:1, V:V). Organic solvents were removed and aliquot of lipid preparations were analysed by mass spectrometry. Lipids were identified using *M. tuberculosis* lipid database {Sartain, 2011 #46}. All of the lipid fractions were used to fabricate AftC proteoliposome.

M.Tb Total lipids	Amount used in making AftC proteoliposomes	Concentration of AftC used in making proteoliposomes (ug)
Band-I	1mg	20ug
Band-II	0.9mg	20ug
Band-III	0.6mg	20ug
Band-IV	1.0mg	20ug
Band-V	1.0mg	20ug
Band-VI	0.2mg	20ug

Note: Reconstitution protocol described in materials and methods

S4B. Chart showing the amounts of lipids obtained after extraction.

Comparable amounts (20 μg) of lipids were used in each case for reconstitution and AftC assay.

Percentage conversion of the enzymatic product

	Acceptor	Product	% conversion
lipid I	1.5*e6	1.8*e3	0.14
lipid II	1.2*e6	1.1*e3	0.09
lipid III	1.33*e6	1.5*e3	0.11
lipid IV	1.5*e6	3.2*e3	0.21
lipid V	1.15*e6	1.2*e3	0.1
lipid VI	1.26*e6	1.2*e3	0.09

S4C. In one preliminary experiment, we could show that AftC activity is somewhat better in comparison to other bands isolated from the preparative TLC (see **S.4A**). We reason that overall poor percentage conversion is due to degradation of proteoliposomes and/or poor incorporation of AftC. This issue is currently being pursued actively with many other lipids.

5. Quantitation of Enzymatic products by MALDI-TOF analysis

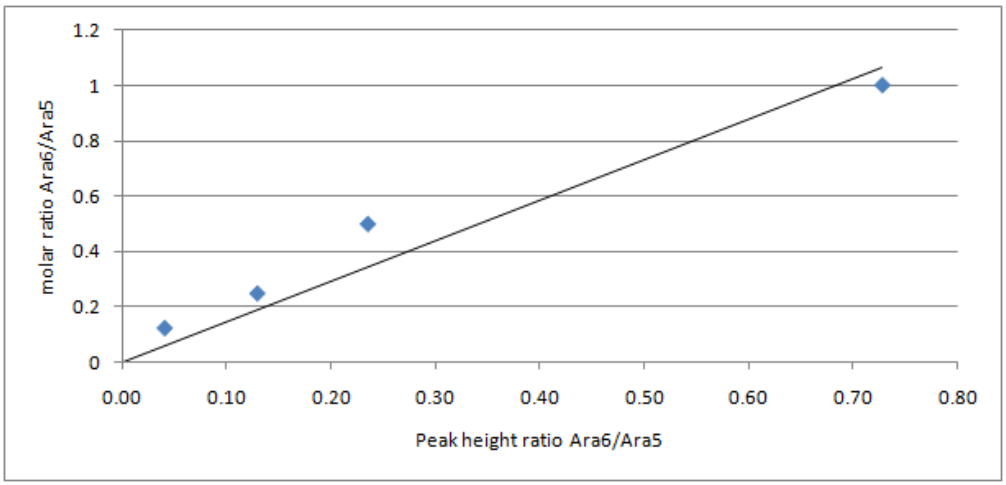


Figure S5. Standard curve obtained after mass spectrometry analysis of Ara_5 and Ara_6. Enzymatic product was quantitated based on the area and height of the peaks of product to substrate. Analysis was performed in sets of triplicates.

Quantitation of Enzymatic Product

	Peak heights					
	Ara6	Ara5	Ratio (height)	Experimental Ratio	Actual Ratio	Enzymatic product % conversion
Z- NPA	790	13386	0.059016883	0.059016883	0.086283	7.9
Z- FPA	1577	9577	0.164665344	0.164665344	0.240741	19.4
E- FPA	0	36595	0	0	0	0.0
Z-HPA	3299	22946	0.143772335	0.143772335	0.210195	17.4

6. Figure S6. ESI-MS Analysis of Band IV. Mass spectra were performed in the negative ion mode.