Supplementary Information

EXPRESSION AND CHARACTERIZATION OF A *MYCOPLASMA GENITALIUM* GLYCOSYLTRANSFERASE IN MEMBRANE GLYCOLIPID BIOSYNTHESIS. POTENTIAL TARGET AGAINST MYCOPLASMA INFECTIONS

Eduardo Andrés, Núria Martínez, Antoni Planas

Laboratory of Biochemistry, Bioengineering Department, Institut Químic de Sarrià, Universitat Ramon Llull, Barcelona, Spain

1.- Cloning of *M. genitalium* ORFs *mg025, mg060,* and *mg517* in *E. coli*

a) Cloning from M. genitalium genome

M.genitalium grown at 37ºC, 5% CO_2 to stationary phase (6 days) was scrapped, washed, and resuspended in Lysis buffer A (67mM Tris-Cl pH8.8, 16.7 mM ammonium sulphate, 5 mM β-mercaptoethanol, 6.7 mM $MgCl_2$, 6.7 mM EDTA pH8.0, 1.67 μM SDS and 50 μg/ml proteinase K). After 1h incubation at 37ºC for cell disruption and treatment at 80ºC for 10 min for proteinase K inactivation, the whole crude mixture was used as template in a hot start polymerase chain reaction (PCR) using primers *#1* to *#6* (Table S1) partially complementary to each ORF, and incorporating *NdeI* and *BamHI* restriction sequences for ligation. The three amplified genes were digested with both restriction enzymes and ligated (T4 DNA ligase) into a pUC18 vector previously digested with *NdeI* and *BamHI* and purified by agarose gel electrophoresis (plasmids pUC18-*mg025*, pUC18-*mg060*, and pUC18-mg51)7. Electrocompetent *E. coli* DH5α cells were transformed with the ligation mixtures and selected transformants verified by DNA sequencing using the Big Dye Terminator kit (Applied Biosystems).

b) Codon adaptation M.genitalium ORFs

All Trp-coding UGA codons present in *mg025* (2 codons), *mg060* (5 codons) and *mg517* (4 codons) were mutated to UGG for recombinant protein expression in *E. coli*.

mg025.- The two UGA codons (W218 and W273) were modified using the Quick Change® methodology (Stratagene) in two rounds of mutagenesis using the pUC18-*mg025* vector with primers #7 and #8 for the *stop*218W mutation, and primers #9 and #10 for the *stop*273W mutation.

mg060.- The five UGA codons (W88, W126, W131, W227, and W287) were replaced as follows. Two separate PCR amplifications of the 5' (from nucleotide -4 to 280 for stop88W mutation, primers #11 and #12) and 3' (from nucleotide 631 to +10 for stop227W and stop287W mutations, primers #13 and #14) regions of the gene (pUC18-*mg060* as template) yielded two fragment incorporating three mutations. The products of both PCR reactions were used to amplify the entire gene using the original pUC18-*mg060* vector as template to afford the gene with three mutations. After restriction digestion, the fragment was ligated into *NdeI*/*BamHI* digested pUC18 vector. Finally, the last two mutations,

*stop*126W and *stop*131W, were introduced by the Quick Change® methodology using primers #15 and #16.

mg517.- The four UGA codons (W170, W197, W268, and W275) were replaced following a similar combined strategy. The *stop*197W mutation was first introduced by the Quick Change® methodology on pUC18-*mg517* vector as template using primers #17 and #18 to give plasmid pUC18-*mg517*(stop197W). Two PCR amplifications of the 5' (from nucleotide -4 to 532 for stop170W mutation, primers #19 and #20) and 3' (from nucleotide 796 to +6 for stop268W and stop275W mutations, primers #21 and #22) regions of the gene (pUC18-*mg517* as template) carried out the other three mutations. The products of both PCR reactions were used to amplify the entire gene with the previously obtained pUC18-*mg517*(*stop*197W) vector as template. After restriction digestion, the fragment was ligated into *NdeI*/*BamHI* digested pUC18 vector. All mutations were verified by DNA sequencing.

c) Subcloning into pET38b(+) vector

The mutated genes in pUC18 vectors were subcloned into pET38b(+) between *NdeI* and *AvrII* sites for protein expression of the full length proteins. Each individual gene was amplified by PCR using flanking primers that incorporate the *NdeI* and *AvrII* sites respectively: primers #23 and #24 for *mg025*, primers #25 and #26 for *mg060*, and primers #27 and #28 for *mg517*. After digestion, each fragment was ligated into *NdeI*/*AvrII*-digested pET38b(+) vector, resulting in pET38b(+)-*mg025*, pET38b(+)-*mg060*, and pET38b(+)-*mg517*. The final constructs were again verified by DNA sequencing. *E. coli* BL21(DE3) were transformed with each vector for protein expression.

d) Subcloning mg517 into pET44b(+) vector

The stop codon of pET38b(+)-*mg157* was removed and a *XhoI* site introduced by the Quick Change® methodology using primers #29 and #30. The resulting plasmid was digested with *NdeI/XhoI* and the gene fragment ligated into *NdeI/XhoI* digested pET44b(+) vector. The final construct was verified by DNA sequencing. *E. coli* BL21(DE3) cells were transformed for protein expression.

2.- Construction of *M. genitalium mg517* knock-out.

The suicidal plasmid pΔMG-517, which contains the *tetM438* selection marker (tetracycline resistance) (ref.1) enclosed by 1kb upstream and downstream flanking regions of the *mg517* gene was constructed for gene replacement by homologous recombination in order to obtain a MG517 null mutant.

Genomic DNA of the wild-type *M. genitalium* strain was used as a template for PCR amplifications. The 1 kb PCR fragment encompassing the 5' upstream region of *mg517* gene was amplified using primers #31 and #32, which incorporate *Sal*I and *Eco*RI restriction sites at their ends. A second 1 kb fragment containing the 3' downstream region of *mg517* gene was also amplified using primers #33 and #34, with *BamH*I and *Nde*I restriction sites at their ends. After restriction digestion, both PCR fragments were ligated with a 2 kb fragment containing the tetM438 selection marker and *Sal*I/*Nde*I digested pUC18 vector. The tetM438 selectable marker was released from the pMTnTetM438 plasmid (ref.1) by digestion with EcoRI and BamHI. *E. coli* DH5α competent cells were transformed with the ligation reaction, and the final pΔMG-517 plasmid isolated.

Transformation of *M. genitalium* strain G37 with the suicide plasmid pΔMG-517 was performed by electroporation as previously described (ref.1,2). Cells were grown to mid-log phase, removed with a cell scrapper, the suspension passed through a 0.45 μm filter, and recultivated for 24h with fresh medium. Then, the medium was removed and the cells washed with electroporation buffer (8 mM HEPES pH 7.2, 272 mM sucrose) and resuspended to a cell density of 10^9 cfu/mL. Plasmid DNA was added and electroporated in a BTX Electro Cell manipulator 600. Transformants were recovered by cultivation in liquid SP4 medium for 2 h at 37ºC, and then plated on solid SP4 medium containing tetracycline until appearance of colonies. Genomic DNAs of the transformant colonies were digested with *Hind*III and electrophoresed on 0.8% agarose gels, and DNA fragments transferred to nylon membranes for Southern blot analysis as previously described (ref.2).

Table S1. Primers used for cloning, mutagenesis, and construction of vectors.

	Experiment	Primer	Sequence
#1	Genomic cloning of *mg025* from *M.gen.*	*P1-direct*	5'-TGGGAATTCCATATGGATAAACTTGTTAGTAT ATT-3'
#2		*P1-reverse*	5'-TAACGCGGATCCTTAGTTATCTGATTTAGAT TCC-3`
#3	Genomic cloning of *mg060* from *M.gen.*	*P2-direct*	5'TGGGAATTCCATATGAAATTATCTGTA ATTATACCTACTT-3'
#4		*P2-reverse*	5'-TAACGCGGATCCTTATTCAGAACTGTTGAAAT-3'
#5	Genomic cloning of *mg517* from *M.gen.*	*P3-direct*	5'-TGGGAATTCCATATGGATA AACTTGTTAGTATATT-3'
#6		*P3-reverse*	5'-TAACGCGGATCCTTAGTT ATCTGATTTAGATTCC-3'
#7	*mg025* SDM for Stop218W	*mg025-218W*	5'- CTACAACAACTTCTTGGGAAAATCCAAATAAATTTAATGC-3'
#8		*mg025-218W-comp*	5'-GCATTAAATTTATTTGGATTTTCCCAAGAAGTTGTTGTAG-3'
#9	*mg025* SDM for Stop273W	*mg025-273W*	5'-CCATAAACCAGCTTACACTTGGTTACCTAAACCTTTAGCG-3'
#10		*mg025-273W-comp*	5'-GCGTAAAGGTTTAGGTAACCAAGTGTAAGCTGGTTTATGG-3'
#11	*mg060* Mutagenesis PCR1 (stopW88)	*mg060-P1dir*	5'-TGGGAATTC CATATGAAATTATCTGTAATTATACCTACTT-3'
#12		*mg060-P1rev*	5'-CCAATACTGTTACCCATTCCCCTTTTGC-3'
#13	*mg060* Mutagenesis PCR2 (stop227W/stop287W)	*mg060-P2dir*	5'-GGTAACTCTATGAGTATCCCTTGACACAGTTCTAGGTTTG-3'
#14		*mg060-P2rev*	5'-CCGGGGATCCTTATTCAGAACTGTTGAAATAACGTTTCAACC AGAAGTTATAC-3'
#15	*mg060* SDM for stop126W and stop131W	*mg060-126/131W*	5'-GACTACTATAAATGCTGGAAAAAGTTTTTGTGGAAAATTCC TACC-3'
#16		*mg060-126/131W comp*	5'- GGTAGGAATTTTCCACAAAAACTTTTTCCAGCATTTATAG TAGTC-3'
#17	*mg517* SDM for stop197W	*mg517-197W*	5'-GGTTATTTGAAGATATACCAATCTGGTATCCGATGTTTTTTT CATC-3'
#18		*mg517-197W comp*	5'-GATGAAAAAAACATCGGATACCAGATTGGTATATCTTCAAAT AACC-3'
#19	*mg517* Mutagenesis PCR1 (stop170W)	*mg517-P1dir*	5'-TGGGAATTCCATATGGATAAACTTGTTAGTATATT-3'
#20		*mg517-P1rev*	5'-CTGTGTTTATAACAATATTCCAAATATATTGATCATTATTTT-3'
#21	*mg517* Mutagenesis PCR1 (stop268W/stop275W)	*mg517-P2dir*	5'-GCTAGGTTTTGGAGAAGGCAAATGTTTGTTTGGTTT GCACTT TTC-3'
#22		*mg517-P2rev*	5'-TTAGCGGGATCCTTAGTTATCTGATTTAGATTCC-3'
#23	*mg025* subcloning into pET38b(+)	*mg025-pET38b dir*	5'-TGGGAATTCCATATGAAATTATCTGTAATTATACCT-3'
#24		*mg025-pET38b rev*	5'-TTTGCGCCTAGGCCGGGGATCCTTACTTAGG-3'
#25	*mg060* subcloning into pET38b(+)	*mg060-pET38b dir*	5'-TGGGAATTCCATATGAAATTATCTGT AATTATACCT-3'
#26		*mg060-pET38b rev*	5'-TAACGCCCTAGGCCGG GGATCCTTATTCAG-3'
#27	*mg517* subcloning into pET38b(+)	*mg517-pET38b dir*	5'-TGGGA ATTCCATATGGATAAACTTGTTAGTATATT-3'
#28		*mg517-pET38b rev*	5'-TAACGCCCTAGGCCGGGGATCCTTAGTTATCTG -3'
#29	*mg517* SDM for subcloning into pET44b(+)	*mg517-XhoI*	5'-GGAATCTAAATCAGATAACCTCGAGTCCCCGGCCTAGG-3'
#30		*mg517-XhoI comp*	5'-CCTAGGCCGGGGACTCGAGGTTATCTGATTTAGATTCC-3'
#31	*mg 517* upstream recombination arm	*mg517RA5'*	5'-TTATGAAGTCGACCTGTTAGATTTGG-3'
#32		*mg517RA3'*	5´-TGATTATGAATTCCAGTTTATCCATT-3'
#33	*mg 517* downstream recombination arm	*mg517LA5'*	5'-TCAGATGGATCCTTAAGTTCTCTTTAAA-3'
#34		*mg517LA3'*	5'-TTTCACCATATGTAAATTCTTTTTGC-3'

3.- Glycosyltransferase activity in *M. genitalium*

The lipid composition of the plasma membrane was analyzed by direct extraction of *M.genitalium* cells grown in SP4 medium to stationary phase with chloroform/methanol/water and TLC separation. In addition to cholesterol and phospholipids, two major glycolipids were detected (GL1 and GL2, Figure S1a). Growth of *M. genitalium* in the presence of ^{14}C-labelled glucose (Figure S1b) showed incorporation of the label into the two glycolipids, GL1 and GL2. They are sequentially synthesized from diacylglycerol by the action of GT activities. Incubation of *M. genitalium* cells (both wet cells after centrifugation or lyophilized cells) with a mixture of labeled UDP-glucose (UDP-[$^{14}C_6$]-Glc) and dioleoylglycerol (DOG) as donor and acceptor substrates in the presence of sodium dodecylsulphate (SDS) as solubilizing detergent followed by autoradiography of the TLC-separated lipid extract showed two labeled spots corresponding to GL1 and GL2 (Figure S1c). In the absence of DOG as acceptor in the reaction mix, no radioactive incorporation into lipidic products was observed. The isolated ^{14}C-labelled GL1 product was then used as acceptor and incubated with *M.genitalium* cell pellet in the presence of SDS and cold UDP-Glc, showing that GL1 is converted to GL2, thus demonstrating the sequential action of the two GT activities.

Glycosyltransferase activities are localized in the plasma membrane as shown by activity assays on different cellular fractions. The cell pellet after growing *M. mycoplasma* to stationary phase in SP4 medium was sonicated and centrifuged to separate the citoplasmic (soluble) fraction from membranes (insoluble fraction). As control, intact cells, and heat inactivated membrane fraction (1h at 90ºC) were also prepared. The four samples were used in the GT activity assay by incubating them with UDP-[$^{14}C_6$]-Glc donor and DOG acceptor and in the presence of SDS at 28ºC for 1 h. The lipid extract was separated by TLC, all eluted compounds scrapped from the TLC plate (to separate unreacted UDP-[$^{14}C_6$]-Glc), and the radioactivity quantified by scintillation counting. Intact cells and membrane fraction incorporated the radioactivity into the lipid products whereas no radioactive products were detected for the cytoplasmic fraction and heat inactivated membranes, confirming that the GT activities are localized in the plasma membrane.

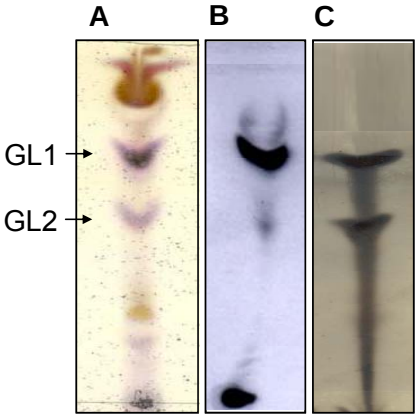
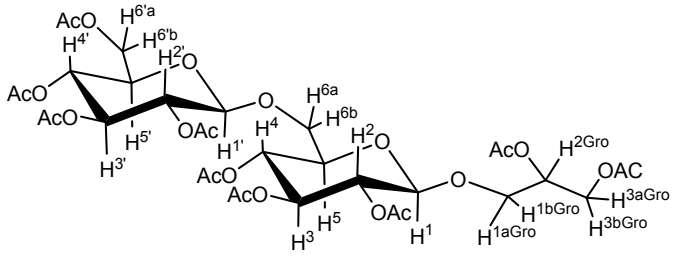


Figure S1.- Analysis of membrane lipids in *M. genitalium.* *A)* Lipid extract from cell grown in SP4 medium separated by TLC eluted with C/M/H_2O 65:35:4 and developed with α-naphtol in sulfuric acid/methanol 2:1 (v/v); *B)* ^{14}C-labeled glycolipids from lipid extracts of *M.genitalium* grown in ^{14}C-glucose supplemented SP4 medium by TLC-autoradiography; *C) In vitro* GT assay by incubating *M.genitalium* cells with UDP-[^{14}C]Glc and DOG in the presence of SDS (TLC-autoradiography).

4.- Structure determination of glycolipids GL1 and GL2 by NMR spectroscopy

a) Glycosidic moiety of GL2 produced in vivo by E.coli BL21(DE3)-pET38b(+)-mg517.- Monodimensional ^{1}H-MNR and COSY spectra showed two characteristic anomeric signals at δ_H 4.59 ppm and 4.49 ppm. The analysis of both spin systems starting from each anomeric signal, plus that of the glycerol moiety in the COSY spectrum (Figure S3) allowed the assignment of all ^{1}H-RMN signals and then the identification, through the HETCOR and HMQC spectra, of each carbon counterpart in the ^{13}C-RMN spectrum (Table S2).

Table S2. Assignment of ^{1}H-NMR and ^{13}C-NMR data of acetylated Glcβ1,6GlcβGro obtained by methanolysis and acetylation of the main glycolipid (GL2) produced *in vivo* by *E.coli* BL21(DE3) cells expressing *mg517*.



^{1}H-NMR		^{13}C-NMR	
Position	δ (ppm) (*J*(Hz))	Position	δ (ppm)
H1	4.49 ($J_{1,2}$ 8.1)	C1	100.6
H2	4.94 ($J_{2,1}$ 8.1, $J_{2,3}$ 9.6)	C2	71.1
H3	5.20 ($J_{3,2}$ 9.6, $J_{3,4}$ 9.3)	C3	72.7
H4	4.89 ($J_{4,3}$ 9.3, $J_{4,5}$ 9.9)	C4	69.0
H5	3.67 *m*(1)	C5	73.4
H6a,b	3.86/3.64 *m*	C6	67.3
H1'	4.59 ($J_{1,2}$ 8.1)	C1'	100.6
H2'	5.00 ($J_{2,1}$ 8.1, $J_{2,3}$ 9.3)	C2'	71.1
H3'	5.20 ($J_{3,2}$ 9.3, $J_{3,4}$ 9.3)	C3'	72.6
H4'	5.08 ($J_{4,3}$ 9.3, $J_{4,5}$ 9.6)	C4'	68.2
H5'	3.71 *m*	C5'	71.9
H6'a,b	4.30/4.14 *m*	C6'	61.8
H1a,b$_{Gro}$	3.97/3.66	C1$_{Gro}$	69.8
H2$_{Gro}$	5.18	C2$_{Gro}$	68.1
H3a,b$_{Gro}$	4.14/4.30	C3$_{Gro}$	62.4

(1) *m*: multiplet, coupling constants not assigned

The coupling constants $J_{1,2}$ = 8.1 Hz in both sugar rings indicates a glucose–glucose–glycerol structure with β linkages, consistent with the inverting mechanism proposed for this family 2 glycosyltransferase. Moreover the displacement of both ^{1}H and ^{13}C signals of H6 and C6 towards a higher field compared to its H6' and C6' equivalents indicates a (1→6) linkage that was confirmed by NOESY experiments (Figure S4), where H1' showed spatial correlation with signals of both H6a and H6b. Therefore the polar head structure of the mayor glycolipid (GL2) found in recombinant *E.coli* is 3-O-(D-glucopyranosyl-β(1→6)-D-glucopyranosyl)-*syn*-glycerol.

Figure S2. COSY spectrum of acetylated Glcβ1,6GlcβGro

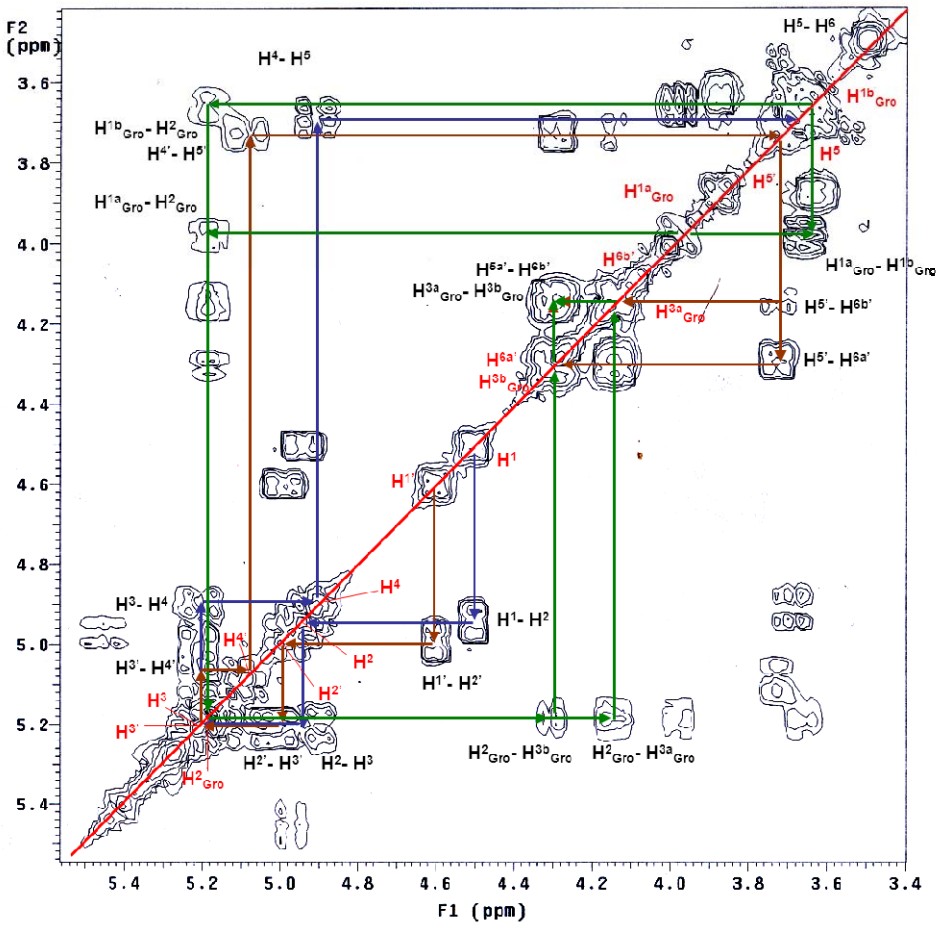
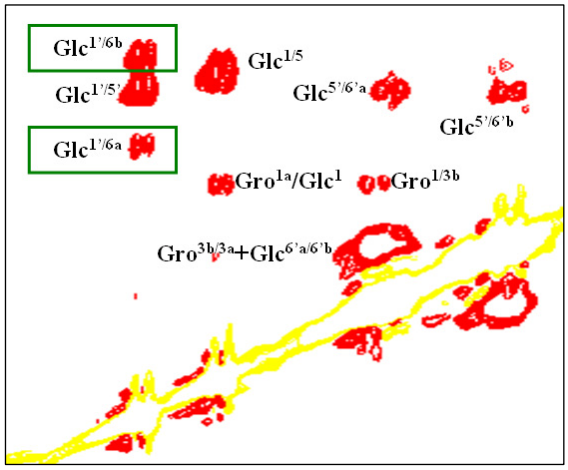


Figure S3. NOESY spectrum of acetylated Glcβ1,6GlcβGro.



b) Glycolipids GL1 and Gl2 produced in vitro by purified GT MG517.- The assignment of all ^{1}H and ^{13}C signals for both glycolipids (Table S3 for GL1 (MGlcDOG) and Table S4 for GL2 (DGlcDOG)) was performed by a similar sequence of NMR experiments as described above. NMR data for peracetylated glycolipids GlcβDAG and Glcβ1,6GlcβDAG have been reported in ref. 3,4, and the de-acetylated disaccharide in ref.5.

Table S3. Assignment of ^{1}H-NMR and ^{13}C-NMR data of GlcβDOG (GL1) obtained by *in* vitro synthesis with purified GT MG517, UDPGlc, DOG, and SDS.

^{1}H-NMR		^{13}C-NMR	
Position	δ (ppm)	Position	δ (ppm)
H1	4.29	C1	103.4
H2	3.19	C2	73.6
H3	3.37	C3	76.4
H4	3.34	C4	70.3
H5	3.30	C5	76.0
H6a,b	3.79/3.66	C6	62.0
$H1a,b_{Gro}$	3.90/3.71	$C1_{Gro}$	67.6
$H2_{Gro}$	5.23	$C2_{Gro}$	70.0
$H3a,b_{Gro}$	4.35/4.17	$C3_{Gro}$	62.4

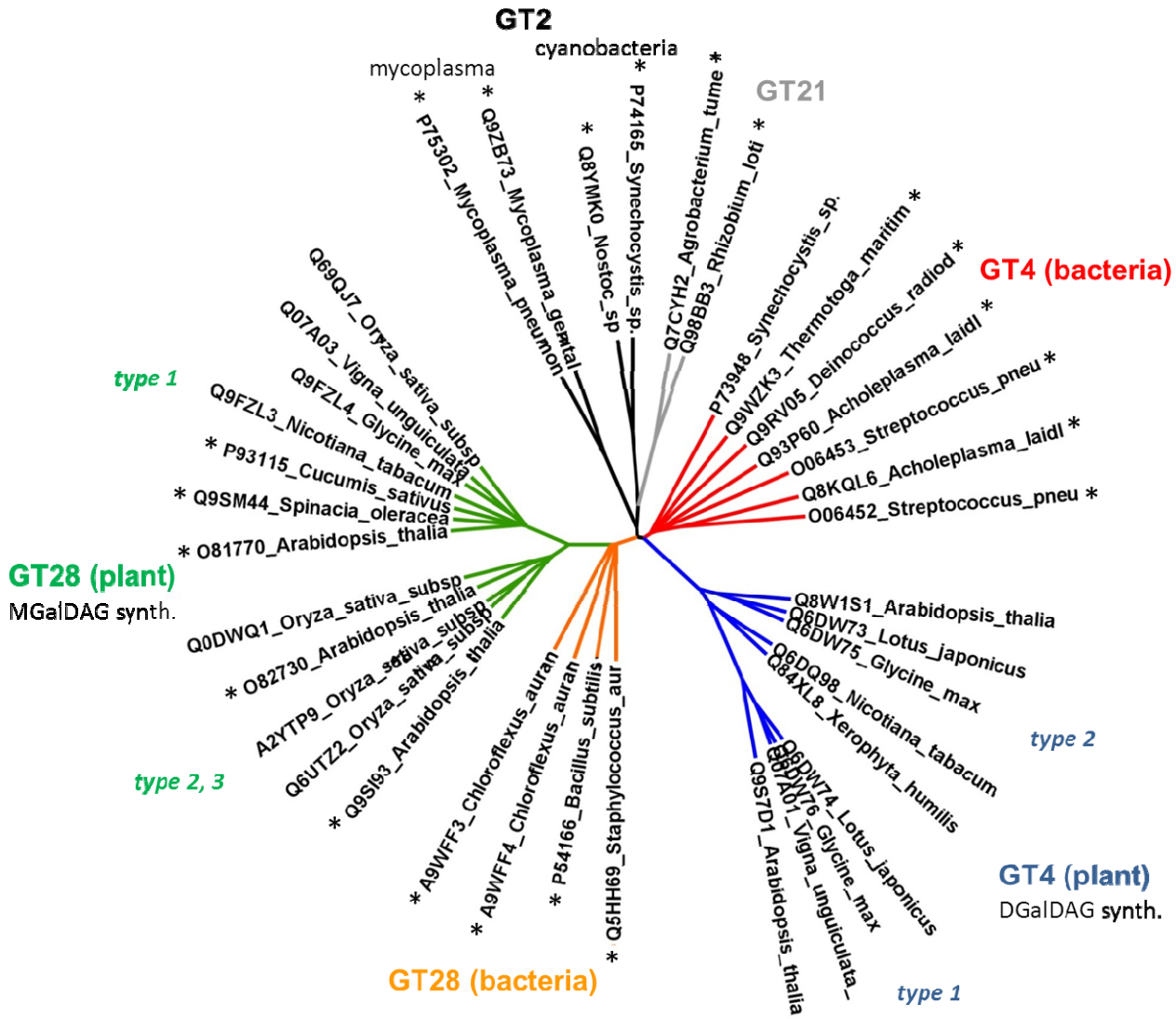
Table S4. Assignment of ^{1}H-NMR and ^{13}C-NMR data of Glcβ1,6GlcβDOG (GL2) obtained by *in vitro* synthesis with purified GT MG517, UDPGlc, DOG, and SDS.

^{1}H-NMR		^{13}C-NMR	
Position	δ (ppm)	Position	δ (ppm)
H1	4.28	C1	102.9
H2	3.21	C2	73.4
H3	3.39	C3	76.3
H4	3.36	C4	70.4
H5	3.30	C5	76.0
H6a,b	4.10/3.78	C6	68.6
H1’	4.36	C1’	103.2
H2’	3.25	C2’	73.6
H3’	3.39	C3’	76.3
H4’	3.30	C4’	70.4
H5’	3.36	C5’	76.0
H6’a,b	3.80/3.69	C6’	61.8
$H1a,b_{Gro}$	3.91/3.71	$C1_{Gro}$	67.7
$H2_{Gro}$	5.2	$C2_{Gro}$	69.8
$H3a,b_{Gro}$	4.33/4.18	$C3_{Gro}$	62.2

5.- Phylogenetic tree for glycosyldiacylglycerol synthases

Glycosyldiacylglycerol synthases (both characterized and annotated) sequences were retrieved from UNIPROT data base and aligned with PROMALS multiple sequence alignment server (J. Pei and N. V. Grishin (2007) *Bioinformatics 23*, 802-808). A neighbor-joining phylogenetic tree was reconstructed from this alignment using the BLOSUM62 score-matrix as a distance metric.

Figure S4. Unrooted phylogenetic tree for glycosyldiacylglycerol synthases (UNIPROT sequences, both characterized (marked with *) and annotated).



The phylogenetic analysis based on amino acid sequence similarity shows that both processive GT21 and GT2 mycoplasma enzymes are closely related to cyanobacterial MGlcD synthases, whereas GT28 processive bacterial enzymes are close to non-processive bacterial GT28 enzymes. Non-processive synthases form isolated groups in plant GT4, bacterial GT4, and plant GT28. Type 1 and type 2 plant enzymes in both GT4 and GT28 are further grouped together on each respective branch in the phylogenetic tree.

6.- References

1. Pich, O.Q., Burgos, R., Planell, R., Querol, E., Piñol, J. (2006) *Microbiology* **152,** 519-527
2. Burgos, R., Pich, O.Q., Querol, E., Piñol, J. (2007) *J. Bacteriol.* **189**, 7014-7023
3. Jorasch, P., Wolter, F.P., Zähringer, U., Heinz. E.(1998) *Mol. Microbiol.* **29**, 419-430
4. Jorasch, P., Warnecke, D.C., Lindner, B., Zähringer, U., Heinz, E. (2000) *Eur. J. Biochem.* **267**, 3770-3783
5. Wicke, C., Hüners, M., Wray, V., Nimtz, M., Bilitewski, U., Lang, S. (2000) *J. Nat. Prod.* 63, 621-626