Disruption of the SucT acyltransferase in *Mycobacterium smegmatis* abrogates succinylation of cell envelope polysaccharides

Zuzana Palčeková[1], Shiva Kumar Angala[1], Juan Manuel Belardinelli[1], Haig A. Eskandarian[2], Maju Joe[3], Richard Brunton[3], Christopher Rithner[4], Victoria Jones[1], Jérôme Nigou[5], Todd L. Lowary[3], Martine Gilleron[5], Michael McNeil[1], Mary Jackson[1*]

From the [1]Mycobacteria Research Laboratories, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523-1682, USA; [2]Global Health Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, VD, CH 1015, Switzerland; [3]Alberta Glycomics Centre and Department of Chemistry, The University of Alberta, Edmonton, AB, T6G 2G2, Canada; [4]Central Instrumentation Facility, Department of Chemistry, Colorado State University, Fort Collins, CO 80523-1872, USA; [5]Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, CNRS, UPS, 205 route de Narbonne, F-31077 Toulouse, France

Running title: *Succinylation of mycobacterial arabinogalactan and LAM*

*To whom correspondence should be addressed: Mary Jackson, Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523-1682, Telephone (970) 491-3582; Fax: (970) 491-1815; E-mail: Mary.Jackson@colostate.edu

Supporting Information

Supplementary Tables

Table S1: Monosaccharidic composition of LAM from WT *M. smegmatis* mc^{2}155, mc^{2}Δ*sucT* and the complemented mutant strain, mc^{2}Δ*sucT*/pMVGH1-*sucT*.

Table S2: Glycosyl linkage analysis of per-*O*-methylated LAM from WT *M. smegmatis* mc^{2}155, mc^{2}Δ*sucT* and mc^{2}Δ*sucT*/pMVGH1-*sucT*.

Table S3: Monosaccharidic composition of mAGP from WT *M. smegmatis* mc^{2}155, mc^{2}Δ*sucT* and the complemented mutant strain, mc^{2}Δ*sucT*/pMVGH1-*sucT*.

Table S4: Glycosyl linkage analysis of per-*O*-methylated mAGP from WT *M. smegmatis* mc^{2}155, mc^{2}Δ*sucT* and mc^{2}Δ*sucT*/pMVGH1-*sucT*.

Supplementary Figures

Figure S1: Amino terminus sequence alignment of *Salmonella typhimurium* OafA with homologous acyltransferases from *M. tuberculosis* using PSI/TM-Coffee.

Figure S2: Construction of acyltransferase knock-outs of *M. smegmatis* and SDS-PAGE analysis of their lipoglycans.

Figure S3: 1D ^{1}H NMR analysis of LM prepared from the WT and mutant strains.

Table S1: Monosaccharidic composition of LAM from WT *M. smegmatis* mc²155, mc²Δ*sucT* and the complemented mutant strain, mc²Δ*sucT*/pMVGH1-*sucT*.

The values reported are for two independent biological repeats.

	WT		**Δ*sucT***		**Δ*sucT* comp**	
	Exp^t I	Exp^t II	Exp^t I	Exp^t II	Exp^t I	Exp^t II
Ara*f*	76.15	79.29	81.04	82.32	70.38	73.06
myo-Ino	0.45	0.65	0.36	0.37	0.68	0.70
Man*p*	23.40	20.06	18.6	17.31	28.94	26.24
Ara*f*/Man*p*	3.25	3.95	4.36	4.76	2.43	2.78

Table S2: Glycosyl linkage analysis of per-*O*-methylated LAM from WT *M. smegmatis* mc²155, mc²Δ*sucT* and mc²Δ*sucT*/pMVGH1-*sucT*.

The values reported are the averages ± SD of three technical repeats

	t-Ara*f*	**2-Ara*f***	**5-Ara*f***	**3,5-Ara*f***	**t-Man*p***	**6-Man*p***	**2,6-Man*p***	**Ara*f*/Man*p***	2,6-α-Man*p*/ [2,6-α-Man*p* + 6-α-Man*p*]	2-α-Ara*f*/ 5-α-Ara*f*
WT	8.1 ± 1.0	8.1 ± 3.8	46.8 ± 1.9	11.2 ± 0.8	9.7 ± 1.3	8.6 ± 0.9	6.9 ± 1.4	2.95 ± 0.16	0.44 ± 0.08	0.18 ± 0.09
Δ*sucT*	11.7 ± 0.8	7.8 ± 3.1	44.8 ± 4.0	12.1 ± 1.1	9.9 ± 1.1	6.0 ± 1.2	7.3 ± 1.1	3.35 ± 0.55	0.55 ± 0.02	0.18 ± 0.08
Δ*sucT* comp	7.1 ± 2.7	5.1 ± 1.3	46.7 ± 3.0	10.5 ± 0.9	12.2 ± 2.9	8.5 ± 4.3	9.2 ± 2.3	2.38 ± 0.55	0.52 ± 0.05	0.11 ± 0.03

Table S3: Monosaccharidic composition of mAGP from WT *M. smegmatis* mc²155, mc²Δ*sucT* and the complemented mutant strain, mc²Δ*sucT*/pMVGH1-*sucT*.

The values reported are averages ± SD of three technical repeats.

	Rha*p*	**Ara*f***	**Gal*f***	**GlcNAc**	**MurNAc**	Ara*f*/Gal*f*	Ara*f*/Rha*p*	Gal*f*/Rha*p*	Mycolic acids/Rha*p*
WT	1.1 ± 0.3	56.8 ± 1.6	21.4 ± 2.9	10.2 ± 0.4	10.5 ± 1.8	2.7 ± 0.4	52.2 ± 11.7	19.3 ± 2.7	22.5
Δ*sucT*	1.0 ± 0.1	57.8 ± 2.3	19.0 ± 2.4	11.6 ± 0.4	10.5 ± 0.8	3.1 ± 0.5	58.9 ± 8.5	19.2 ± 1.7	22.2
Δ*sucT* comp	1.1 ± 0.3	55.0 ± 3.0	24.8 ± 1.8	9.6 ± 1.3	9.4 ± 1.4	2.2 ± 0.3	51.1 ± 11.8	22.7 ± 3.3	29.2

Table S4: Glycosyl linkage analysis of per-*O*-methylated mAGP from WT *M. smegmatis* mc²155, mc²Δ*sucT* and mc²Δ*sucT*/pMVGH1-*sucT*.

The values reported are averages ± SD of three technical repeats.

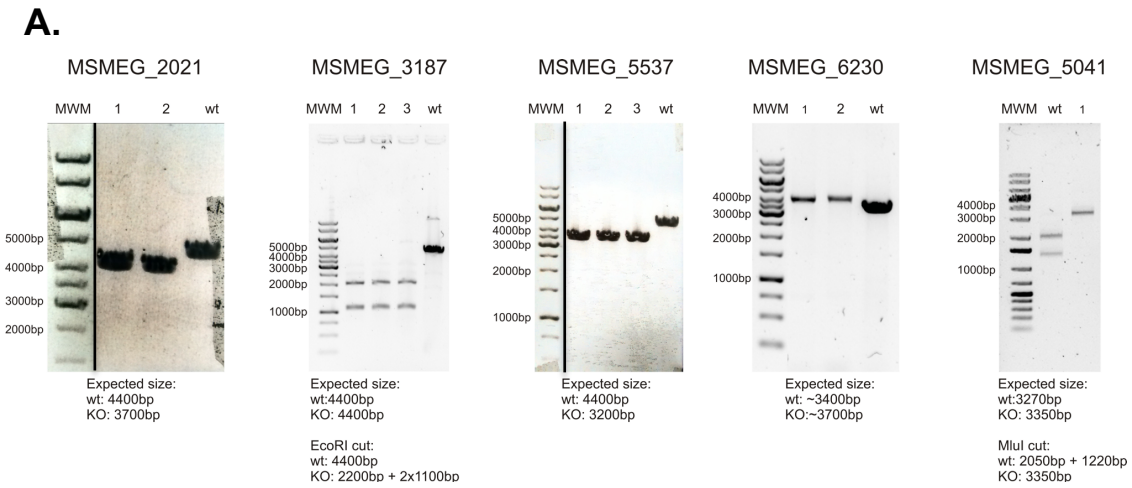
	t-Ara*f*	**2-Ara*f***	**5-Ara*f***	**3,5-Ara*f***	**t-Gal*f***	**5-Gal*f***	**6-Gal*f***	**5,6-Gal*f***	Ara*f*/Gal*f*
WT	9.6 ± 0.3	6.4 ± 0.6	31.3 ± 0.2	9.0 ± 0.1	4.0 ± 0.2	18.7 ± 0.4	11.4 ± 0.4	9.6 ± 1.6	2.95 ± 0.16
Δ*sucT*	9.9 ± 0.3	8.4 ± 0.3	34.0 ± 0.8	9.7 ± 0.4	3.6 ± 0.6	16.4 ± 1.8	10.4 ± 1.2	7.6 ± 5.1	3.35 ± 0.55
Δ*sucT* comp	9.2 ± 0.9	7.7 ± 1.2	27.8 ± 0.9	8.7 ± 0.4	5.0 ± 1.2	21.4 ± 2.9	13.6 ± 2.0	6.8 ± 5.9	2.38 ± 0.55

Figure S1: Amino terminus sequence alignment of *Salmonella typhimurium* OafA with homologous acyltransferases from *M. tuberculosis* (and MSMEG_3187 from *M. smegmatis*) using PSI/TM-Coffee. Conserved sequence (*), conservative mutations (:), semi-conservative mutations (.), and non-conservative mutations ( ). Essential functional residues in OafA are boxed in green (corresponding residues in Rv1565c: S67, G68, R98 and Y335). MSMEG_3187 (SucT; 719 amino acid residues) shares 75% identity (86% similarity) on a 728 amino acid overlap with Rv1565c from *M. tuberculosis* (729 amino acid residues).

<u>Figure S2</u>: Construction of acyltransferase knock-outs of *M. smegmatis* and SDS-PAGE analysis of their lipoglycans.

(A) Allelic replacement at the *MSMEG_2021*, *MSMEG_3187* (*sucT*), *MSMEG_5041*, *MSMEG_5537* and *MSMEG_6230* loci of *M. smegmatis* mc^{2}155. Allelic replacement at all loci was confirmed by PCR using sets of primers located outside the allelic exchange substrates (with or without subsequent restriction digest of the PCR products). The expected sizes of the products for the WT and different mutants are indicated under each gene. Splicing on the gels showing the *MSMEG_2021* and *MSMEG_5537* knock-outs was performed to remove unrelated lanes and add the molecular weight markers.

(B) Lipoglycan profiles of the WT and mutant strains. Lipoglycans extracted from WT mc^{2}155 and each of the knock-out mutants were run on a 10-20% Tricine gel followed by periodic acid-silver staining. The lipoglycan profile of the *MSMEG_3187* (*sucT*) mutant is shown in Figure 2.



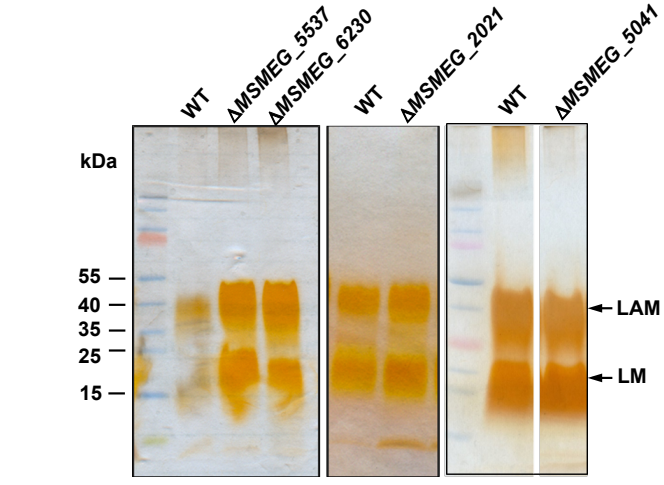
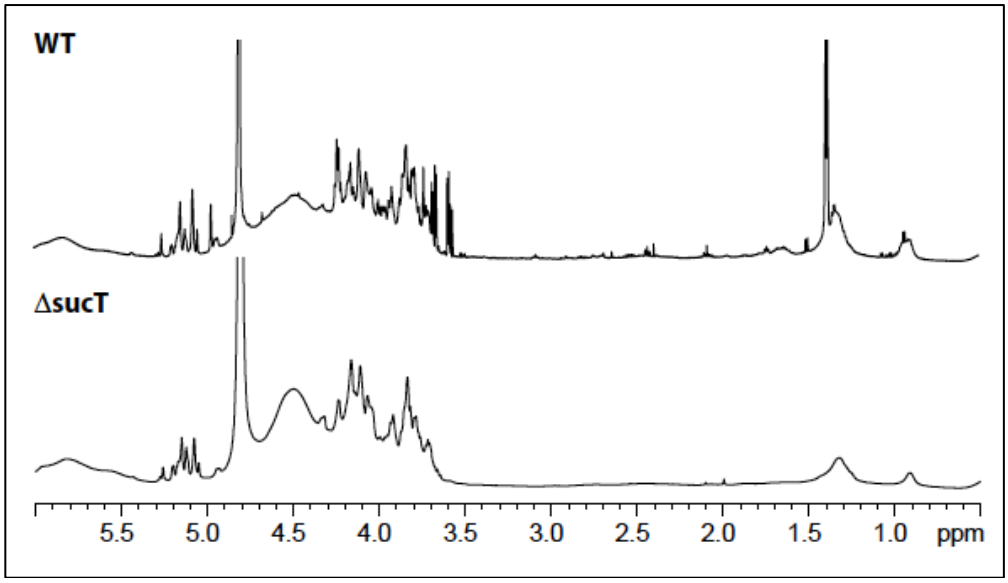
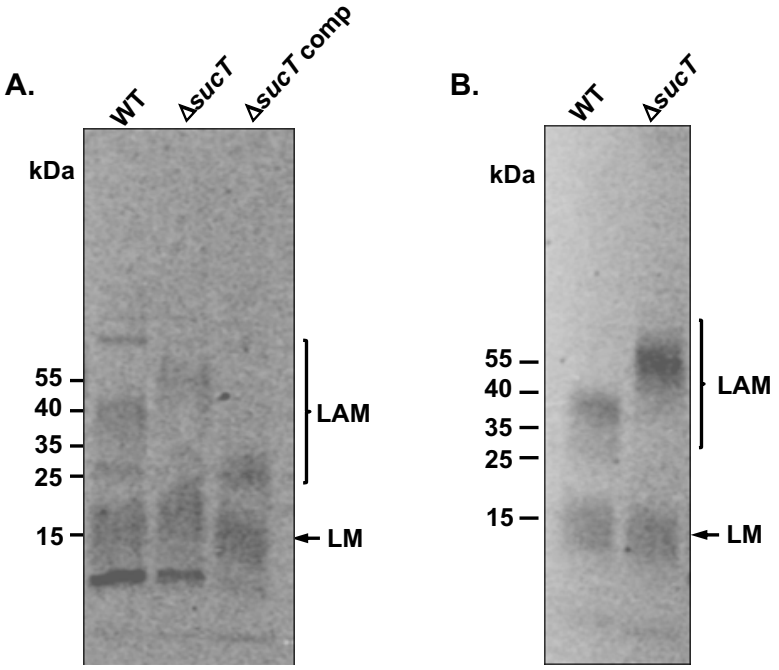


Figure S3: 1D ^{1}H NMR analysis of LM prepared from the WT and mutant strains.



<u>Figure S4</u>: [^{14}C]acetate (A) and [1,2-^{14}C]glucose (B) incorporation in the lipoglycans from WT mc^{2}155, mc^{2}Δ*sucT* and mc^{2}Δ*sucT*/pMVGH1-*sucT*.

M. smegmatis cells were metabolically labeled as described under Methods and their lipoglycans run on 10-20% Tricine gel followed by autoradiography. The strains show no significant difference in lipoglycan acylation or *de novo* synthesis.



<u>Figure S5</u>: Analysis of LAM phospho-inositol capping in the WT, mutant and complemented mutant strains.

Figure S5A: Negative ion LC/MS analysis of oligoarabinosides after *C. gelida* endoarabinanase digestion of LAM from the *M. smegmatis* WT and Δ*sucT* mutant strains.
Panels (a–b) are the extracted ion chromatograms (EICs) showing the presence of Ara_4 oligosaccharide (actual $[M\text{-}H]^-$ at mass *m/z* 545.1723) and Ara_4-PI (actual $[M\text{-}H]^-$ at mass *m/z* 787.1915) in both *M. smegmatis* strains. Panels (c-d) show the exact mass of the $[M\text{-}H]^-$ ions of detected Ara_4-PI, and the panels (e-f) show the exact mass of the $[M\text{-}H]^-$ ions of Ara_4 in both strains.

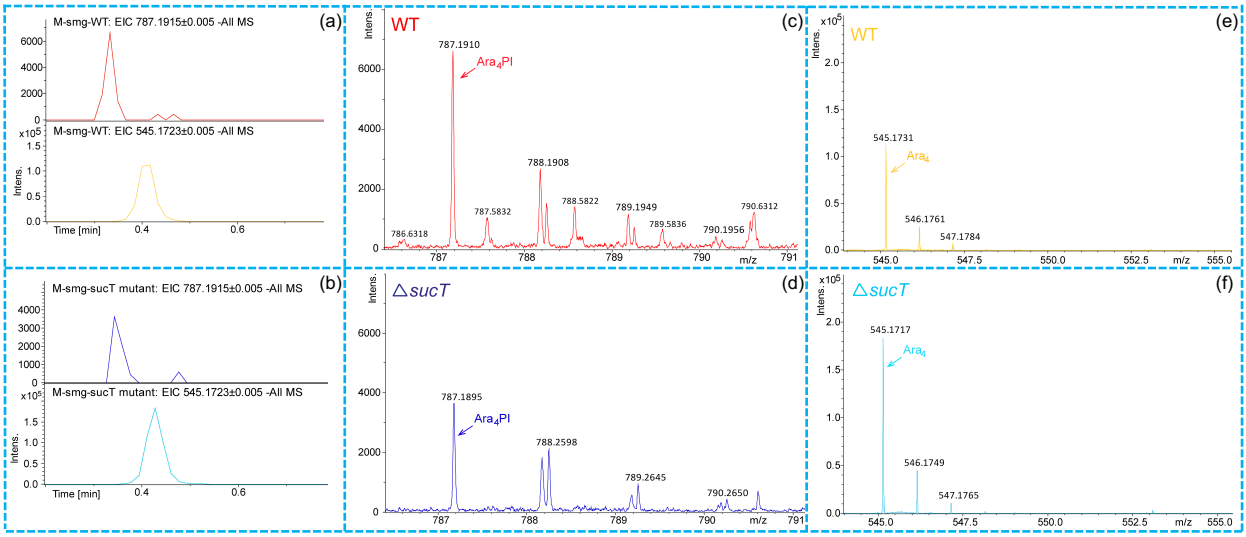
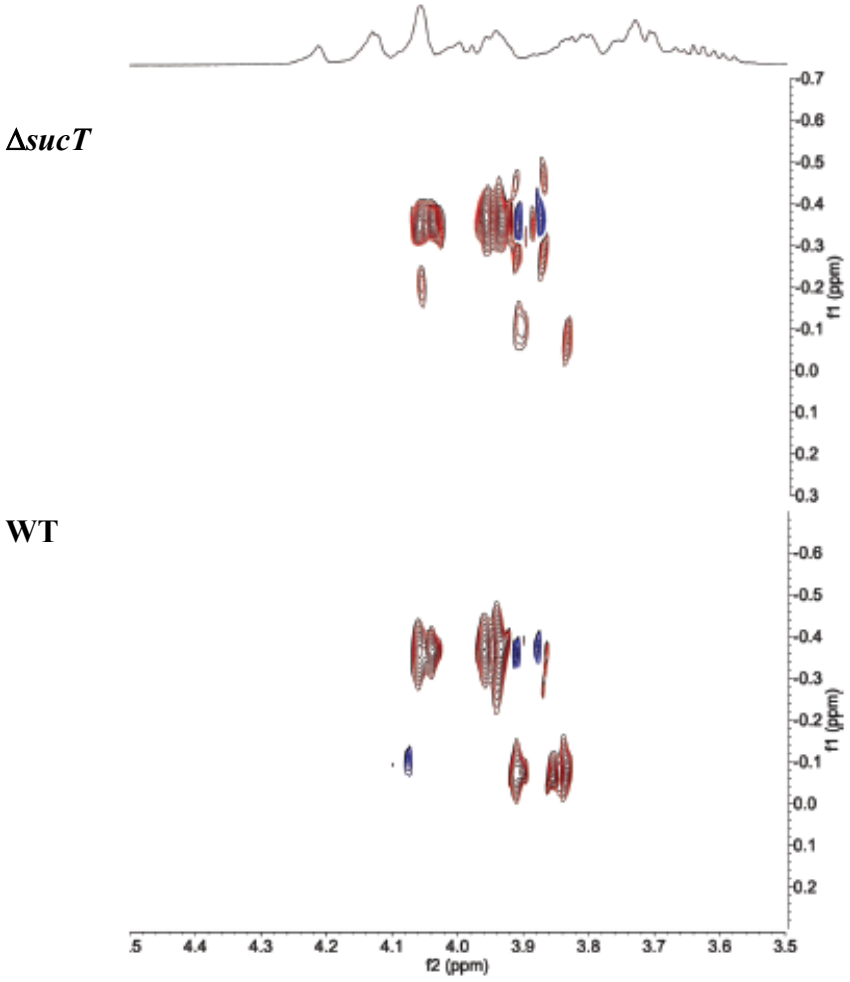
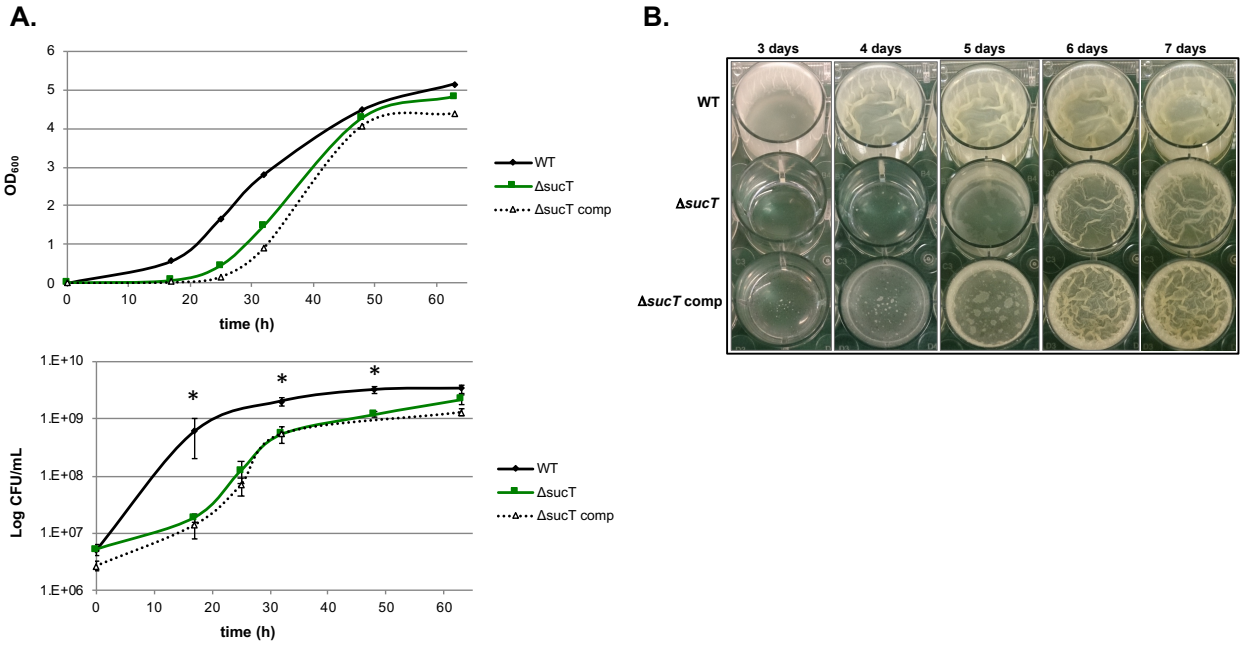


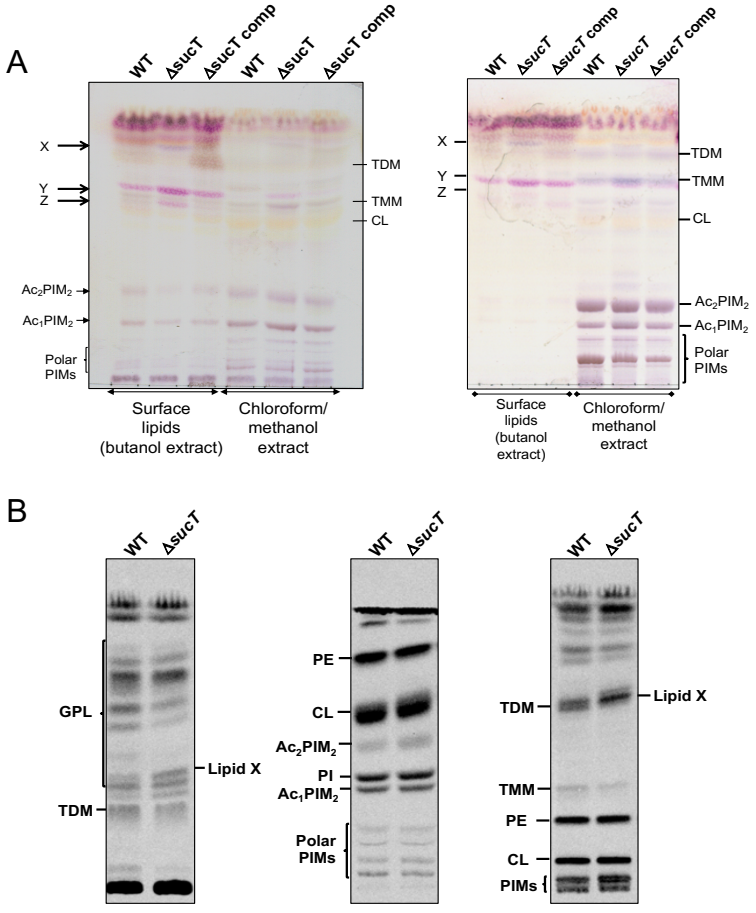
Figure S5B: Two-dimensional ^{1}H-^{31}P HMQC NMR of Δ*sucT* mutant and WT base-treated LAM. Phosphorus resonances at -0.08 and -0.35 (vertical axis) are present. These resonances correspond to the inositol anchor phosphate and the inositol arabinosyl phosphate, respectively, as determined by Gilleron *et al.* (1997). The proton correlations (horizontal axis) of each phosphate are also similar to those reported by Gilleron *et al.* except that the upfield proton correlations for the anchor phosphate were not seen in the mutant spectrum. However, the presence of the phosphorous present as inositol-phosphate-arabinose, including its coupling with protons in LAM as deciphered by Gilleron *et al*. (1997), is very clearly seen in the *M. smegmatis* Δ*sucT* mutant.



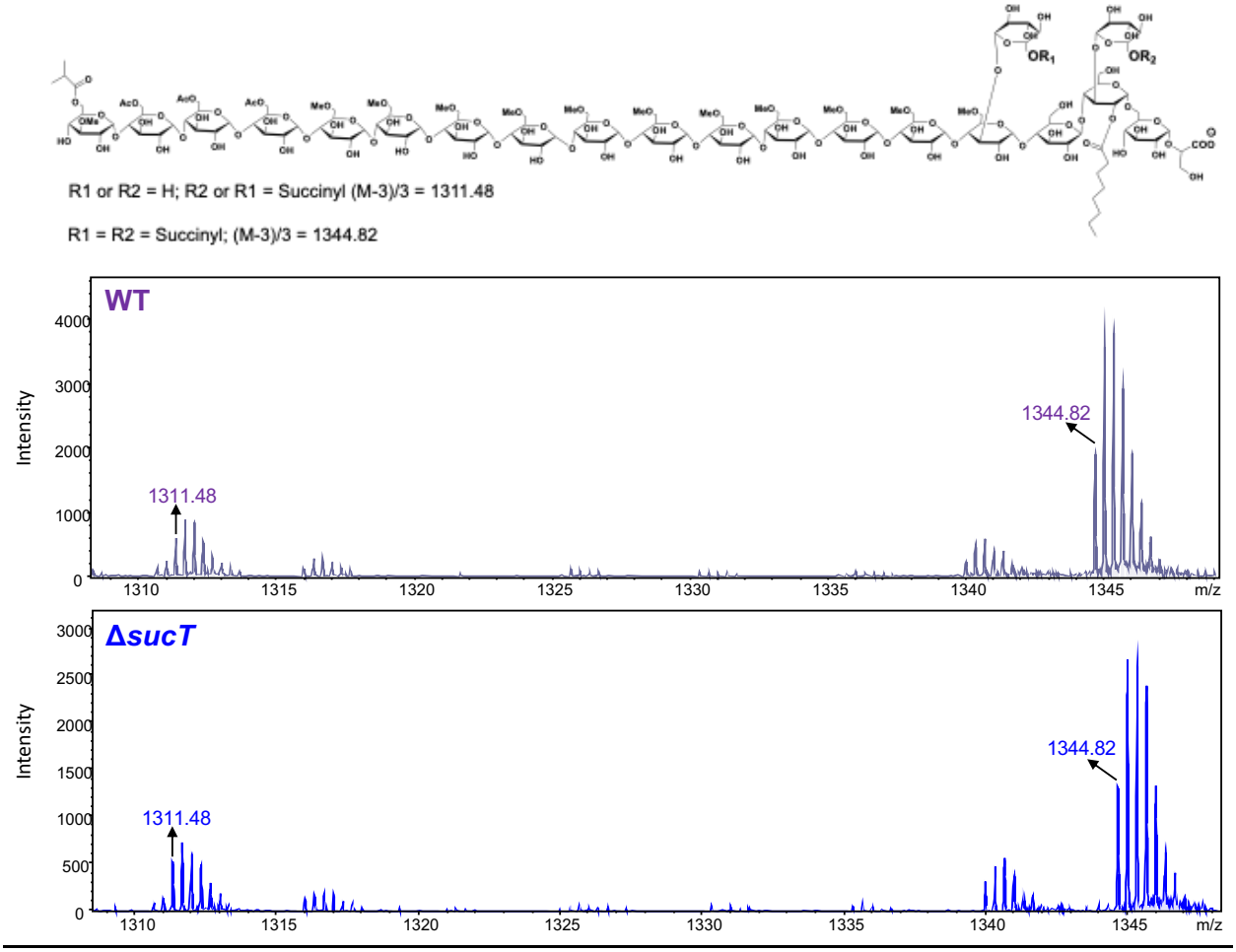
<u>Figure S6</u>: Planktonic and biofilm growth characteristics of the *M. smegmatis sucT* mutant.
(A) Growth kinetics of WT mc^{2}155 (solid black line), mc^{2}Δ*sucT* (green line) and mc^{2}Δ*sucT*/pMVGH1-*sucT* (dotted black line) in 7H9-ADC-tyloxapol at 37°C. Growth kinetics was monitored by absorbance at 600 nm and CFU counting. The CFU values reported are averages ± SD of three technical repeats. Asterisks denote statistically significant differences between WT and *sucT* mutant pursuant to the Student's *t*-test ($P < 0.05$).
(B) Surface pellicle formation in Sauton's medium at 37°C.



<u>Figure S7</u>: Surface and total lipid content of WT mc²155, mc²Δ*sucT* and mc²Δ*sucT*/pMVGH1-*sucT*.
(A) Comparative surface lipid content of the wild-type, mutant and complemented mutant strains grown in 7H9-ADC-Tyloxapol. The TLC analysis of surface (extracted with water-saturated 1-butanol) and other extractable lipids (extracted from 1-butanol-treated cells with chloroform/methanol 2/1, by vol.) from the WT, mutant and complemented mutant strains points to a reproducible increase in the abundance of three compounds (labeled X, Y and Z) in the surface material of the *sucT* mutant. TLC plates were developed once in chloroform/methanol/water/ammonium hydroxide (65/25/4/0.5, by vol.) and revealed by spraying with α-naphthol and heating. The results of two independent experiments performed on different culture batches are shown.
(B) *M. smegmatis* cells were metabolically labeled with [1,2-^{14}C]acetate as described under Methods and their total lipids analyzed by TLC in different solvent systems followed by autoradiography. From left to right, the solvent systems used are: chloroform/methanol (90/10, by vol.), chloroform/methanol/water (65/25/4, by vol.) and chloroform/methanol/water (20/4/0.5, by vol.).
The overall incorporation of radiolabel in the total lipids of each strain was comparable (16.9% in the WT, 20.4% in the *sucT* mutant and 18.7% in the complemented mutant, respectively). All strains show comparable lipid (including phosphatidylinositol mannoside) profiles. The position of lipid X whose abundance is increased in the surface lipids of the mutant strain is shown.
GPL, glycopeptidolipids; TMM, trehalose monomycolates; TDM, trehalose dimycolates; PE, phosphatidylethanolamine; CL, cardiolipin; Ac_1PIM_2, triacylated phosphatidylinositol dimannosides; Ac_2PIM_2, tetraacylated phosphatidylinositol dimannosides.

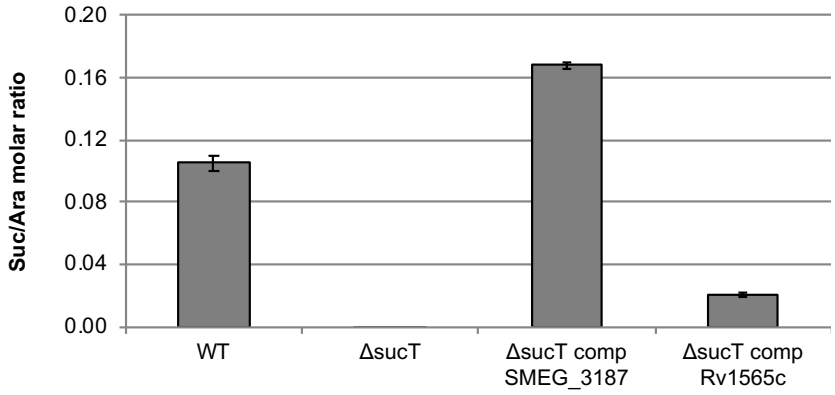


<u>Figure S8</u>: MGLP profiles of the *M. smegmatis* WT and *sucT* mutant.
Negative ion LC/MS of native MGLP. Exact masses of succinylated isoforms were calculated from the MGLP structure published by De *et al.* (2018). Mass spectra are showing the presence of $[M\text{-}3H]^{-3}$ ions corresponding to mono- and di-succinylated MGLP isoforms in both WT and mutant strains.

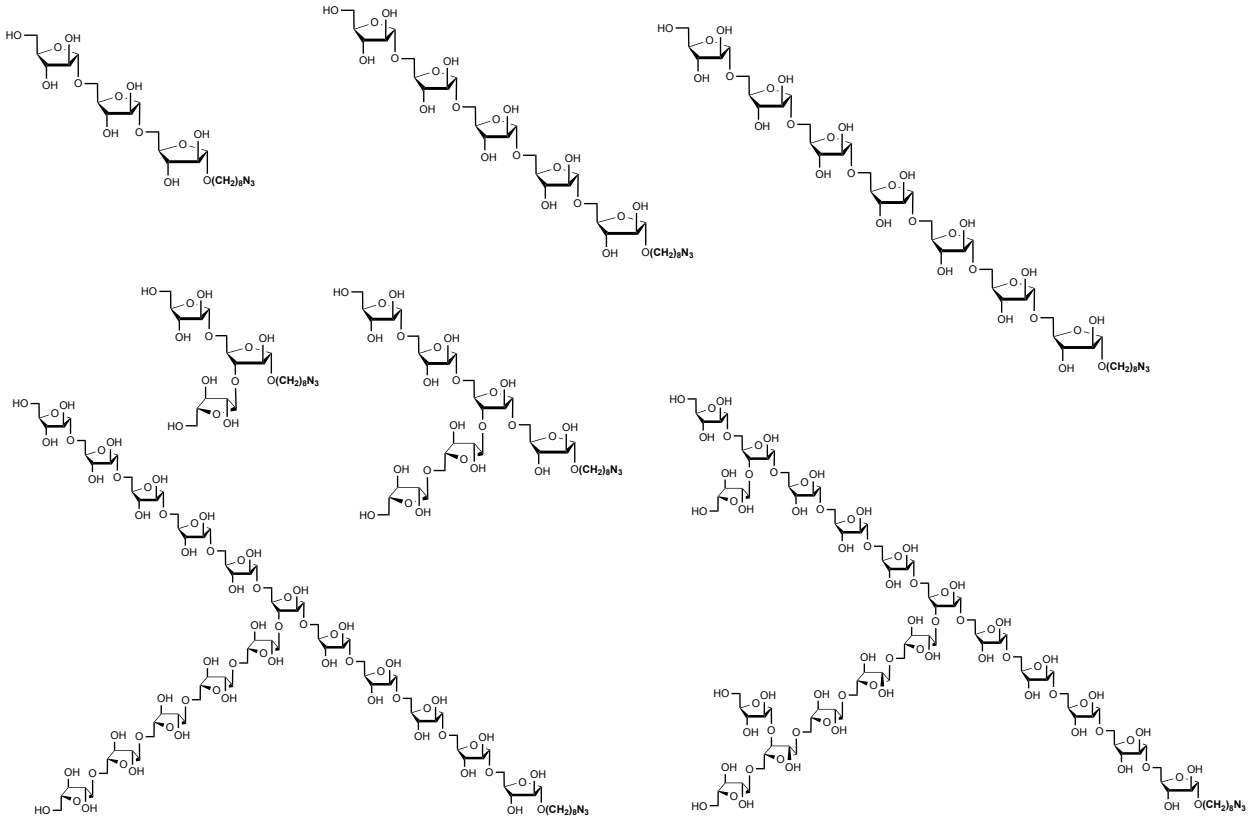


<u>Figure S9</u>: Complementation of the *M. smegmatis sucT* mutant with the *Rv1565c* gene from *M. tuberculosis*.

The amount of succinates and arabinose residues in the same LAM samples prepared from the WT, mutant, and *MSMEG_3187* and *Rv1565c* complemented mutant strains were quantified using the butanolysis procedure described in the Supplementary Methods. The *MSMEG_3187* and *Rv1565c* genes were both expressed from the *hsp60* promoter in the replicative plasmid pMVGH1. Results are expressed as average ± SD succinate/arabinose molar ratios from three technical replicates.



<u>Figure S10</u>: Structures of the neoglycolipid acceptors used in the succinyltransferase assays.



Supplementary Methods

Analytical procedures - Determination of the monosaccharide composition of mAGP and LAM, and glycosyl linkage patterns followed earlier procedures (Kaur *et al*., 2007). Alditol acetates and per-*O*-methylated alditol acetates were analyzed by GC/MS on a Thermo Scientific TRACE 1310 Gas Chromatograph paired with a Thermo Scientific TSQ 8000 Evo Triple Quadrupole GC-MS/MS. Samples were run on a Zebron ZB-5HT Inferno 30 m x 0.25 mm x 0.25 µm capillary column (Phenomenex) at an initial temperature of 100°C. The temperature was increased to 150°C at a ramp rate of 20°C min^{-1}, then to 240°C at a ramp rate of 5°C min^{-1} and was held at this temperature for 3 min to be finally increased to 300°C at a rate of 30°C min^{-1} and held at the final temperature for 5 min. Data handling was carried out using the Thermo Scientific Chromeleon Chromatography Data System software.

Succinates were detected and quantified by GC/MS of the butyl succinate derivatives obtained from either 20 µg of purified LAM or 1 mg of mAGP after butanolysis with 2.2 M acetyl chloride in 1-butanol (100 µl) at 100°C for 1 h. Arabinose residues in the same butanolized samples were quantified upon conversion to their TMS derivatives using the Tri-Sil HTP reagent (Thermo Fisher Scientific). Samples dissolved in *n*-hexanes were analyzed by GC/MS (see instrument's description above). Samples were injected at an initial temperature of 60°C held for 1 min followed by a temperature increase to 275 °C at a ramp rate of 10°C min^{-1}. Retention times and mass spectra were compared with those of standard succinic acid and D-arabinose and quantified by comparison with the areas of internal standards, deuterated succinic acid (Sigma-Aldrich) and D-[1-^{13}C]arabinose, processed by the same method.

Digestion of 10 µg of native LAM with 2.5 µg *Cellulomonas gelida* endoarabinanase to analyze the non-reducing arabinan termini of LAM followed earlier procedures (Torrelles *et al*., 2004). Digestion reactions were carried in 50 µL of 20 mM Tris-HCl buffer pH 7.8 and incubated for 16 h at 37°C. Reactions were stopped with ethanol (50% vol/vol., final concentration) and dried. Digestion products were acetylated with 100 µl acetic anhydride and 100 µl pyridine for 2 h at 80°C and dried. Acetylated digestion products were extracted with $CHCl_3/H_2O$ (1:1, v/v) and analyzed by LC-MS on an Agilent 1260 Infinity chromatograph using a 2.1 mm × 150 mm (3.5 µm particle size) XBridge reverse phase C18 column (Waters) heated to 40°C. Separation was done with a flow rate of 0.32 mL min^{-1} using solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in methanol). 4 µl out of 50 µl of a sample were injected in 10% solvent B (90% solvent A) followed by a gradient over 25 min to 100% solvent B which was held for an additional 10 min. Only those sample components eluted by LC after 12 min and later were introduced into the Agilent 6224 time-of-flight (TOF) mass spectrometer equipped with dual electrospray ionization (dual ESI) operated in positive ion mode. Mass spectra were acquired at a rate of 1.32 spectra per sec from *m/z* 450 to *m/z* 3,200. Data were analyzed using the Agilent MassHunter software.

The presence of phosphoinositol caps on LAM was determined by LC/MS analysis of the digestion products obtained after the treatment of deacylated LAM with *Cellulomonas gelida* endoarabinanase. 150 µg of LAM was deacylated with 0.2 M sodium hydroxide for 30 min at 37°C and then neutralized to pH~7 with glacial acetic acid. Salts were removed by diluting the material in water and passing through an Amicon Ultra-0.5 Centrifugal Filter Unit (MWCO 3K, Millipore) three times at 20,000 x g at room temperature. Deacylated LAM was treated with 5 µg of *Cellulomonas gelida* endoarabinanase in 100 µl of water for 20 h at 37°C and directly analyzed using ultra-performance liquid chromatography (UPLC) on a Waters Acquity UPLC H-Class system coupled to a Bruker MaXis Plus QTOF MS instrument. Separation was performed in a gradient mode on Waters Acquity UPLC BEH C18 column (2.1 mm x 50 mm, 1.7 µm particle size) with a flow rate of 0.4 mL min-^{1} using solvent A (water), solvent B (acetonitrile) and solvent C (0.5 M ammonium acetate). Solvent C was kept at 2% during the whole gradient. 5 out of 100 µl of each sample were injected in 10% solvent B (88% solvent A, 2% solvent C). The proportion of solvent B was increased from 10% to 70% in 3 min, to 98% in 3.4 min, and held at 98% for 1.4 min. The post-time was held for 2 min. Data acquisition in the negative electrospray ion mode with a mass-to-charge ratio (*m/z*) range of 110-4,000 at 1Hz was performed in the full MS scan mode. Source settings were as follows: capillary voltage 3,500V, endplate offset 500V, nebulizer gas pressure 3 bar, drying gas flow 10 L min^{-1},

and drying temperature 300°C. Data was processed using the Bruker Compass 2.0 Data Analysis 4.4 software.

Digestion of mAGP with endogenous *M. smegmatis* endoarabinanase was performed by incubating 1.5 mg of mAGP from the WT, mutant and complemented mutant strains with 5 mg of crude cell wall protein extracts prepared from the corresponding strains (Dong *et al*., 2006) in MOPS buffer (pH 8) containing 10 mM $MgCl_2$ and 5 mM β-mercaptoethanol at 37°C for 24 h. Reactions were stopped by the addition of $CHCl_3/CH_3OH$ to reach a final $CHCl_3/CH_3OH/H_2O$ ratio of 10:10:3 (by vol.), followed by centrifugation at 7,000 x *g* for 15 min. The released oligoarabinosides were dried, re-suspended in 200 µl 50% ethanol and finally analyzed using the same method and instruments as for the analysis of phosphoinositol caps.

NMR experiments were performed at 298K with a cryo-probed Bruker DRX600 spectrometer (Karlsruhe, Germany) according to 2D ^{1}H-^{13}C HMQC sequences previously reported (Gilleron *et al*., 1999) and a Prodigy™ cryo-probed Bruker Avance-IV 400 MHz NEO spectrometer for the 2D ^{1}H-^{13}P HMQC sequences. Native molecules were dissolved in D_2O and analyzed in 200 x 5 mm 535-PP NMR tubes. Proton and carbon chemical shifts are expressed in ppm downfield from the signal of external acetone (δH 2.22 and δC 30.89).

Supplementary References

De, P., McNeil, M., Xia, M., Boot, C. M., Hesser, D. C., Denef, K., Rithner, C., Sours, T., Dobos, K. M., Hoft, D., and Chatterjee, D. (2018) Structural determinants in a glucose-containing lipopolysaccharide from Mycobacterium tuberculosis critical for inducing a subset of protective T cells. *J Biol Chem* **293**, 9706-9717

Dong, X., Bhamidi, S., Scherman, M., Xin, Y., and McNeil, M. R. (2006) Development of a quantitative assay for mycobacterial endogenous arabinase and ensuing studies of arabinase levels and arabinan metabolism in *Mycobacterium smegmatis*. *Appl Environ Microbiol* **72**, 2601-2605

Gilleron, M., Himoudi, N., Adam, O., Constant, P., Venisse, A., Riviere, M., and Puzo, G. (1997) *Mycobacterium sm*egmatis phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. *J Biol Chem* **272**, 117-124

Gilleron M., Nigou J., Cahuzac B.and Puzo G. (1999) Structural study of the lipomannans from *Mycobacterium bovis* BCG: characterisation of multiacylated forms of the phosphatidyl-myo-inositol anchor. *J Mol Biol* **285**, 2147-2160

Kaur, D., McNeil, M. R., Khoo, K. H., Chatterjee, D., Crick, D. C., Jackson, M., and Brennan, P. J. (2007) New insights into the biosynthesis of mycobacterial lipomannan arising from deletion of a conserved gene. *J Biol Chem* **282**, 27133-27140

Torrelles, J. B., Khoo, K. H., Sieling, P. A., Modlin, R. L., Zhang, N., Marques, A. M., Treumann, A., Rithner, C. D., Brennan, P. J., and Chatterjee, D. (2004) Truncated structural variants of lipoarabinomannan in Mycobacterium leprae and an ethambutol-resistant strain of *Mycobacterium tuberculosis*. *J Biol Chem* **279**, 41227-41239