Design of a new biological platform for the production of glycoglycerolipids

Nuria Orive Milla

UNIVERSITAT RAMON LLULL

DOCTORAL THESIS

Title	**Design of a new biological platform for the production of glycoglycerolipids**
Presented by	**Nuria Orive Milla**
Centre	**IQS School of Engineering**
Department	**Bioengineering**
Directed by	**Dra. Magda Faijes and Dr. Antoni Planas**

Amb el suport del programa nacional BFU2016-77427-C2-1-R i BIO2013-49022-C2-1-R
del Ministerio de Economía y Empresa, d'Espanya i els Ajuts IQS pels estudis de Doctorat i la Societat
Econòmica Barcelonesa d'Amics del País

Good books don’t give up all their secrets at once

Stephen King

There's nothing you can do that can't be done
Nothing you can sing that can't be sung.
Nothing you can say but you can learn how to play the game.
It's easy.

Nothing you can make that can't be made.
No one you can save that can't be saved.
Nothing you can do but you can learn how to be you in time.
It's easy.

Nothing you can know that isn't known.
Nothing you can see that isn't shown.
Nowhere you can be that isn't where you're meant to be.
It's easy.

The Beatles

Table of contents

List of figures ..... **VII**

List of tables ..... **IX**

Publications and communications ..... **XI**

Acknowledgements ..... **XIII**

Summary ..... **XIX**

Resum ..... **XXI**

Resúmen ..... **XXIII**

List of abbreviations ..... **XXV**

List of strains ..... **XXVII**

1. Introduction **3**

1.1. Glycolipids **3**

1.2. Applications of glycolipids **5**

1.2.1. Glycolipids as biosurfactants 5

1.2.2. Glycolipids as drug delivery systems 7

1.2.3. Glycolipids as antimicrobial agents 8

1.2.4. Vaccine adjuvants 9

1.2.5. Glycolipids in tumor pathogenesis 11

1.3. Glycolipid production **11**

1.3.1. Chemical synthesis 12

1.3.2. Enzymatic synthesis 13

1.3.3. Cell factory approach 14

1.4. Framework of the project **16**

1.4.1. Fatty acid metabolism 16

1.4.2. Glycosidic metabolism 23

2. Objectives **27**

3. Proposed strategies to increase the production of GGL **31**

3.1. Designed strategies **31**

3.2. Molecular biology to follow the strategies 34

3.3. Analytical methods used to characterize the different strains obtained **35**

3.3.1 Glycolipid quantification 35

3.3.1.1 Setting-up the analytical method 36

3.3.2 Glycosyltransferase activity determination 36

4. Strategy 1: Increase of the DAG availability by removing competing reactions **41**

4.1. Study basis **41**

4.2. Obtaining knockout strains **43**

4.2.1. Δ*fadE* 43

4.2.2. Δ*tesA* 44

4.2.3. Δ*tesA* Δ*fadE* 45

4.3. Strain characterization **46**

4.4. GGL production 47

5. Strategy 2: Increase of FA availability by modulating transcription factors 53
5.1. Strategy basis 53
5.2. Molecular biology related to the strategy 54
5.2.1. Obtaining the Δ*tesA* Δ*fabR* strain 54
5.2.2. FadR plasmid 55
5.3. Strain characterization 55
5.4. GGL Production 56
5.5. Lipid profile analysis 57

6. Strategy 3: Increase the conversion of acyl donors to PA by overexpressing PlsC and PlsB 61
6.1. Strategy basis 61
6.2. Molecular biology related to the strategy 61
6.2.1. Study basis 61
6.2.2. Low copy number plasmids 62
6.2.2.1. $PlsC^{L}$ 62
6.2.2.2. $plsB^{L}$ 63
6.2.2.3. $plsC^{L}$·$plsB^{L}$ 63
6.2.3. High copy number plasmids 64
6.2.3.1. $plsB^{H}$ 65
6.2.3.2. $plsC^{H}$ ·$plsB^{H}$ 66
6.3. Study of the GGL production using different acyltransferases 67
6.4. Study of the acyltransferase enzymatic activity 68

7. Strategy 4: Increase the production of phosphatidic acid from phospholipids 73
7.1. Strategy basis 73
7.2. Molecular biology related to the strategy 74
7.2.1. mg517·cdh 74
7.2.2. $plsCxpgpB^{H}$ 75
7.3. Testing the enzymatic activities 76
7.3.1. CDH enzymatic activity 76
7.3.2. Phosphatidylglycerophosphatase enzymatic activity 77
7.4. Effect of CDH and PgpB on the GGL production 78

8. Strategy 5: Increasing UDP-Glc availability 83
8.1. Strategy basis 83
8.2. Molecular biology related to the strategy 84
8.2.1. GalU 84
8.2.2. Obtaining Δ*tesA* Δ*ushA* strain 84
8.3. GGL production 86
8.3.1. GalU overexpression 86
8.3.2. GGL production in Δ*tesA* Δ*ushA* 87
8.5. Study of the different acyltransferases 88
8.6. The effect of *fadR* overexpression in Δ*tesA* Δ*ushA* strain 90
8.7. CDH effect on the production 91
8.8. Study of the $plsCxpgpB^{H}$ in Δ*tesA* Δ*ushA* 93

9. Further characterization of the engineered strains 99
9.1. Glycosyltransferase characterization..... 99
9.1.1. Obtaining mg517xmCherry 99
9.2.2. MG517 characterization..... 100
9.3. Phospholipid analysis DAG..... 103

10. Discussion 109

11. Phosphatidylserine decarboxylase library..... 115
11.1. Introduction..... 115
11.2. Regulation of *psd* 116
11.3. Design of the library..... 117
11.4. Obtaining the KI plasmid..... 119
11.5. Obtaining the library..... 120
11.6. Library characterization 121

12. Conclusions 127

13. Material and methods 131
13.1. Molecular biology 131
13.1.1. Genomic DNA extraction and purification 131
13.1.2. Plasmid DNA extraction and purification 131
13.1.2.1. High Purity Plasmid Miniprep kit (Clinisciences, Ref. NB-03-002) 131
13.1.2.2. InnuPREP Plasmid Mini Kit 2.0 (AnalytikJena, Ref. 845-KS-5041250)..... 131
13.1.3. DNA Gel extraction..... 131
13.1.4. DNA quantification 131
13.1.4.1. Qubit..... 131
13.1.4.2. Nanodrop 132
13.1.5. Subcloning techniques..... 132
13.1.5.1. Gibson assembly..... 132
13.1.5.2. CPEC assembly..... 132
13.1.5.3. Golden Gate Assembly 133
13.1.5.4. Single Strain Annealing..... 134
13.1.6. Weight marker..... 134
13.1.7. Competent cells..... 134
13.1.7.1. Electrocompetent cells..... 134
13.1.7.1.1. Electrocompetent cells to maintain in-80°C..... 134
13.1.7.1.2. Fresh electrocompetent cells 135
13.1.7.2. Chemical competent cells..... 135
13.1.8. DNA digestion..... 136
13.1.8.1. DpnI 136
13.1.8.2. Double digestion..... 136
13.1.9. Ligation 136
13.1.10. Datsenko and Wanner protocol to obtain KO strains 136
13.1.10.1. Procedure: transform with pKD46..... 136
13.1.10.2. Procedure: insert linear DNA..... 137
13.1.10.3. Procedure: removal antibiotic resistance cassette 138

13.2. Cultures 139
13.2.1. Preinocculum 139
13.2.2. Minimal medium 139
13.2.3. Antibiotics 140
13.2.4. IPTG 141
13.2.5. Preparation of cultures for further analysis 141
13.2.6. Lyophilization 141
13.3. GGL analysis 141
13.3.1. Folch extraction 141
13.3.2. Anthrone reaction 142
13.3.2.1. Sample preparation 142
13.3.2.2. Standard curve 142
13.3.2.3. Anthrone reagent preparation 143
13.3.2.4. Anthrone assay 143
13.3.3. TLC analysis 143
13.3.3.1. One-dimension TLC 143
13.3.3.1.1. Mobile phase preparation 143
13.3.3.1.2. Sample preparation 143
13.3.3.1.3. Standards 143
13.3.3.1.4. Thin Layer chromatography procedure 143
13.3.3.2. Two-dimension TLC 144
13.4. Enzymatic activities 144
13.4.1. Glycosyltransferase activity 144
13.4.1.1. Preparation of cell extract for glycosyltransferase activity 144
13.4.1.2. Glycosyltransferase determination 144
13.4.1.3. Quantification of the formed products 145
13.4.2. Acyltransferase activity 145
13.4.2.1. Cell extract preparation for the acyltransferase enzymatic activity assay 145
13.4.2.2. Acyltransferase reaction 145
13.4.3. PgpB activity 145
13.4.4. CDH activity 146
13.5. Lipid analysis 146
13.5.1. Lipid profile by gas chromatography 146
13.5.2. Lipid profile identification by GC-MS 147
13.5.3. Phospholipid analysis by radiolabeled ^{14}C-acetate 147
13.6. Protein quantification 148
13.6.1. MG517 quantification by fluorescence 148
13.6.2. Total protein quantification 149
13.6.2.1. BCA 149
13.6.2.2. Bradford 149

14. References 153

15. Annexes 167
15.1. Strains used in this project 167
15.1.1. BL21 Star (DE3) (WT) 167

15.1.2. BL21 Star (DE3) Δ*fadE* ... 167
15.1.3. BL21 Star (DE3) Δ*tesA* ... 168
15.1.4. BL21 Star (DE3) Δ*tesA* Δ*fadE* ... 168
15.1.5. BL21 Star (DE3) Δ*tesA* Δ*fabR* ... 169
15.1.6. BL21 Star (DE3) Δ*tesA* Δ*ushA* ... 169
15.1.7. Summary ... 170
15.2 Production summary of the strains ... 171
15.3. Plasmids used in this project ... 173
15.4. Primers used in this project ... 177
15.5. Characterization of mg517xmCherry protein ... 183
15.6. Gas chromatography chromatograms and confirmation ... 184
15.6.1. Myristic acid ... 185
15.6.2. Palmitic acid ... 185
15.6.3. Palmitoleic acid ... 186
15.6.4. 2 - hexyl - cyclopropaneoctanoic acid ... 186
15.6.5. Stearic acid ... 187
15.6.6. Oleic acid ... 187
15.6.7. 2 - octyl - cyclopropaneoctanoic acid ... 188
15.6.8. Subproducts ... 188
15.6.8.1. Benzenepropanoic acid ... 188
15.6.8.2. 9-octadecenamide (Oleamide) ... 189
15.7. Areas and percentage of the radiometric assay ... 191
16. Publication ... 193

List of figures

Figure 1. Glycolipid structure 5
Figure 2. Biosurfactant structure examples 6
Figure 3. Barriers that drug delivery systems need to overcome to reach the target 7
Figure 4. Types of nanotechnology used for drug delivery 8
Figure 5. Structures of α-galactosylceramides 10
Figure 6. Structure of the GGL isolated from *Meiothermus taiwanensis* 10
Figure 7. Example of how to synthesize glycolipids by chemical synthesis 12
Figure 8. Glycoglycerolipids synthesis 13
Figure 9. Sequential reaction of MG517 14
Figure 10. Metabolic pathway to synthetize glycoglycerolipids in *E. coli* 16
Figure 11. Lipid metabolic pathway in *E. coli* 17
Figure 12. *Escherichia coli* UDP-Glc biosynthesis pathway 23
Figure 13. Metabolic engineering strategies to increase the production of GGL 33
Figure 14. Procedure to obtain knockouts 34
Figure 15. Anthrone reaction 35
Figure 16. Summary of the analytical procedure followed to quantify GGL 36
Figure 17. Reaction between C6-NBD-ceramide and UDP-Gal by the catalysis of MG517 37
Figure 18. Chromatogram of MG517 reaction 37
Figure 19. Overlap of the different chromatograms 37
Figure 20. *Escherichia coli*'s metabolic pathways for GGL precursors, UDP-Glc and DAG. 41
Figure 21. Designed primers to knockout of *fadE* in BL21 Star (DE3) 43
Figure 22. Agarose gels and sequencing of BL21 Star (DE3) Δ*fadE* 44
Figure 23. Agarose gels and sequencing of BL21 Star (DE3) Δ*tesA* 45
Figure 24. Agarose gels of BL21 Star (DE3) Δ*tesA* Δ*fadE* 46
Figure 25. Growth of ΔtesA/mg517-plsC^{H} 47
Figure 26. Microscopic analysis of the engineered strains 47
Figure 27. GGL production using MG517 and PlsC^{H} proteins 48
Figure 28. Glycolipid profile 49
Figure 29. Strategy used to increase the homologous sequences to recombine lineal DNA 54
Figure 30. Agarose gels of BL21 Star (DE3) Δ*fabR* 54
Figure 31. Microscopical analysis of strains 55
Figure 32. GGL production obtained when modulating the transcription factors 56
Figure 33. Lipid profile in the new engineered strains 58
Figure 34. PA synthesis pathway 61
Figure 35. Agarose gel of PlsC^{L} 62
Figure 36. Agarose gel of PlsB^{L} 63
Figure 37. Design of the region between PlsC and PlsB proteins 64
Figure 38. Agarose gel of PlsC^{L}·PlsB^{L} 64
Figure 39. Primer design to subclone plsB into pRSF-1b plasmid 65
Figure 40. Agarose gel of PlsB^{H} 66
Figure 41. Agarose gel of PlsC^{H}·PlsB^{H} 67
Figure 42. Production in BL21 Star (DE3) ΔtesA using MG517 and different acyltransferase 68
Figure 43. DTNB coupled assay 69
Figure 44. Folds increase of the enzymatic activity of each acyltransferase 70
Figure 45. Phospholipid synthesis in *E. coli* 73
Figure 46. *pgpB* and *cdh* involvement in the phospholipid biosynthesis metabolism 74
Figure 47. Design of the primers and linker zone to obtain PlsCxPgpB fusion protein 75
Figure 48. Agarose gel of PlsCxPgpB^{H} 76
Figure 49. CDP-diacylglycerol synthesis reaction 76
Figure 50. Effect of using CDH in BL21 Star (DE3) ΔtesA strain 77

Figure 51. Assay principle to quantify by Malachite green PgpB enzymatic activity 78
Figure 52. Enzymatic activity of PlsCxPgpB protein 78
Figure 53. GGL production using CDH and fusion protein PlsCxpgpB 79
Figure 54. Biosynthesis of UDP-glucose in *Escherichia coli* 83
Figure 55. Primer design to obtain larger homologous sequences 84
Figure 56. Agarose gel to confirm the removal of *ushA* from the genome 85
Figure 57. Agarose gel to confirm the presence of the resistance cassette 85
Figure 58. Agarose gel to confirm the removal of the resistance cassette 86
Figure 59. Production of Δ*tesA* with or without the expression of GalU enzyme 87
Figure 60. GGL production using Δ*tesA* and Δ*tesA* Δ*ushA* engineered strains 88
Figure 61. Microscopic analysis of the different engineered strains 88
Figure 62. Production in Δ*tesA* Δ*ushA* overexpressing different acyltransferases 89
Figure 63. GGL production in Δ*tesA* and Δ*tesA* Δ*ushA* strains overxpressing *fadR* 90
Figure 64. Fatty acid profile in the engineered strains and using FadR in Δ*tesA* Δ*ushA* 91
Figure 65. Effect of using CDH 92
Figure 66. Production using PlsCxPgpB protein in different strains 94
Figure 67. Enzymatic activity of PgpB enzyme in the different strains 94
Figure 68. Agarose gels of Mg517xmCherry 100
Figure 69. Relation between production and enzymatic activity of MG517 101
Figure 70. MG517 quantification by using mCherry as a reporter gene 101
Figure 71. MG517 enzymatic activity when expressing PlsC or not 103
Figure 72. One-dimension TLC for radioactivity measure 104
Figure 73. Membrane composition of the different engineered strains 105
Figure 74. Metabolic pathway for the synthesis of glycoglycerolipids in *Escherichia coli* 109
Figure 75. Phospholipid biosynthesis in *Escherichia coli* 115
Figure 76. Information about *psd* 117
Figure 77. Design of the promoter and RBS library for *psd* and its screening system 118
Figure 78. PCR colony from the colonies obtained by GGA 120
Figure 79. Scheme to obtain the linear DNA to perform the KI 120
Figure 80. Summary of the procedure to obtain the library 121
Figure 81. Monitorization of the growth and fluorescence overtime of 264 colonies 122
Figure 82. Fluorescence/growth vs maximal growth rate 122
Figure 83. Strain characterization by fluorescence determination and growth rate 123
Figure 84. Sequencing of BL21 Star (DE3) Δ*fadE* 167
Figure 85. BL21 Star (DE3) Δ*tesA* sequencing 168
Figure 86. FadE region sequencing of the doble knockout BL21 Star (DE3) Δ*tesA* Δ*fadE* 168
Figure 87. FabR sequence of BL21 Star (DE3) Δ*tesA* Δ*fabR* 169
Figure 88. UshA removal sequence of BL21 Star (DE3) Δ*tesA* Δ*ushA* 169
Figure 89. Characterization of MG517-mCherry fusion protein (mg517xmCherry) 183
Figure 90. Gas chromatogram of the methyl ester fatty acids 184
Figure 91. Myristic acid mass spectrum (C14:0) 185
Figure 92. Palmitic acid mass spectrum (C16:0) 185
Figure 93. Palmitoleic acid mass spectrum (C16:1) 186
Figure 94. 2 – hexyl – cyclopropaneoctanoic acid mass spectrum (C17:0Δ) 186
Figure 95. Stearic acid mass spectrum (C18:0) 187
Figure 96. Oleic acid mass spectrum (C18:1) 187
Figure 97. 2 – octyl – cycopropaneoctanoic acid (C19:0Δ) 188
Figure 98. Subproduct 1 mass spectrum identified as benzenepropanoic acid 188
Figure 99. Subproduct 2 mass spectrum identified as oleamide 189

List of tables

Table 1. Metabolic strategies used to increase the production of free fatty acids bacteria 18
Table 2. Data of the assay to correlate biomass with glucose 36
Table 3. Growth rate in the engineered strains 46
Table 4. Production of GGL in the engineered strains 48
Table 5. Primers and template used to obtain *fadR* plasmid 55
Table 6. Growth rate in the engineered strains 55
Table 7. Production of GGL modifying genetic regulators 56
Table 8. Percentage abundance of the fatty acid methyl esters 57
Table 9. Primers and template used to obtain the fragments for plsC^{L} 62
Table 10. Primers and template used to obtain the fragments for plsB^{L} 63
Table 11. Summary of the required templates and primers to obtain the desired fragments .. 64
Table 12. Primers and templates used to obtain plsB^{H} 65
Table 13. Primers and templates used to obtain plsC^{H}-plsB^{H} 66
Table 14. Production of GGL of engineered Δ*tesA* strain 67
Table 15. Acyltransferase activity of the Δ*tesA* strains 69
Table 16. Primers and templates used to obtain mg517·cdh plasmid 74
Table 17. Primers and templates used to obtain the desired fragments for plsCxpgpB^{H} 75
Table 18. Production of GGL produced by ΔtesA strain when expressing mg517·cdh-plsC^{H} and mg517-plsCxpgpB^{H} 78
Table 19. Summary of the production comparing Δ*tesA* and Δ*tesA* Δ*ushA* strains 88
Table 20. Glycoglycerolipid production using different acyltransferases in Δ*tesA* Δ*ushA* 89
Table 21. Production in the different strains using *fadR* 90
Table 22. Fatty acid content in the different engineered strains 91
Table 23. Production in the different engineered strains overexpressing CDH 92
Table 24. GGL production in the different strains when plsCxpgpB protein was used 93
Table 25. Primers and template used to obtain mg517xmCherry plasmid 99
Table 26. Glycosyltransferase activity of the different strains coexpressing mg517-plsC^{H} 100
Table 27. Specific activity using MG517-mCherry fusion protein 102
Table 28. Enzymatic activity in BL21 Star (DE3) when only using MG517 or MG517 and PlsC 103
Table 29. Conditions and equipment used to reveal the radiolabeled TLC 105
Table 30. Primers designed to subclone psd into KI plasmid by GGA. 119
Table 31. Mix composition for the Gibson Assembly reaction mixture 132
Table 32. 5x ISO buffer preparation 132
Table 33. Gibson Assembly reaction mix, volumes 133
Table 34. Golden Gate Assembly reaction mix 133
Table 35. PCR conditions for GGA 133
Table 36. Conditions for the doble digestion using NcoI-HF and EcoRI-HF enzymes 136
Table 37. Stock for the minimal medium preparation 139
Table 38. Volumes used for the different cultures 140
Table 39. Antibiotic stocks 140
Table 40. Volumes used to extract the lipid fraction in the first extraction 142
Table 41. Volumes used for the second extraction 142
Table 42. Glucose standard curve 142
Table 43. Conditions for the Gas Chromatography analysis 146
Table 44. Conditions for the mass analysis 147
Table 45. Stocks used for the minimal mèdium supplemented with ^{14}C-acetate 147
Table 46. BCA standard curve 149
Table 47. Strain genotypes 170
Table 48. Overall glycoglycerolipid production of the engineered strains 171
Table 49. Properties of mCherry protein 183

Table 50. Identified fatty acids 184

Publications and communications

Article

Orive-Milla N, Delmulle T, de Mey M, Faijes M, Planas A. Metabolic engineering for glycoglycerolipids production in E. coli: Tuning phosphatidic acid and UDP-glucose pathways *Metab Eng*. 2020;61:106-119. doi:10.1016/j.ymben.2020.05.010

Poster

Tuning a metabolically engineered *E. coli* strain for glycolipids production. 12th Carbohydrate Bioengineering Meeting (23- 26 April 2017)

Tuning a metabolically engineered *E. coli* strain for glycolipids production. 19th Eurocarb Symposium (6th of July 2017)

Tuning a metabolically engineered *E. coli* strain for glycolipids production. Metabolic Engineering 12 (June 2018)

Tuning a metabolically engineered *E. coli* strain for glycolipids production. Applied Synthetic Biology (24-26 October 2018)

Design of a new biological platform for production of high added value products: glycolipids. I Jornada de doctorands IQS (16-17 May 2019)

Oral presentation

Design of a new biological platform for production of high added value products: glycolipids. I jornada de doctorands IQS (16th May 2019)

Acknowledgements

Per fi ha arribat la part que tenia més ganes d'escriure, els agraïments! És molt probable que em deixi a algú ja que durant aquests anys de tesi hi ha hagut molta gent que m'ha ajudat i amb la que he tingut la sort de conviure experiències tant al laboratori com a fora d'ell. Demano perdó ja per avançat i si veieu que no us he posat, m'ho dieu i ja veurem si us invito a una cervesa després de demanar-vos perdó mil vegades!

Sense enrotllar-me més començo. En primer lloc m'agradaria donar les gràcies tant el Dr. Antoni Planas com a la Dra. Magda Fajes, ambdós directors de la tesi, per tota la seva contribució en el desenvolupament d'aquest. Per totes les hores dedicades i per donar-me sempre idees i estratègies per utilitzar i augmentar la producció. Donar les gràcies sobretot per ajudar-me a créixer professionalment ja sigui per deixar-me donar un tastet al món de la investigació o per dur-me a congressos on veure que s'estava fent actualment en el nostre camp. A tu Toni vull agrair-te sobretot l'oportunitat d'haver-me deixat formar part del Laboratori de Bioquímica, on he tingut l'oportunitat de conèixer i treballar al costat d'un gran equip. A la Magda cal agrair-li especialment la paciència que ha tingut sempre amb mi, sense arribar mai a desesperar-se i per les quedades per corregir la tesi als llocs més insospitats (restaurants, vienes, etc.) que em van ajudar a donar-li forma (tot i que cada cop que sortia d'una correcció tenia ganes de plorar de la quantitat de coses que havia de canviar).

I would like to thank Prof. Marjan de Mey, for allowing me to stay, not once but twice (!!) at the MEMO laboratory. I had the most amazing time there and thanks to you I had the opportunity to learn new techniques that I never used before. Thanks not only for being a super tutor but also for showing me the Gentsian lifestyle (i.e., cheese and wine event). I would also like to thank Chiara and Lien for being such incredible people, you looked shy at the beginning but then, you were so funny! It has been a pleasure sharing the bench with you Chiara (and your Hollywood look alike Student) and laugh everytime something did not work out (which happened quite a lot to be honest). Lien thanks for being so friendly and for being also a pet person! (I'm glad I wasn't the only one). Also thank you for being, I think, one of the few who didn't mess with me. Is actually nice to have someone who doesn't need to do it! So, thank you!! I would like to thank David for teaching me his idea of the library and helping me into developing mine, without your help it would not have been possible! I would also like to thank Tom for being so nice to me. I am sorry that you always have to supervise me! But you have to admit that I improved quite a lot from the first time that I was in Ghent. Thanks for helping me, at the beginning of the PhD, in designing the knockouts and then never losing hope in that something will work (or at least I hope you didn't lose it!). Now we have a paper together. See? Your effort was rewarded! I also want to acknowledge Andries for doing all the characterization assays of the library. To Marteen, thanks for being so funny when you were so stressed about your PhD and how the time to finish was coming. To Brecht I would like to thank him for being always so nice and for not complaining (a lot) when he came to Barcelona and we ate at crazy times, I know that it had to be very hard

for you considering that you were always going for lunch at 12.0h (if not earlier!). I think that I can say this in the name of all the biochemistry lab, that we had so much fun having you in Barcelona! I have to thank Dries as well for always being so supportive (to Real Madrid specially) and for always messing with me. It felt like being home honestly! I would also like to thank Jo for always being so kind and helpful to me. I didn't know you very well but I heard amazing things about you, so I wish I had the opportunity. Finally, last but not the least, I would like to thank Wouter for being such a crazy person. I think that since the moment I saw that we had to share bench I knew that crazy (funny) times were coming. Thanks for always helping me in all the technical problems I had and for helping Carlos and I in designing primers to achieve the knockouts. Without you it would have take longer and we didn't have enough time. Also thank you for letting me steal your DMSO, water, eppis and so... a sorry, you didn't know right? Well, that was my revenge for always messing with me!! Also, thank you for being the most amazing Christmas donkey, you know, if you want to switch careers in the future, you really have options there! Overall, thanks to each one of you for always making it so easy for me to be there. It's always difficult to be an outsider and you made it crazily easy! Dank u wel!

Respecte al Laboratori de Bioquímica m'agradaria donar les gràcies a tots els sèniors que m'han ajudat en cada dubte que he tingut i m'han donat idees en cada un dels seminaris que he fet. Al Xevi per ajudar-me sobretot amb les coses del programa de doctorat, és una sort tenir-te allà per saber com funciona la burocràcia! Al Pau Leivar, moltes gràcies per totes les suggerències realitzades als seminaris que sempre han sigut d'utilitat, sobretot perquè al treballar també en glicolípids ajuda bastant a veure que no havíem pensat i que podria ser útil! Al Marc Carnicer per ajudar-me tant amb els càlculs estadístics com amb les simulacions de fluxos metabòlics (i per les magdalenes tan bones que portes quan és el teu aniversari). A tu Teresa, mira que era difícil omplir el forat que va deixar l'Estela quan va acabar com a companya de taula però realment has estat molt bona compi! Moltes gràcies per tots els cafès que ens has dut per fer qualsevol cosa random, per fer-nos sentir sempre valorats, per posar-nos banda sonora quan practiques cant i, sobretot, per pujar-me sempre l'autoestima arreglant-te problemes tecnològics. Sempre que penso que no en tinc ni idea de informàtica recordo aquells moments i dic "no estamos tan mal". També cal donar les gràcies a la Patricia per ajudar-nos a desconnectar durant els esmorzars i dinars, per deixar-nos agafar la teva cadira, sempre amb la precaució de tornar-te el coixí i per saber tots els cotilleos de la universitat, que com saps, ens donen la vida a tots!

Respecte a la família biocher que dir? Aquesta part serà molt dura d'escriure eh? Ho faré per ordre cronològic. A l'Almudena, moltes gràcies per les bogeries que fas i dius. El teu riure contagiós és algo que hem trobat molt a faltar des de que te'n vas anar i sobretot les teves expressions úniques. També agrair-te l'anàlisi setmanal de The walking dead durant els cafès que em va ajudar a no perdre'm ni un detall de la sèrie. A l'Estela que dir de tu xurri que no sàpigues ja? Moltes gràcies per ser com ets, per ser tan especial i per sempre veure el millor de

cada situació. Per ser la meva persona preferida per anar als concerts del James Morrison i per sempre ajudar en tot allò que podies.
A tu Cristina només agrair-te que siguis sempre així com ets. Per ser una de les millors persones que he conegut i sobretot per sempre tenir paraules amables per tothom. El teu nivell de sacrifici no té límits i sempre ho fas tot des de la bondat. Moltes gràcies Sergi per riure't sempre de mi! No hi ha res com arribar a les 8 del matí i trobar-te a un Sergi a punt de fotre's un tret preguntant-te "Creus que la vida és una puta merda? Estàs molt enfonsada?" realment sense tu el Laboratori no hagués estat el mateix. Gràcies per tots els riures que aportes i per tots els moments de "nenis" que hem tingut. Sobretot per recordar-me cada cop que pots com et vas sacrificar per venir a veure'm a Bèlgica, recordant-me tot seguit la fresqueta que feia.
A tu Mireia no sé ni per on començar, hem viscut tantes coses, moltes de surrealistes que realment no sé que agrair-te. Suposo que t'ho podria resumir tot en un gràcies per ser com ets, per sempre donar un toc d'humor a cada desgràcia que passa, a ser una companya excel·lent de poiata i per robar absolutament qualsevol material que necessitis i que vegis que està en una altra poiata (aquest gràcies era més aviat irònic). Gràcies per ser tan increïblement dramàtica i que d'aquesta manera ningú noti que jo tinc tendència a fer el mateix! Gràcies per sempre ser positiva i per fer-nos riure a tots amb les teves desgràcies, per compartir tants riures i per compartir tants concerts. Òbviament, gràcies també per donar-nos sempre versions alternatives a totes les cançons que cantes i de les quals t'inventes totes les lletres. Al Marc, Bernat i Álvaro, tot i havier coincidit menys temps al laboratori donar-vos les gràcies per l'alegria que sempre heu aportat. Per tota la feina i bon rotllisme que aporteu al dia a dia. Gràcies per ser tan bons companys i per sempre estar disposats en ajudar en tot allò que podeu!
Finalment, a tu Laia no sé ni que dir que no sàpigues ja però ho intentarem igualment! Moltes gràcies per ser com ets, per haver estat cada dia d'aquest llarg (casi que interminable) camí. Per estar al meu costat en cada etapa que obro i tanco i sobretot per la bondat amb la que ho fas tot. Moltes gràcies per totes les converses, per tots els riures, els viatges i sobretot per deixar-me formar part de la teva vida! Si amb algú havia de viure aquesta experiència, no podria haver demanat a ningú millor! Moltes gràcies per estar sempre al meu costat.
Moltes gràcies a tots els Biochers per haver-me adoptat i deixat formar part d'un grup tan maco. També gràcies per haver-me deixat aprendre coses de cada un de vosaltres que m'han ajudat a millorar (espero) com a persona i com a professional. Perquè fer un doctorat al vostre costat ha estat increïble. No podria haver desitjat uns millors companys! Gràcies, gràcies i gràcies!

Aquest treball no hagués sigut possible sense l'ajuda de diverses persones, expertes en diferents tècniques analítiques. Totes elles m'han ajudat, a mi que sóc una persona del món bio totalment, a augmentar els meus coneixements en anàlisis. Així doncs, per començar m'agradaria donar les gràcies a la Dra. Cristina Ribas del departament de cromatografia de IQS per haver-me permès no només analitzar els àcids grassos provinents de les soques sinó també per totes les hores que va dedicar a explicar-me com funcionava l'equip, com s'analitzaven els

cromatogrames i, en el cas del masses, en ajudar-me a identificar cada un dels compostos. Gràcies per tota l'ajuda Cristina perquè sense tu no sé si ho hauria aconseguit!
També vull agrair a la Dra. Gemma Arsequell i a l'Ellen del CSIC tota la seva ajuda per tal d'intentar analitzar per cromatografia els glicolípids. Va ser molt maco treballar amb vosaltres i és una experiència que recordaré sempre. Sobretot moltes gràcies Ellen per totes les hores que em vas dedicar que, especialment sent cap al final de la teva tesi, van suposar un gran sacrifici.
Per últim, m'agradaria agrair al Dr. Josep Antoni Pérez de la UAB per ajudar-me amb l'anàlisi de les capes fines per radioactivitat. Tot i que posar a punt les condicions del mètode va ser una mica complicat sense la teva ajuda no hagués estat possible.

Respecte a tots els estudiants que han format part del projecte al llarg d'aquests anys. En primer lloc m'agradaria agrair a la Mireia, Marta i Eulàlia tot el seu esforç i totes les seves hores dedicades. De cada una de vosaltres he après moltes coses mentre hem treballat juntes. Moltes gràcies per tot l'esforç que hi heu dedicat i per tota la feina que heu aportat. Us desitjo de tot cor el millor en les feines que esteu actualment desenvolupant. Jo que he tingut la sort de treballar amb vosaltres, sé que sou capaces d'aconseguir tot allò que vulgueu.
A tu Carlos t'haig d'agrair moltíssimes coses. No només a nivell professional sinó també a nivell personal. Moltes gràcies per les hores que has dedicat al projecte, a fer-lo teu i a aconseguir els boníssims resultats que vas aconseguir. Qui hauria dit al setembre, quan tot just començaves a fer el TFM, que series capaç de fer un TFM com el que vas fer? La teva evolució va ser extraordinària i realment estic molt contenta d'haver-te pogut ajudar. Personalment et vull agrair el tros de persona que ets. Crec que mai m'havia rigut tant treballant amb algú al laboratori. Mai vas perdre el somriure per res i sempre estaves disposat a tot per tal de que els KO funcionessin. També dir-te que moltes gràcies pels cines, billars (on sempre perdies òbviament), cerveses i especialment per acabar-te sempre els meus plats!! Va ser una sort anar a Ghent amb algú com tu ja que no només vaig tenir un minion per a que m'ajudés amb la part experimental sinó que també vaig tenir la sort de conèixer una grandíssima persona!

Pel que fa a la família que elegim, els amics (de fora del laboratori), m'agradaria començar donant les gràcies a la Laura per ser sempre al meu costat. Perquè sempre has estat un referent a la meva vida i perquè una de les millors coses que m'ha passat mai va ser anar a aquella famosa festa major ara fa ja més de 12 anys. Perquè hem viscut moltes etapes juntes i perquè ens queden moltes més per passar! Gràcies per la serenor que sempre m'aportes i per ajudar-me sempre a filtrar! M'agradaria també agrair a la Cristina el seu suport durant tots aquests anys. Per tots els sopars, dinars, esmorzars i mil coses més que hem viscut. Perquè els teus comentaris sempre m'ajuden a posar-ho tot en perspectiva i perquè sempre m'has ajudat en tot el que he necessitat.
Vull agrair també al Joan i al Dani la seva constant presència. Per tots els anys que fa que ens coneixem i pel milió de coses que hem viscut. Per tots els riures que ens hem fet, pels viatges, per les marujeades i sobretot per ser sempre tan únics!

A tu Mercè et vull agrair directament que siguis com ets. Crec que no ets conscient del tros de persona que ets i lo afortunats que som tots de tenir-te ben a prop. Moltes gràcies per ser sempre una roca a la vida de la gent que t'envolta i per sempre estar disposada a ajudar a tot allò del que siguis possible per als demés. Perquè sense tu, Sant Llorenç no seria el mateix i probablement m'hauria estressat molt més amb la tesi del que m'he estressat.

A la Cristina (kita), Mabel i Laia R. us vull donar les gràcies per haver sigut gent tant important i per sempre preguntar-me per la tesi. Tot i que amb el temps ens hem anat veient menys, cada cop que ens veiem és com si no hagués passat ni un dia. Moltes gràcies pel vostre suport!

A tu Nuri que dir-te? No et vull dir coses massa maques perquè després t'ho creuràs i m'ho aniràs retraient sempre però vaja... Donar-te les gràcies per estar sempre aquí. Qui havia de dir que quart d'ESO seria un curs tan genial al teu costat? Ha plogut molt des de llavors però l'única cosa que no ha canviat és que hem continuat juntes, veient-nos, tu recomanant-me series i llibres i sempre explicant-me els marujeos de tot el poble. Gràcies també per la correcció de l'anglès que vas fer!

Agrair a la Fani, la Marta, la Mire i la Sílvia per sempre mostrar interès per la tesi i, especialment la Fani, per donar-me suport quan passava moments molt negres.

Gràcies Mire per sempre donar-me perspectiva des de la distància, per sempre preocupar-te, preguntar-me i per formar sempre part de la meva vida. Tot i que m'hagis abandonat per anar-te'n a València sempre tinc l'esperança de que tornaràs a prop i em deixaràs que mimi molt al meu Brunete!

A tu Pau, gràcies per haver format part de la meva vida. Tot i que em va posar molt trista que marxessis a Estrasburg a fer el teu doctorat, realment mai hem perdut el contacte. Gràcies per tots els cafès que hem arribat a fer junts, per sempre recordar-te de mi, sempre preocupar-te per la tesi i per sempre saber com treure'm un somriure. Haig de confessar que el cactus que em vas regalar no va durar pas massa però crec, i tots sabem, que la culpa no va ser meva.

A tu Jose gràcies per recolzar-me sempre amb el doctorat i donar-me sempre una opinió diferent a la que esperaria. Sempre em fas replantejar coses!

Por último, obviamente, no podia olvidarme de mi siame favorita. Muchas gracias Si Wen por todo. A pesar de la distancia, sigues siendo una persona importantíssima en mi vida. Muchas gracias por estar siempre animándome cuando tenia bajones y por indignarte conmigo frente a las injustícias. Si algo de bueno va a tener la defensa de la tesis es que me permitirá verte de nuevo (o eso espero)!

Per últim, vull agrair a la meva família el seu suport incondicional. Al meu germà Roger per sempre estar a peu de guerra i donar-me sempre consells. Al meu tiet per ser sempre un constant a les nostres vides i sempre ajudar-nos amb tot el que podia. A la meva mare òbviament vull agrair-li el seu suport, la seva comprensió i la seva paciència. Sempre m'has ajudat amb tot i sempre has renunciat a tot per nosaltres. Mai has demanat res a canvi i sempre has sigut un gran referent a les nostres vides. Moltes gràcies pel teu sacrifici i per haver-nos ajudat a convertir-nos en les persones que som avui en dia. Finalment, tot i que ja no hi és, agrair al meu

pare tot el que em va ensenyar a la vida. Totes les lliçons que em va donar, els viatges que vam fer i l'ajuda que em va brindar. Gran part de la persona que sóc avui en dia és gràcies a tu i vull pensar que et sentiries orgullós de la persona en la que m'he convertit.

Agrair també al tribunal de la tesi per aportar el seu coneixement i estar disposats a valorar el present projecte. També vull agrair a qui s'estigui llegint aquesta tesi per tal de que la feina que hem desenvolupat al llarg d'aquests anys sigui útil per a noves vies d'investigació.

Per últim, gràcies a IQS School of Engineering per l'ajut predoctoral així com a la SEBAP per la seva ajuda de mobilitat.

Moltes gràcies a tots!

Summary

Glycolipids are products of high-added value due to their amphipathic properties, which endow them with a broad range of applications in the chemical (i.e., biosurfactants) and biomedical sectors (i.e., vaccine adjuvants). Depending on their lipidic moiety, glycolipids are classified in different families. If the lipid moiety is a ceramide or diacylglycerol, the glycolipids are known as glycosphingolipids or glycoglycerolipids respectively. While glycosphingolipids have shown to play essential roles in many biological processes, glycoglycerolipids (GGL) are interesting compounds due to their potential use as vaccine adjuvants or tumor suppressors.

Although the interest of these compounds is very high, their applications are hampered by their low availability and high production costs. Chemical synthesis requires complex protection and deprotection steps to achieve the desired regio- and stereospecificity of the glycosidic linkage, which consequently lower the yield and efficiency of the process. Therefore, we considered metabolic engineering as a potential strategy for the production of glycolipids and we aimed at building up a metabolic engineering platform in *E. coli* to achieve these complex structures of interest. In previous studies, our group reported that the glycolipid synthase MG517 from *Mycoplasma genitalium* was functional and glycoglycerolipids were obtained from UDP-glucose (UDP-Glc) and diacylglycerol (DAG). In addition, the first generation of engineered strains demonstrated that the availability of the DAG was the key bottleneck in GGL production (Mora-Buyé et al., 2012).

In the present project, five different metabolic strategies were proposed to increase the production of GGL using *E. coli*. The first four strategies were aimed at increasing the available pool of the lipidic precursor, DAG. Thus, the first strategy was based on increasing DAG availability by removing competing reactions. To achieve so, different genes involved in the ß-oxidation and activation of fatty acids were knocked out (Δ*tesA* y Δ*fadE)* reporting an almost 2-fold production increase. The second strategy was based on increasing fatty acid availability by modulating different transcriptional factors (*fabR* y *fadR*). Although this strategy did not report an improvement of GGL yield, it showed a change in the fatty acid profile with an increase of unsaturated fatty acids. The third strategy was based on increasing the conversion of acyl donors to phosphatidic acid, precursor of DAG, by overexpressing PlsC and PlsB acyltransferases. The fourth strategy was based on increasing diacylglycerol availability by overexpressing the fusion PlsCxPgpB protein that could redirect the flux to DAG or CDH promoting the hydrolysis of phospholipids. Among the different engineered strains, the Δ*tes*A strain co-expressing MG517 and a fusion PlsCxPgpB protein was the best producer, with a 350% increase of GGL titer compared to the parental strain expressing MG517 alone. Interestingly, the strains co-expressing CDH showed a shift in the GGL profile towards the diglucosylated lipid (up to 80% of total GGLs).

Finally, a metabolic strategy was proposed to increase the availability of the other precursor, UDP-Glc. This fifth strategy was based on overexpressing GalU enzyme, which is responsible for the biosynthesis of UDP-Glc, and by removing the UDP-sugar diphosphatase encoding gene *ush*A. However, none of these modifications further improved the GGL titers.

Finally, as it was also reported by our group that phosphatidylethanolamine was exchangeable in the membranes of *E. coli* by the new GGL compounds, a library of promoters and RBS was designed to decrease the production of this phospholipid trying at the same time to increase the production of glycolipids.

Resum

Els glicolípids són producte d'alt valor degut a les seves propietats amfipàtiques que els doten d'un ampli rang d'aplicacions en els sectors químic (ex., biosurfactants) o biomèdic (ex., adjuvant de vacunes). Depenent de la unitat lipídica que els forma aquests compostos poden ser classificats en diferents famílies. Si la unitat lpídica és una ceramida o diacilglicerol, el glicolípid resultant es coneixerà com a glicoesfingolípid o glicoglicerolípid (GGL) respectivament. Mentre que els glicoesfingolípids han demostrat jugar un paper clau en diversos processos biològics, els glicoglicerolípids són interessants degut al seu ús potencial per a ser usats com a adjuvants de vacunes o supressors tumorals.

Tot i que l'interès per aquests compostos és alt, la seva aplicació es veu obstaculitzada per la seva baixa disponibilitat i alt cost de producció. La síntesi química requereix de complexes passos de protecció i desprotecció per tal d'aconseguir la desitjada regio- i estereoespecificitat de l'enllaç glicosídic que, conseqüentment, comporta una reducció del rendiment i eficiència del procés. Per això, vam considerar l'enginyeria metabòlica com a estratègia potencial per a la producció de glicolípids i ens vam centrar en l'obtenció d'una plataforma d'enginyeria metabòlica en *E. coli* per tal d'obtenir aquests complexes productes d'interès. En estudis previs, el nostre grup va reportar que la sintasa de glicolípids MG517 de *Mycoplasma genitalium* era funcional i que s'obtenien glicoglicerolípids a partir de UDP-glucosa (UDP-Glc) i diacilglicerol (DAG). Addicionalment, una primera generació de soques modificades va demostrar que la disponibilitat de DAG era limitant per a la producció de GGL (Mora-Buyé et al., 2012).

En el present projecte, cinc estratègies diferents d'enginyeria metabòlica van ser proposades per tal d'augmentar la producció de GGL utilitzant *E. coli*. Les primeres quatre estratègies tenien com a objectiu incrementar el *pool* del precursor lipídic, DAG. Sent així, la primera estratègia es basà en augmentar la disponibilitat de DAG a través de l'eliminació de reaccions competitives. Per aconseguir-ho, es van knockejar diferents gens involucrats en la ß-oxidació i l'activació d'àcids grassos (Δ*tesA* i Δ*fadE*) reportant un increment en la producció de casi el doble. La segona estratègia es va basar en incrementar la disponibilitat d'àcids grassos mitjançant la modulació de factors de transcripció (*fabR* i *fadR*). Aquesta estratègia no va reportar un increment de la producció però si un canvi en el perfil lipídic amb un increment d'àcids grassos insaturats. La tercera estratègia es basava en incrementar la conversió dels donadors d'acils a àcid fosfatídic, precursor del DAG, sobreexpressant les aciltransferases PlsC i PlsB. La quarta estratègia es centrà en augmentar la disponibilitat del diacilglicerol per la sobreexpressió de la proteïna de fusió PlsCxPgpB, capaç de redirigir el flux cap a DAG, o CDH promovent la hidròlisi de fosfolípids. D'entre les diferent soques modificades, Δ*tesA* co-expressant MG517 i la proteïna de fusió PlsCxPgpB va ser la soca més productora, amb un 350% d'increment en la producció de GGL comparant-la amb la soca parental expressant únicament MG517. Especialment interessant és que les soques coexpressant CDH van presentar un canvi en el perfil de GGL cap al lípid diglucosilat (representant al voltant d'un 80% del total de GGLs).

Finalment, es va proposar una estratègia metabòlica per incrementar la disponibilitat de l'altre precursor, UDP-Glc. Aquesta cinquena estratègia es va basar en sobre expressar l'enzim GalU, responsable de la biosíntesi d'UDP-Glc, i eliminant l'enzim codificant per la UDP-sucre difosfatasa *ushA*. No obstant, cap d'aquestes modificacions va aconseguir millorar els nivells de GGLs.

Per últim, tal i com va ser reportat pel nostre grup que la fosfatidiletanolamina era intercanviable en les membranes d'*E. coli* pels nous producte GGL, una llibreria de promotors i RBS va ser dissenyada per tal de disminuir la producció d'aquest fosfolípid, augmentant al mateix temps de la producció de glicolípids.

Resúmen

Los glicolípidos son productos de alto valor debido a sus propiedades amfipáticas, que los dota en un amplio rango de aplicaciones en los sectores químicos (ej., biosurfactantes) o biomédicos (ej., adyuvante de vacunas). Dependiendo de la unidad lipídica que los forma, los glicolípidos son clasificados en diferentes familias. Si la unidad lipídica es una ceramida o diacilglicerol, los glicolípidos son conocidos como glicoesfingolípidos o glicoglicerolípidos (GGL) respectivamente. Mientras que los glicoesfingolípidos han demostrado jugar papeles clave en diversos procesos biológicos, los glicoglicerolípidos son compuestos interesantes debido a su potencial uso como adyuvantes de vacunas o supresores tumorales.

Aunque el interés por estos compuestos es muy alto, su aplicación se ve obstaculizada por su baja disponibilidad y altos costes de producción. La síntesis química requiere de complejos pasos de protección y desprotección para conseguir la deseada regio- y estereoespecificidad del enlace glicosídico, que conlleva a una reducción del rendimiento y eficiencia del proceso. Por ello, consideramos la ingeniería metabólica como estrategia potencial para la producción de glicolípidos y nos centramos en construir una plataforma de ingeniería metabólica en *E. coli* para conseguir estas complejas estructuras de interés. En previos estudios, nuestro grupo reportó que la sintasa de glicolípidos MG517 de *Mycoplasma genitalium* era funcional y que glicoglicerolípidos podían ser obtenidos a partir de UDP-glucosa (UDP-Glc) y diacilglicerol (DAG). Adicionalmente, la primera generación de cepas modificadas demostró que la disponibilidad de DAG era limitante en la producción de GGL (Mora-Buyé et al., 2012).

En el presente proyecto, cinco estrategias diferentes de ingeniería metabólica fueron propuestas para aumentar la producción de GGL en *E. Coli*. Las primeras cuatro estrategias se centraron en aumentar el *pool* del precursor lipídico, DAG. Para ello, la primera estrategia se basó en incrementar la disponibilidad de DAG a través de la eliminación de reacciones competitivas. Para lograrlo, se knockearon diferentes genes relacionados con la ß-oxidación y la activación de ácidos grasos (Δ*tesA* y Δ*fadE*) reportando un incremento de casi el doble. La segunda estrategia se basó en incrementar la disponibilidad de ácidos grasos mediante la modulación de factores de transcripción (*fabR* y *fadR*). Aunque estrategia no reportó una mejora en el rendimiento de GGL, sí mostró un cambio en el perfil de los ácidos grasos con un incremento de los ácidos grasos insaturados. La tercera estrategia se basó en incrementar la conversión de los donadores de acilos a ácido fosfatídico, precursor del DAG, mediante la sobreexpresión de las aciltransferasas PlsC y PlsB. La cuarta estrategia se centró en aumentar la disponibilidad de diacilglicerol mediante la sobreexpresión de la proteína de fusión PlsCxPgpB capaz de redirigir el flujo de DAG, o CDH promoviendo la hidrólisis de fosfolípidos. Entre las diferentes cepas modificadas, la cepa Δ*tesA* coexpresando MG517 y la proteína de fusión PlsCxPgpB fue la mayor productora, con un incremento de los niveles de GGL del 350%, comparándola con la cepa parental expresando únicamente MG517. Interesantemente, las cepas coexpresando CDH mostraron un cambio en el perfil de GGL hacia el lípido diglucosilado (hasta el 80% del total del GGLs).

Finalmente, una estrategia metabólica fue propuesta para aumentar la disponibilidad del otro precursor, UDP-Glc. La quinta estrategia se basó en la sobreexpresión de la enzima GalU, responsable de la biosíntesis de UDP-Glc, y la eliminación de la UDP-azúcar difosfatasa codificada por el gen *ushA*. Sin embargo, ninguna de estas modificaciones mejoró los niveles de GGL.

Por último, tal y como reportó nuestro grupo que la fosfatidiletanolamina era intercambiable en las membranas de *E. coli* por los nuevos compuestos GGL, una librería de promotores y RBS fue diseñada para disminuir la producción de este fosfolípido intentando al mismo tiempo aumentar la producción de glicolípidos.

List of abbreviations

cdh	CDP-diacylglycerol diphosphatase
Δ*fadE*	*E. coli* BL21 Star (DE3) Δ*fadE*
Δ*fadE* Δ*tesA*	*E. coli* BL21 Star (DE3) Δ*fadE* Δ*tesA*
Δ*tesA*	*E. coli* BL21 Star (DE3) Δ*tesA*
Δ*tesA* Δ*fabR*	*E. coli* BL21 Star (DE3) Δ*tesA* Δ*fabR*
Δ*tesA* Δ*ushA*	*E. coli* BL21 Star (DE3) Δ*tesA* Δ*ushA*
1D	1 dimension
2D	2 dimension
ACN	Acetonitrile
AP	Alkaline phosphatase
BL21 Star (DE3)	*Escherichia coli* B
CDP-DAG	CDP-diacylglycerol
Cer	Ceramide
CerNBD	Ceramide-NBD
CL	Cardiolipin
CMP	Cytidine monophosphate
DAG	Diacylclycerol
DGalCer	Digalactosylceramide
DGalCerNBD	Digalactosylceramide
DGalDAG	Digalactosyldiacylglycerol
DGDAG	Diglucosyldiacylglycerol
DGlcDAG	Diglucosyldiacylglycerol
E. Coli	*Escherichia coli*
FA	Fatty acid
fabR	DNA-binding transcriptional repressor FabR
fadE	acyl-CoA dehydrogenase
fadR	DNA-binding transcriptional dual regulator FadR
FFA	Free fatty acids
galU	UTP- glucose-1-phosphate uridylyltransferase
GGL	Glycoglycerolipids
GPIs	Glycosylphosphatidylinositols
GSL	Glycosphingolipids
GT	Glycosyltransferase
H_2O	Water
LPA	Lysophosphatidic acid
MetOH	Methanol
mg517	diacylglycerol beta-glycosyltransferase
MGalCer	Monogalactosylceramide

MGalCerNBD	Monogalactosylceramide-NBD
MGalDAG	Monogalactosyldiacylglycerol
MGDAG	Monoglucosyldiacylglycerol
MGlcDAG	Monoglucosyldiacylglycerol
MQ	MilliQ water
PA	Phosphatidic acid
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
pgm	Phosphoglucomutase
PGP	Phosphatidylglycerol phosphate
pgpA	Phosphatidylglycerophosphatase A
pgpB	Phosphatidylglycerophosphatase B
pgpC	Phosphatidylglycerophosphatase C
PL	Phospholipids
plsB	glycerol-3-phosphate 1-*O*-acyltransferase
plsC	1-acylglycerol-3-phosphate *O*-acyltransferase
psd	Phosphatidylserine decarboxylase
Reference strain	BL21 Star (DE3) pET44b-mg517
tesA	acylCoA thioesterase I
TetraGDAG	Tetraglucosyldiacylglycerol
TGDAG	Triglucosyldiacylglycerol
TLC	Thin Layer Chromatography
UDP-Gal	UDP-Galactose
UDP-Glc	UDP-Glucose
ushA	UDP-sugar hydrolase
WT	BL21 Star (DE3)

List of strains

	Engineered strain name	Chassis (host strain)	Plasmids
#0	WT/ mg517	BL21 Star (DE3)	pET44b-*mg*517
#1	WT/ mg517-plsC^{H}	BL21 Star (DE3)	pET44b-*mg*517 + pRSF1b-*pls*C
#2	ΔtesA/ mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C
#3	ΔfadE/ mg517-plsC^{H}	BL21 Star (DE3) Δ*fad*E	pET44b-*mg*517 + pRSF1b-*pls*C
#4	ΔtesA ΔfadE/ mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fad*E	pET44b-*mg*517 + pRSF1b-*pls*C
#5	ΔtesA ΔfabR/ mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fab*R	pET44b-*mg*517 + pRSF1b-*pls*C
#6	ΔtesA/fadR-mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A	p5T7-*fad*R + pET44b-*mg*517 + pRSF1b-*pls*C
#7	ΔtesA ΔfabR/ fadR-mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fab*R	p5T7-*fad*R + pET44b-*mg*517 + pRSF1b-*pls*C
#8	ΔtesA/ mg517	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517
#9	ΔtesA/ mg517-plsC^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*pls*C
#10	ΔtesA/ mg517-plsB^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*pls*B
#11	ΔtesA/ mg517-plsC^{L}·plsB^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*plsC·pls*B
#12	ΔtesA/ mg517-plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*B
#13	ΔtesA/ mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*plsC·pls*B
#14	ΔtesA/ mg517·cdh-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517·cdh + pRSF1b-*plsC·pls*B
#15	ΔtesA/ mg517-plsCxpgpB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C*xpgp*B
#16	ΔtesA/ mg517-plsC^{H}-galU	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C + pCDF1b-*gal*U
#17	ΔtesA/mg517-plsCxpgpB^{H}-galU	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*plsCxpgp*B + pCDF1b-*gal*U
#18	ΔtesA ΔushA/ mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C
#19	ΔtesA ΔushA/ mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*plsC·pls*B
#20	ΔtesA ΔushA/ mg517-plsC^{H}-galU	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C + pCDF1b-*gal*U
#21	ΔtesA ΔushA/ fadR-mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	p5T7-*fad*R + pET44b-*mg*517+ pRSF1b-*plsC·pls*B
#22	ΔtesA ΔushA/ mg517·cdh - plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517·*cdh* + pRSF1b-*plsC·pls*B
#23	ΔtesA ΔushA/ mg517-plsCxpgpB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*plsCxpgp*B

INTRODUCTION

1. Introduction

1.1. Glycolipids

Glycolipids, first discovered by Ernst Klenk in 1942 after their isolation from brain tissue glycolipid (GL), is defined by the International Unit of Pure and Applied Chemistry (IUPAC) as "any compound containing one or more monosaccharide residues bound by glycosyl linkage to a hydrophobic moiety" (Chester, 1998).

These compounds can be found in the membranes of all organism kingdoms and their functions depend on the structure of the glycolipid. Between them, related to the membranes itself, it is of special interest its ability to stabilize the membranes and cell surface rigorization (Holst, 2008).

Due to its amphipathic structure they tend to cluster in different supramolecular structures depending on the environment surrounding them and their chemical properties (Rand and Luzzati, 1968). Some examples of these arrangements could be lamellar, micellar, non-lamellar or hexagonal among others (Faivre and Rosilio, 2010). These supramolecular structures are interesting since it is possible to use these products for different applications being examples of this drug delivery system.

As it happens with phospholipids, glycolipids are placed into the cell membranes but a major difference between both compounds is that in GL, hydrogen bonds have a strong influence in the transition temperature that confers them a structural integrity in the cell membranes (Holst, 2008).

Glycolipids can be divided into two different grades of complexity being those classified as simple glycolipids, sometimes called saccharolipids (i.e., rhamnolipids), and complex glycolipids (i.e., glycosphingolipids). Complex glycolipids are structurally more heterogeneous since they contain, in addition to the glycosyl and fatty acid, other groups such as glycerol (e.g., glycoglycerolipids) or acylated-sphingosine (e.g., glycosphingolipids) (Abdel-Mawgoud and Stephanopoulos, 2018).

Moreover, due to the high complexity of the glycosylation patterns among the complex glycolipids, its classification is based on their natural lipidic part being the following one the most prominent categories:

- Glycoglycerolipids (GGL): this compound is referred to any glycosylation pattern bound to diacylglycerol. Mono- and digalactosyldiacylglycerols are the main glycolipids found in chloroplasts membranes and other photosynthetic tissues (Hölzl and Dörmann, 2007; Kalisch et al., 2016). A classic example of GGL are galactolipids, found in photosynthetic tissues who play a role in their photosynthetic properties, and sulfolipids which contain sulphur functional groups and are associated with the sulphur cycle in plants. There are also glycoglycerolipids that are sulphated, known as sulfoquinovosyl diacylglycerols and which function is related to photosynthetic tissues and control of the ionic membrane's

permeability. Glycoglycerolipids based on fatty acids found in *E. coli* are the main compounds studied in this project.

- Glycosphingolipids (GSL): referred to glycolipids that contain at least one monosaccharide moiety linked to a ceramide or sphingoid (Wennekes et al., 2009). These compounds are essential structure components of the mammalian cell membranes. Besides their structural function in the lipidic bilayer they have also reached interest due to their biological activities. Probably the most known compound of this group are cerebrosides which were first discovered by Thudichum in 1884 when he was studying the chemical composition of the brain. These cerebrosides are abundantly found in the nerve cell membranes where they play an important role in cell signaling. Other subclassifications of these compounds are gangliosides, globosides, glycophosphosphingolipids and glycophosphatidylinositols.

Depending on the organism where the glycolipid is present, the function and the abundance is different:

- In prokaryotes, the main glycolipids found are derived of alkylglycerols, thylakoids and sulfoquinovasyl diacylglycerols. All of them are found in bacteria and also, it has been reported that these organisms contain digalactosyldiacylglycerol that is a known inhibitor of type I HIV reverse transcriptase (Zhang et al., 2014). A phosphocholine-containing glycoglycerolipid was isolated in *Mycoplasma fermentans* and a very unusual one was isolated from *Borrelia burgdoferi* (cholesteryl 6-o-acyl-ß-D-galactopyranoside) that have been reported to play a role in the host immune response against Lyme disease where it can be used as a diagnostic or a treatment agent. Regarding to other genre, in Mycobacteria there are three major types of glycolipids: lipooligosaccharides, glyceropeptidelipids, phenolic glycolipids and lipoarabinomannan, that is a lipoglycan very important of the cell wall.
- In plants, glycolipids can be grouped in four main groups: glucose and sucrose esters, stearyl glucosides, glycosphingolipids and glycosyldiacylglycerols. It seems that glycosphingolipids of plants may play a role in the communication between cells (Kalisch et al., 2016). On the other hand, glycoglycerolipids, specifically galactolipids, are key lipids in the membranes where they substitute phospholipids and play a role in the photosynthesis. They are also substrate for the production of oxylipins which are responsible for plants' response, especially, against stress.
- In animals, glycolipids have been reported to play an important role in cell membranes since they are responsible for key functions such as cell communication, cell differentiation or proliferation among others. It has also been reported that alterations in their composition and mutations in glycosyltransferases might play a role in oncogenesis.

  The most important glycolipids in animals are glycosphingolipids. Along with sphingomyelin and cholesterol, these glycolipids can be found in the microdomains in membrane-linked processes.

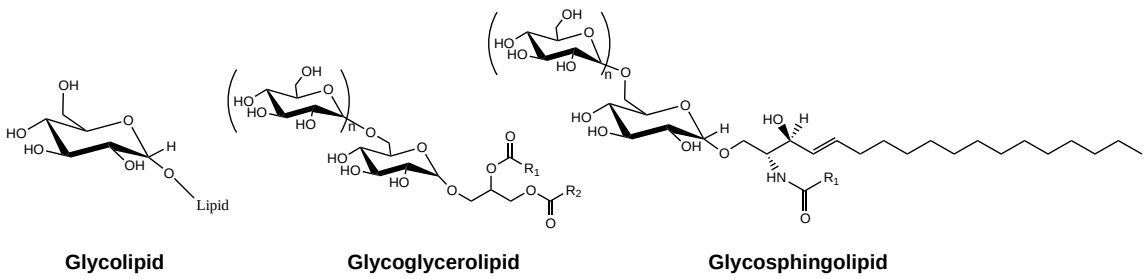

Figure 1. Glycolipid structure. Glycoglycerolipid is formed by diacylglycerol moiety bound to glucose or galactose while glycosphingolipid is formed by ceramide bound to glycoside units that can be very complex. R1 and R2 refer to the apolar part of the compound

Regarding the biological roles of glycolipids, one of the most important characteristics of a living cell is its ability to prompt and adequate the behavior during formation, maintenance and pathogenesis of tissues. Cell to cell interactions play an important role during the development. These interactions are highly dynamic processes that include migration, recognition, signaling, adhesion and attachment. All these abilities are possible due its capacity to interact with other cells and components. In this line, carbohydrates, the most prominently exposed structures, serve as important players in these events where they have been recognized as interaction sites in cell adhesion. As an example, in lymphocyte cell homing, carbohydrate increase the adhesion of circulating lymphocytes to the specialized tissues, tissue maintenance or host-pathogen interaction (Karlsson, 1989, 1986; Springer, 1990). In glycolipids, due to their amphipathic structure, the saccharide part is exposed to the medium where it can interact with different proteins, hormones, toxins or viruses among others while the hydrophobic part remains attached to the membranes by weak bounds such as electrostatics or Van der Waals linkages (Faivre and Rosilio, 2010).

1.2. Applications of glycolipids

In the last decades, a growing interest in renewables energies and less toxic compounds in order to reduce the human impact in the environment has been observed. Chemical or pharmaceutical industries need to reconsider the way they produce their products to try to become eco-friendlier with the environment considering living cells or bioproducts as an alternative way of production with a less toxic approach. Glycolipids, due to their amphipathic structure that provides them with some interesting properties, have been one of the promising topics of study since its potential use in different fields and industries.

Among their potential applications it is of special interest to their roles as a biosurfactants drug delivery systems, antimicrobial agents and, also, signaling agents.

1.2.1. Glycolipids as biosurfactants

The high demand of surfactant nowadays is being filled by a numerous, mainly petroleum-based, chemical surfactants that are usually toxic to the environment and also, non-biodegradable (e.g., sodium dodecyl sulphate and linear alkylbenzene sulfonates) (Schramm, 2010). The interest of glycolipids in this field has increased in the last decade since their properties confers them low

toxicity, high biodegradability and environment compatibility and also prompted applications not only in food, cosmetic and pharmaceutical industries but also in environmental protection and energy-saving technologies (Banat et al., 2000).

One of the major producer sources of GL biosurfactants are microbes where actinobacteria are known to produce almost 50% of the total production. Compared to the chemical produced ones, biosurfactants present the following properties: they are sophisticated structures with functional groups and chiral centers that present lower toxicity and higher biodegradability rates. They also present lower critical micelle concentration with higher surface activity and more gradual adsorption and activity while an increased resistance to physical factors such as temperature and ionic strength and biological functions (i.e., antitumor and antimicrobial activities) (Muthusamy et al., 2008).

Biosurfactants can be classified as glycolipids (rhamnolipids, sophorolipids, trehalose lipids), lipopeptides (i.e., surfactin) and polymeric compounds (i.e., emulsan). Among them, the most studied compounds are rhamnolipids due to their properties (Chong and Li, 2017). These compounds are produced by hydrocarbon-degradating organisms being the most one biosurfactant producer *Pseudomonas aeruginosa*. Since they are promising alternatives to petroleum based surfactants, Bahia *et al* engineered *Saccharomyces cerevisiae* to obtain mono-rhamnolipids in a non-pathogenic strain (Bahia et al., 2018).

Its main action mechanism is based on surface interactions mediated by the amphiphilic nature of the molecules that contain hydrophobic and hydrophilic regions. Surfactants are able to interact at the interfaces between aqueous and non-aqueous components of a complex system and at liquid gas interfaces (Marchant and Banat, 2012). The biosurfactants present hydrophilic parts composed by carbohydrates and peptides while fatty acids (saturated, unsaturated, branched or hydroxylated) are the hydrophobic ones. Examples of the GL structures of biosurfactants can be found in Figure 2.

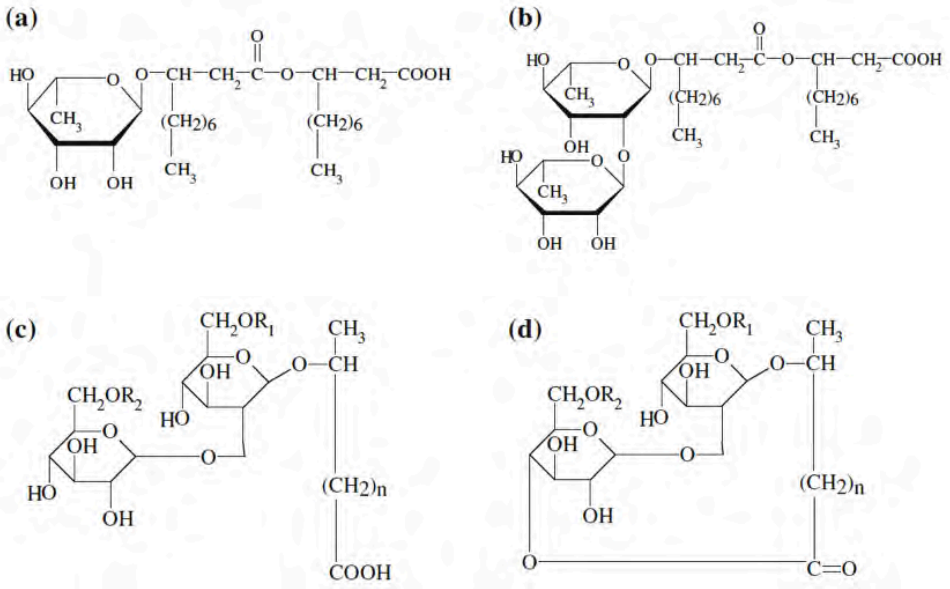

Figure 2. Examples of biosurfactants structure: (a) mono-rhamnolipid (b) di-rhamnolipid (c) sophorolipid in open chain (d) sophorolipid in lactonic form (Marchant and Banat, 2012)

One of the majors fields of interest until now for biosurfactants has been the bulk industries where domestic cleaning products, laundry detergents (Sajna et al., 2013), plastics, textiles, cosmetics (Lourith and Kanlayavattanakul, 2009), nutrition (Mnif and Ghribi, 2016; Nitschke and Costa, 2007) and pharmaceutical are the main fields of study. However, new uses for these products are currently being studied being an example of the formation and disruption of bacterial biofilms and other medically related applications that will be further discussed in the next sections (Rodrigues et al., 2007).

1.2.2. Glycolipids as drug delivery systems

In the last decades, nanotechnology has become a challenging innovation in different industries such pharmaceutical and chemical ones. One of the main goals of this innovation pursues the delivery of bioactive compounds inside the body in a controlled way. Drugs need to go through different biological barriers in order to reach the desired target. Normally, these devices need to protect the drug from degradation, improve the absorption and modify the pharmacokinetics and dynamics of it (Figure 3). It is still difficult for some drugs to reach the tissue and cell population due to the biological barriers present in the organism.

Glycolipids have the ability to self-organize into complex supramolecular structures because of their amphiphilic nature both in solvent (lyotropic properties) or temperature depending (thermotropic behavior). Thanks to their properties, synthetic glycolipids have been specially designed for the preparation of self-assembled drug delivery systems (Corti et al., 2007; Muthusamy et al., 2008; Vill and Hashim, 2002).

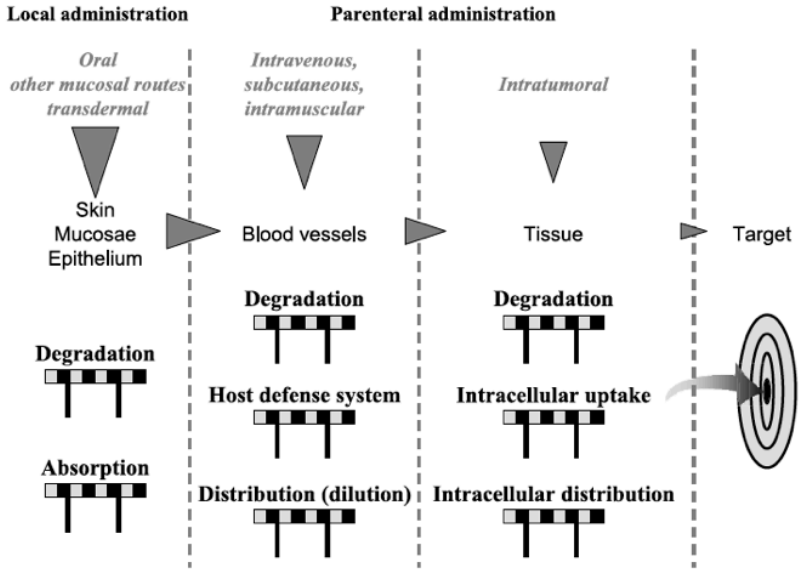

Figure 3. Barriers that drug delivery systems need to overcome to reach the target (Couvreur and Vauthier, 2006)

From the first liposomes proposed in 1974 by Tavill and coworkers (Gregoriadis et al., 1974), there was an exponential increase in the number of drug delivery systems (Couvreur and Vauthier, 2006) (Figure 4). Liposomes mimic natural cell membranes and have been investigated

as drug carriers due to the excellent capacity to trap the drugs, biocompatibility and safety suggesting that lipids and glycoconjugates such as sulfated lactosyl archaeol (SLA) could be an excellent drug carrier (He et al., 2019; Jia et al., 2019).

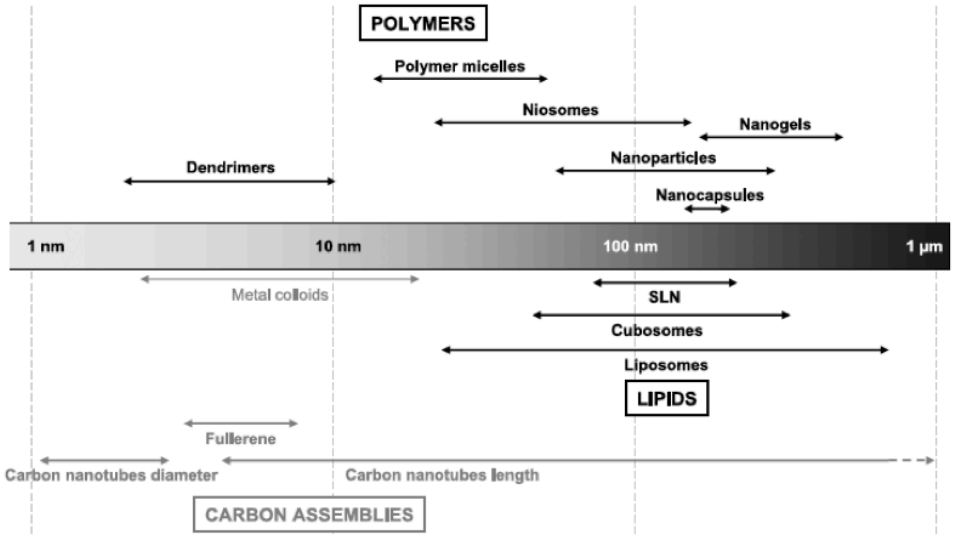

Figure 4. *Types of nanotechnology used for drug delivery and targeting (Couvreur and Vauthier, 2006)*

1.2.3. Glycolipids as antimicrobial agents

It has been reported that various rhamnolipids can exhibit antibacterial and antiphytoviral effects. These properties are related to their capacity to prevent the microbial adhesion, avoiding their penetration into the cells (Li et al., 2004; Linington et al., 2006).

Hossain et al reported two antimicrobial glycolipids in their study of the lipids from *Borrelia burgdoferi*, bacterium responsible for the Lyme disease. Immunoassays showed antibody reactivity against glycolipids while lipids of other organisms seemed to be non-immunogenic, suggesting a possible role for the glycolipids of this bacterium as promising candidates for the diagnosis of this disease (Hossain et al., 2001).

Furthermore, Reid reported that the premenopausal vagina of women was rich in lactobacilli flora, bacteriocins and biosurfactants (i.e., glycolipids) that appeared to protect the host. This work emphasized a possible probiotic role of these compounds and suggested an alternative treatment and preventive treatment to antibiotics in the future (Reid, 2001).

Moreover, Gan et al applied the biosurfactant secreted by *Lactobacillus fermentum* RC-14 to inhibit surgical implant infections caused by *Staphylococcus aureus* showing a significant inhibition and a possible role for biosurfactants in preventing microbial adhesion (Gan et al., 2002). Attachment to the mucous membranes seems to be the initial event of the infectious processes and is considered as an important virulence factor. Therefore, by inhibiting this adhesion the infection can be avoided. The substances involved in this attachment are in their majority proteins of virus, bacteria, yeast or parasite while, in the case of the host, the majority

of structures exposed that play allow this attachment are glycoconjugates (Karlsson, 1989, 1986).

Among glycolipids, the ones that have been reported to impact more in the immune system are glycosphingolipids since they play an important role in the pathogen invasion and modulation of the immune system. Recently, these lipids can be used to produce antibody-based immunity approaches in cancer (Zhang et al., 2019).

1.2.4. Vaccine adjuvants

Vaccines are compounds containing antigens synthetized with the aim to either prevent or treat illness by inducing a strong immunologic response (humoral and cellular). Adjuvants, from Latin *adjuvare*, (meaning enhance), are helpers that potentiate the effect of the vaccine into the immune system. Among these adjuvants and depending on their origin we can find microorganism and nucleic acid-based derivatives or even mineral salts. Regarding this last one, the most used in research is Alum, a mixture of aluminum salts that has been widely used. The main problem of this mixture is that cannot stimulate Th1 response. The use of carbohydrates could suppose an improvement regarding its compatibility and good tolerability in the organism while promoting both Th1 and Th2 responses (Wang et al., 2015).

Glycoconjugate vaccines are composed of sugar antigens that are covalently bound to a carrier. Normally the bond between the sugar and the carrier is made randomly even though new studies suggest that the area where this takes place has an influence on the immunological response (Stefanetti et al., 2015). An example of a glycolipid that can be used as an adjuvant is trehalose–6,6'–dibehenate which is an analogue of trehalose–6,6'–dimycolate (TDM) from *Mycobacterium tuberculosis.* This compound presents antitumoral activity and can stimulate host resistance against infections (Lemaire et al., 1986; Sueoka et al., 1995). Another interesting glycolipid of this organism is lipoarabinomannan, which is abundant in the bacterial cell wall, and is able to induce high antibody titters upon infection (Behren and Westerlind, 2019).

Using glycolipids (i.e., natural or synthetic α-galactosylceramides) as a powerful agonist for iNKT, a subpopulation of T lymphocytes, has led new approaches to increase a wide variety of immune responses, including those involved in the vaccination against infections and cancer (Carreño et al., 2014; Kharkwal et al., 2016). The ability of iNKT cells to respond under certain conditions to normal CD1d expressing cells is believed to reflect their recognition of self-lipids. Among the endogenous lipids and phospholipids candidates, cellular glycosylphosphatidylinositols (GPIs) and glycosphingolipids have been responsible for iNKT selection and homeostatic maintenance (Gumperz et al., 2000; Joyce et al., 1998; Liu and Guo, 2017; Zhou et al., 2004).

The activation of iNKT cells by these compounds leads to the rapid secretion of multiple cytokines, including those typical of Th1 and Th2 responses. This activation also causes transactivation of other immune cells, such as dendritic, T and B cells, thereby enhancing cellular and humoral immune responses. Despite the potential of using α-galactosylceramide as an activator of iNKT cells, a limited success in this area has been achieved in humans (Nair and Dhodapkar, 2017).

Howell and coworkers reported a new generation of synthetic α-galactosylceramides that combines a sphinganine base with a hydrocinnamoyl ester on C6" of the sugar that are able to maintain the activity seen *in* vitro in humans (Chennamadhavuni et al., 2018) (Figure 5).

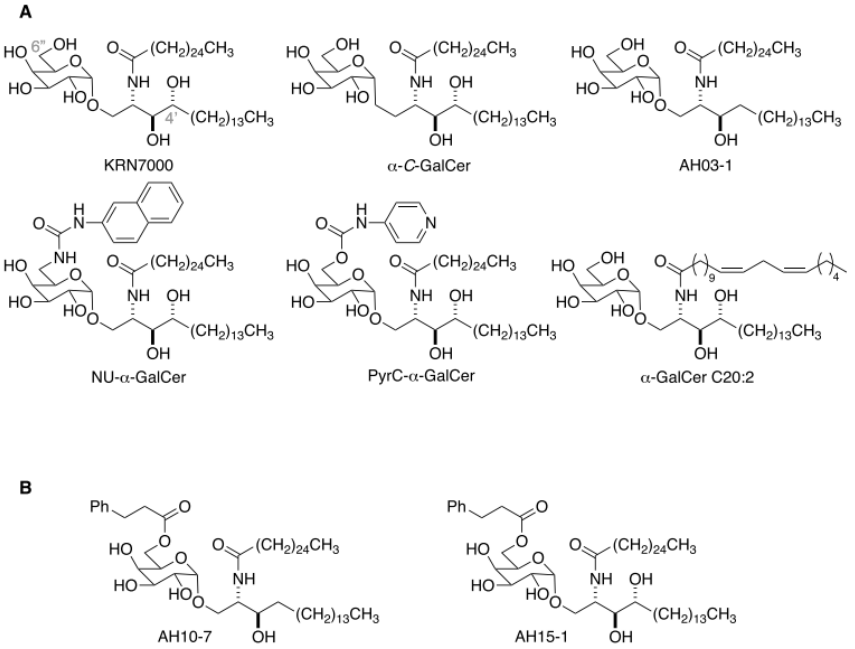

Figure 5. *Structures of α-galactosylceramides. (A) Structures of KRN7000 (with 4' and 6'' positions noted) and some examples of cytokine biasing derivatives. (B) Structure and synthetic strategy for C600 modified α-GalCer derivatives AH10-7 and AH15-1 (Chennamadhavuni et al., 2018).*

It has also been reported that a mixture of mono- and di- forms of seven different glycoglycerolipids (galactolipids) from *Ficus microcarpa* showed an inhibitory effect on TNF-α-induced IL-8 secretion raging up to 40% of inhibition when using concentrations of 50 μM of galactolipids helping to prevent the epithelial cell injury and apoptosis (Kiem et al., 2012). In this line, Ghosh et al reported that glycoglycerolipids and its truncated forms from *Meiothermus taiwanensis* were able to suppress the cytokine production of monocytes and therefore the immune response to avoid inflammation (Ghosh et al., 2013).

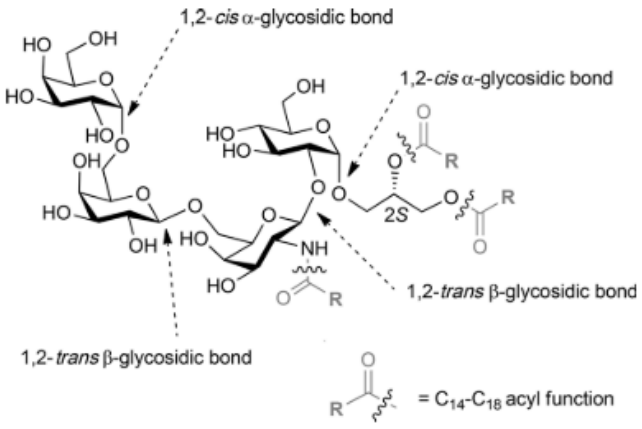

Figure 6. Structure of the GGL isolated from Meiothermus taiwanensis (Ghosh et al., 2013)

1.2.5. Glycolipids in tumor pathogenesis

Akasaka et al reported that the glycoglycerolipid (monogalactosyldiacylglycerol, MGDAG) extracted from spinach enhances the anti-cell proliferation effect of gemcitabine (Akasaka et al., 2016, 2013). This effect is possible since MGDAG is reported to inhibit some DNA polymerases that are highly expressed in cancer cells (Maeda et al., 2011). In addition, they also reported that MGDAG additionally increases the cytotoxicity of radiation in pancreatic cancer cells. In this line of work, Maeda et al also reported a growth inhibition in colon tumor growth in mice when MGDAG was provided (Maeda et al., 2013).

Moreover, it has been reported that glycoglycerolipid analogues, derived from 2-O-ß-D-galactosylglycerol are able to inhibit Protein Kinases C (PKC), enzymes involved in the phosphorylation of several proteins and signaling cascades inducers, translocation to the plasma membrane and downstream signaling pathways in glioblastoma (Colombo et al., 2011).

It has been seen that invasive/metastatic properties of tumoral cells are controlled by multiple factors with complex machinery, many of which operate at cell surface membranes (Hakomori and Handa, 2002). It is particularly important the high expression of aberrant glycosphingolipids in a specific type of tumors which are involved in tumoral cell adhesion and signal transduction defining then the stage of tumor progression and invasiveness (Hakomori, 1996; Sakakura et al., 1998). Furthermore, it has also been reported that altered glycosylation patterns found in gangliosides are involved in aggressiveness and metastasis of neuro-ectoderm derived tumors such as melanoma and neuroblastoma, indicating that it is possible to target these compounds for therapeutic antibodies (Furukawa et al., 2006; Krengel and Bousquet, 2014; Rabu et al., 2012). Dewald et al reported in 2018 that Tumor Necrosis Factor (TNF) is involved in the expression of complex gangliosides in breast cancer relating the tumoral microenvironment and gangliosides with the development of the tumor (Dewald et al., 2018).

Related to this work, Dellanoy and coworkers reviewed the relationship between gangliosides and cancer along with the role of cytokine-induced glycosylation changes in both inflammation and cancer diseases due to changes in the signaling pathways (Dewald et al., 2016; Groux-Degroote et al., 2018, 2017).

An interesting work in breast cancer has been reported by Huang et al where they demonstrated that it was possible to increase the immune response against this type of cancer by using a mixture of Globo H as an antigen (specifically expressed in cancer cells) with α-galactosylceramide (Huang et al., 2013). This combination promoted an increase in the production of anti-GH IgG that specifically attacked these cancerous cells becoming a promising candidate for immunotherapy.

1.3. Glycolipid production

Glycoglycerolipids, as it is said before, are abundant membrane components in the photosynthetic tissues of plants and cyanobacteria, where they are mainly galactolipids (Andrés et al., 2012; Hölzl and Dörmann, 2007). Wider structural diversity is found in Gram-positive

bacteria such *Mycobacterium sp.* or *Lactobacillus plantarum* (Rakhuba et al., 2009; Wicke et al., 2000) where the head group linked to diacylglycerol contains glucosyl, galactosyl or mannosyl units with a diverse types of glycosidic linkages (Hölzl and Dörmann, 2007).

Regarding the production, it has been reported that the amount that can be extracted from plants varies from 8 to 700 mg for each 100 g of plants (Yunoki et al., 2009). Due to its potential uses, studying new ways to obtain these products has become one of the main fields of study. Nowadays, organic and chemoenzymatic synthesis can produce glycolipids but in low quantities (Du et al., 2007; Faivre and Rosilio, 2010; Klement et al., 2007; Mao et al., 2006). Due to the variation of the fatty acids profile of glycoglycerolipids is difficult to quantify the amount GGL produced.

1.3.1. Chemical synthesis

The chemical synthesis of GGL has become really difficult because of the regio- and stereoselectivity required for the product. To achieve the degree of pureness required it is necessary to do multi-step protections and deprotections steps, which consequently lowers the yield and efficiency. In addition, using high amounts of organic solvents and reagents is far from the goals of green chemistry (Anastas and Kirchhoff, 2002).

Glycosyl fluorides and trichloroacetamidates are common reagents in these glycosylation reactions and typically afford moderate yields (30-60%). Purification of these products can become challenging due to α/ß mixtures. Glycosyl iodide studies led to important improvements in reaching the desired stereoselectivity. Reactions of per-O-benzylated galactosyl iodide with an azido sphingosine exclusively produced the α-anomer in over 90% yield (Figure 7) (Du et al., 2007). Due to unfavorable hydrogen bonding interactions, an azido group is used to place the amide because of the amide deactivates the primary hydroxyl of the acceptor.

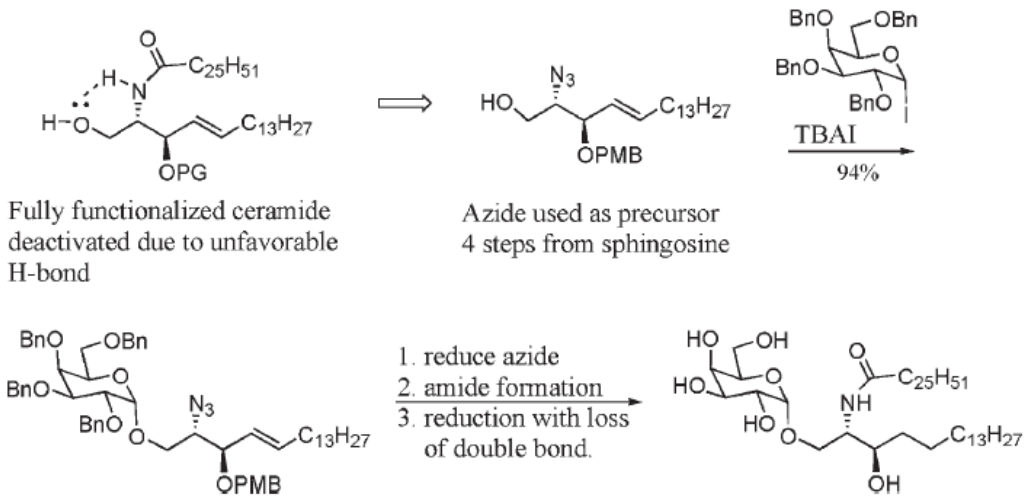

Figure 7. Example of how to synthesize glycolipids by chemical synthesis using per-O-benzylated galactosyl iodide and an azido sphingosine (Du et al., 2007)

It is known from previous studies that per-O-benzyl galactosyl iodide would not glycosylate ceramide acceptors with the amide group, but per-O-silyated glycosyl iodides are more reactive and are able to react onto them obtaining yields between 30 to 89% and high α-stereoselectivity (Du and Gervay-Hague, 2005). Advantages like easy preparation in large scale and mild acidic conditions for group deprotection (to protect the glycosidic bond) make this kind of donor promising candidates.

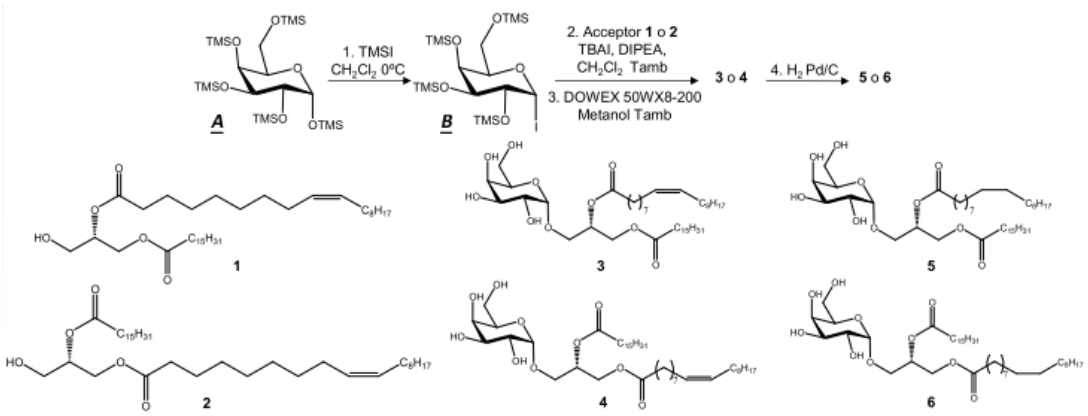

Figure 8. Glycoglycerolipids synthesis example. Beginning the synthesis with compound A, it reacts with trimethylsilyl iodide (TMSI) to give galactosyl iodide (compound B).

1.3.2. Enzymatic synthesis

By using enzymes, it is possible to reach the desired stereo- and regioselectivity. This strategy presents several drawbacks such as purification procedures, cost of substrates required to obtain the desired product or difficult conditions to produce them.

This approach has been widely used regarding the capabilities of enzymes to select specifics and produce a desired product with a stereo- and regioselectivity. In this line, several efforts have been directed in redesigning enzymes to increase its activity or affinity to certain substrates while also providing products with the desired structure. Withers and coworkers reported in 2001 a strategy to rapidly screen the aglycone specify of a glycosidase and thereby determine which enzymes were best suited to catalyze specific transglycosylation reactions (Blanchard and Withers, 2001).

Our group has also been involved in the redesign of several enzymes (Aragunde et al., 2014; Codera et al., 2015; Faijes et al., 2006; Faijes and Planas, 2007; Pérez et al., 2011; Pozzo et al., 2017; Val-Cid et al., 2015). Recently, C. Alsina et al reported in 2019 the redesign of a ChiA from *Thermococcus kodakaraensis* to increase the production of a chitooligosaccharides with higher degrees of polymerization. This is a product of interest due to the biological activities that this oligosaccharide present (Alsina et al., 2019). The best producer enzyme reported was ChiA D1022A with a formation of insoluble polymers, high polymerized product, 45% yield (w/w) and containing about 55% of DP10 (target chitooligosaccharide). This work showed that it was possible to modify a hydrolytic enzyme to a synthetic enzyme obtaining high yields.

Another example of reported work on enzymatic engineering can be found in Yu et al who reported a highly efficient streamlined chemoenzymatic strategy for total synthesis of four

prioritized ganglioside cancer antigens GD2, GD3, fucosyl GM1 and GM3 from commercially available lactose and phytosphingosine by using one-pot multienzyme (OPME) glycosylation reactions (Yu et al., 2018).

Regarding specifically to the production of GGL, it has already been reported by our group that by using MG517 is possible to sequentially transfer glucosyl or galactosyl units to a diacylglycerol acceptors obtaining GGL (Andrés et al., 2012). The first reaction takes place between glucosyl or galactosyl moieties and diacylglycerol forming monoglucosyldiacylglycerol or monogalactosyldiacylglycerol. Afterwards, another saccharide is linked to the sugar moiety of the glycolipid forming diglucosyldiacylglycerol or digalactosyldiacylglycerol and so on (Figure 9). The main drawback of this method is the highly cost of the substrates (UDP-Glc, UDP-Gal and DAG) and the difficulty to purify and maintain the activity of this enzyme due that glycosyltransferases are enzymes normally associated to cell membranes. This makes difficult the purification of these enzymes regarding to its hydrophobic required environment.

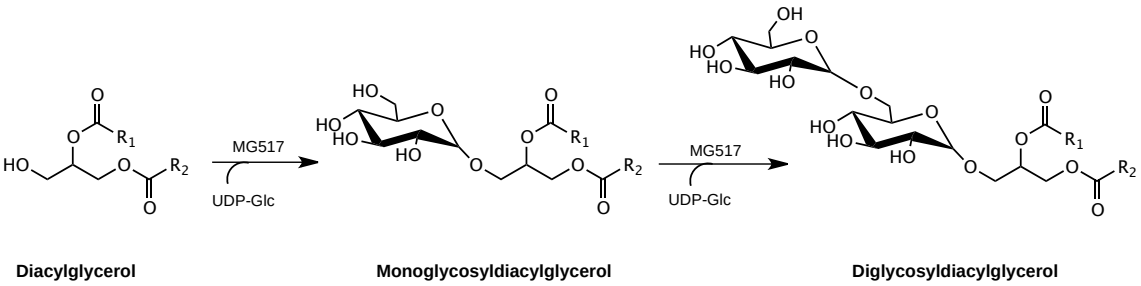

Figure 9. Sequential reaction of MG517 using diacylglycerol as acceptor and UDP-glucose as a donor

1.3.3. Cell factory approach

Due to the high cost that enzymatic synthesis may suppose and the complexity related to it in terms of required substrates and purification procedures, it was proposed to use cells as factories. To achieve so, it is required that the cells to be used have the precursors needed to produce the product and, especially, the enzymes to catalyze the desired reaction. Being so, several studies have reported the use of cells to produce products of added value. Several projects have been focused in the production of lipidic products for the biofuel industry, but little has been done in the field of glycolipids. One example of cell factory approach focused in the production of glycolipids can be found in the production of rhamnolipids. This type of glycolipids are produced in large quantities by *Pseudomonas aeruginosa*. The main drawback about using this organism is its pathogenicity, which may cause safety and health concerns during large-scale production and applications. To solve so, several metabolic engineering efforts have been applied to either decrease the pathogenicity of this strain or use non-pathogenic strains by introducing key genes in the rhamnolipids production (Chong and Li, 2017).

Regarding the GGL production, it has been reported that *Mycoplasma sp.* is able to produce these compounds in a very high stereo- and regioselectivity. Mycoplasma is a genus of bacteria that lacks cell walls around their cell membrane, and it is known as a parasitic lifestyle of

eukaryotic organisms. Without cell walls, these organisms are not affected by many common antibiotics, which makes them really difficult to eliminate. These Gram-positive related bacteria contain glycoglycerolipids in their membranes as key structural components involved in the bilayer properties and stability (Razin et al., 1998; Razin and Hayflick, 2010). As previously mentioned, this organism contains a membrane-associated glycosyltransferase, MG517, able to produce glycoglycerolipids but its pathogenicity and difficulties of growth by its own makes it not recommended to be used in large-scale production.

Our group previously reported that it was possible to use *Escherichia coli* as a cell factory to produce GGL by subcloning MG517 into a pET44b(+) plasmid (Mora-Buyé et al., 2012). The idea of using *E. coli* is due to their easy culturing, short doubling time population, non-pathogenesis and easily genetically manipulated. In addition, *Escherichia coli* has been widely used in the production of several products including those to be used in biopharmaceutical applications where it was reported that from more of 100 recombinant proteins a 34% of these were obtained by *E. coli* (Meyer and Schmidhalter, 2012). Furthermore, *E. coli* contains the required precursors, UDP-Glc and DAG, to produce GGL.

To produce glycoglycerolipids it is necessary to have two main precursors. On one hand, the glycoside donor, UDP-glucose (UDP-Glc) is obtained from glucose-1-phosphate by uridilytransferase, GalU while on the other hand, the other required precursor, diacylglycerol (DAG), can be obtained by PgpB from phosphatidic acid (PA). This phosphatidic acid is also the precursor of phospholipid synthesis.

Regarding the biosynthesis of these precursors, to produce the lipid acceptor diacylglycerol, everything starts with glycerol-3-phosphate. This compound is then acylated by PlsB acyltransferase forming lysophosphatidic acid (LPA), which is sequentially acylated by PlsC acyltransferase to form phosphatidic acid (PA). This PA is one of the main lipidic precursors since all the phospholipids synthetized starts from him. If this phosphatidic acid is then dephosphorylated by phosphatidylglycerophosphatase B, also known as PgpB, diacylglycerol is obtained. All these enzymes are associated with cell membranes.

On the other hand, to obtain the glycosyl donor, UDP-Glucose, it is required *pgm* to obtain glucose-1-phosphate from glucose-6-phosphate. This compound is then transformed into UDP-Glc by the union of a UDP group by the action of UTP-glucose-1-phosphate uridilytransferase (GalU). This reaction is reversible and the enzyme in charge of doing so is UDP-sugar hydrolase (UshA).

Planas and coworkers tried different conditions in order to reach the maximal production and finally concluded that the highest amount of glycoglycerolipids was obtained when MG517 was co-expressed along with PlsC acyltransferase (Mora-Buyé et al., 2012). The required enzymes to produce DAG and UDP-Glc are already naturally found in *E. coli* so to produce GGL it was necessary to subclone MG517 into a suitable plasmid for *E. coli* and see if this enzyme was functional in *E. coli* to produce GGL in this organism (Figure 10).

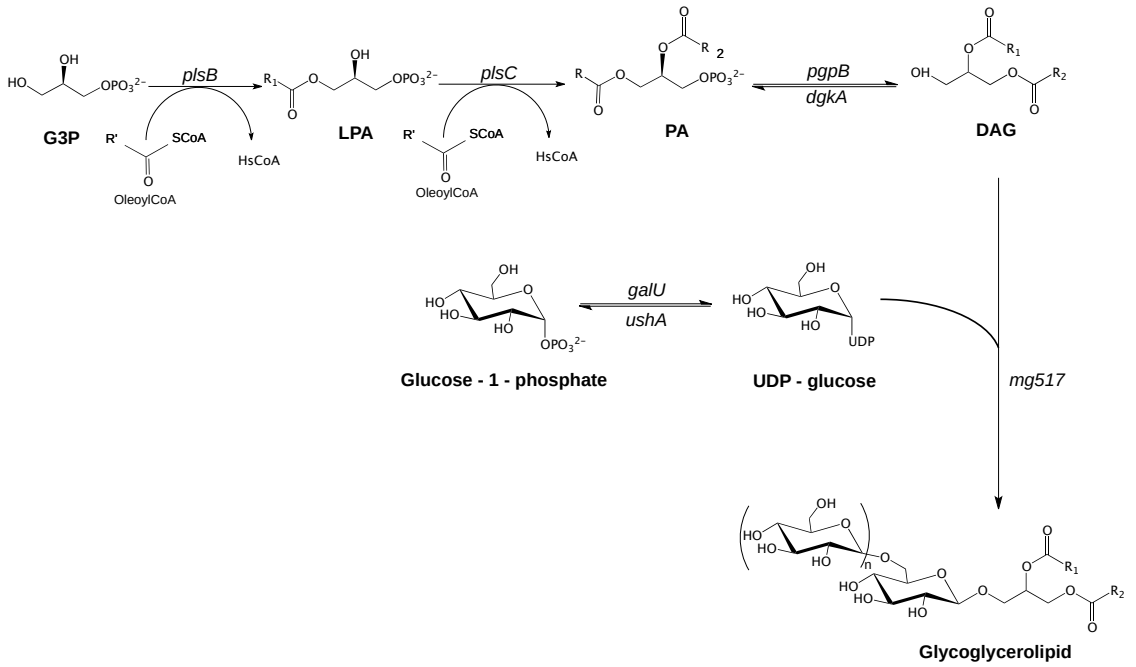

Figure 10. Metabolic pathway to synthetize glycoglycerolipids in E. coli

1.4. Framework of the project

Glycoglycerolipids are highly complex molecules with high interest due to their use as drug delivery systems or antitumoral activity. They present sugar moieties with a regio- and stereoselectivity linked to the lipidic part which in addition presents chirality. Nowadays, enzymatic and chemical synthesis still present difficulties to synthesize these types of molecules. Our group proposes to use *E.* coli as cell factories in a new strategy to produce these complex structures. As a proof of concept, it was already reported by our group that when MG517 glycosyltransferase was expressed in *E.* coli this organism was able to produce GGL, compounds that were not produced naturally by this organism (Andrés et al., 2012; Mora-Buyé et al., 2012). Results showed *E. coli* as a promising organism to use metabolic engineering strategies to increase the UDP-glucose and diacylglycerol pools to be further used to produce GGL.

In the study of Mora-Buyé et al, tested *in vivo*, it was reported that the highest amount of GGL was achieved when MG517 was co-expressed along with PlsC, the acyltransferase responsible for the conversion of lysophosphatidic acid into phosphatidic acid, precursor of diacylglycerol.

In this project, studies of the lipid and glycoside metabolisms were performed to identify the state-of-the-art and identify the most promising genes to be removed from the *E. coli* genome or which genes should be overexpressed to increase as maximum as possible the availability of the precursors that could later be used for GGL synthesis.

1.4.1. Fatty acid metabolism

The biosynthesis of fatty acids to produce the membranes is an energy-intensive and vital phase of the cell physiology. Bacteria and plants are able to do it by using a very conservative collection of enzymes called Type II Fatty Acid Synthase system (FAS II) (White et al., 2005). The knowledge

on this field has gone from a rudimentary understanding of the major enzymes and lipid structural classes to a high-resolution crystal structures of the representative members of all major protein categories (Rock and Jackowski, 2002). The fatty acid metabolism in *E. coli* is well known and regarding its fast growth, simple nutritional requirements and available genetic tools, this organism serves as an excellent host for the fatty acid production. There are other organisms that have been engineered to produce fatty acids such as *Saccharomyces cerevisiae* and *Yarrowia lipolytica*, a yeast able to produce large amounts of fatty acids (Coelho et al., 2010; Papanikolaou and Aggelis, 2010; Qiao et al., 2015). The main drawback of these organisms is that its higher complexity leads to more complex engineering processes. Figure 11 presents the lipid metabolism in *E. coli* reported by Janßen et al in 2014 (Janßen and Steinbüchel, 2014).

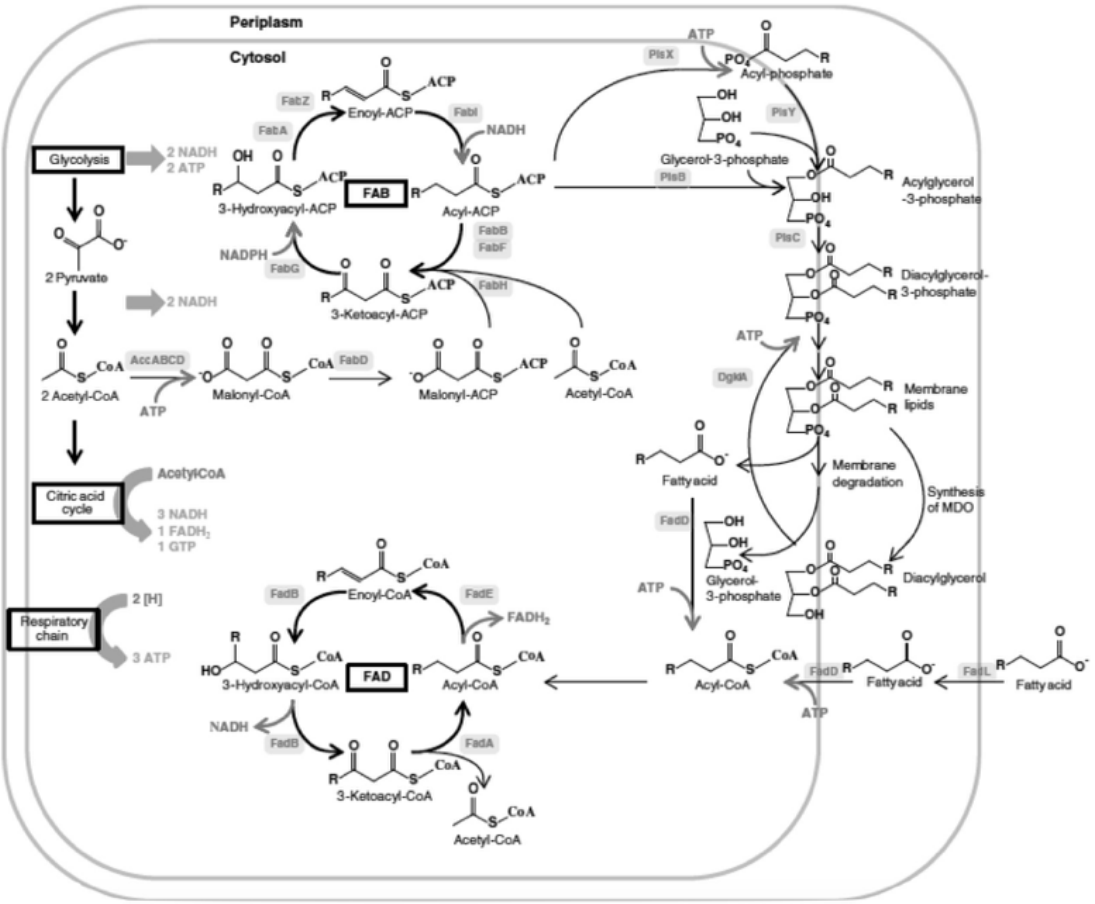

Figure 11. Lipid metabolic pathway in E. coli (Janßen and Steinbüchel, 2014)

As it can be seen, many enzymes are involved in these pathways making it very complex to target which genes would be key to increase the production of fatty acids. A literature review of the main metabolic strategies used to increase the production of fatty acids is presented in Table 1.

Table 1. Metabolic strategies used to increase the production of free fatty acids using different bacteria

Strategy		Increase folds	Strain	Compared to	Yield	Productivity	Reference
ΔfadD, plant thioesterase heterologous expression, ACC overexpression		20 fold	BL21(DE3)	WT	2.5 g/L	0.17 g/L·h	(Lu et al., 2008)
overexpression ACC mutant *'tesA*		6 fold	BL21(DE3)	WT	6 nmol	-	(Davis et al., 2000)
overexpression of *fabD*		11% more (relative quantity)	ML103 (MG1655 without fadD)	MG1655	0.3-0.5 g/L	0.012-0.020 g/L·h	(Zhang et al., 2012)
GLY module overexpression	GLY in medium (*mGLY-hACA-hFAS)*	18 fold	BL21(DE3)	WT	1.3 g/L	0.0188 g/L·h	(Xu et al., 2013)
ACA module overexpression							
FAS module overexpression	*m*GLY-*h*ACA-*h*FAS	20 fold	BL21(DE3)	WT	1.4 g/L	0.0203 g/L·h	
FadR overexpression		7.5 fold	FRT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1, pE8a-fadR)	LT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1)	5.2 g/L	0.072 g/L·h	(F. Zhang et al., 2012b)
FabA overexpression		1.25 fold	FAT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1, pE8a-fadA)	LT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1)	0.9 g/L	0.012 g/L·h	
FabB overexpression		3 fold	FBT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1, pE8a-fadB)	LT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1)	1.8 g/L	0.016 g/L·h	
FabF overexpression		3 fold	FFT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1, pE8a-fadF)	LT-ΔfadE (*E. coli* DH1:ΔfadE, pKS1)	2 g/L	0.027 g/L·h	

Strategy	Increase folds	Strain	Compared to	Yield	Productivity	Reference
ΔfadD + Acyl-ACP thioesterase overexpression (Control strain)	-	ML103_18	-	3.1 g/L	0.064 g/L·h	(San and Li, 2017)
ΔfadD + ΔsucC + FabZ and Acyl-ACP thioesterase overexpression	1.81 fold	MLK163_18Z	ML103_18	5.6 g/L	0.116 g/L·h	
ΔfadD + ΔsucC + ΔfabR + fabZ and acyl-ACP thioesterase overexpression	1.65 fold	MLK212_18Z		5.2 g/L	0.108 g/L·h	
ΔfadD + FabZ and Acyl-ACP thioesterase overexpression	1.48 fold	ML103_18Z		4.61 g/L	0.096 g/L·h	
ΔfadD + fadR and acyl-ACP thioesterase overexpression	1.34 fold	ML103_18fadR		4.2 g/L	0.087 g/L·h	
ΔfadD + ΔsucC + Acyl-ACP thioesterase overexpression	1.27 fold	MLK163_18		4 g/L	0.083 g/L·h	
ΔfadD + ΔsucC + ΔfabR + Acyl-ACP thioesterase overexpression	1.23 fold	MLK212_18		3.8 g/L	0.079 g/L·h	
ΔfadD + ΔfabR + Acyl-ACP thioesterase overexpression	1.20 fold	MLK211_18		3.7 g/L	0.077 g/L·h	
ΔfadD + ΔfadR + FabZ and Acyl-ACP thioesterase overexpression	1.19 fold	MLK225_18Z		3.7 g/L	0.077 g/L·h	
ΔfadD + ΔfabR + FabZ and Acyl-ACP thioesterase overexpression	1.16 fold	MLK211_18Z		3.6 g/L	0.075 g/L·h	
tesA' expression	12 fold	tesA' (*E. Coli DH1)*	*E. Coli* DH1 (wt)	0.3 g/L	0.006 g/L·h	(Steen et al., 2010)
tesA' + ΔfadD	35 fold	tesA' ΔfadD (E. Coli DH1)	*E. Coli* DH1 (wt)	0.7 g/L	0.014 g/L·h	
tesA' + ΔfadE	50 fold	tesA' ΔfadE (E. Coli DH1)	*E. Coli* DH1 (wt)	1 g/L	0.02 g/L·h	
Δ*fadE*, *fadR* + *tesA* overexpression	15 fold	*E. coli* DH1 Δ*fadE*/pA58C-TR	*E. coli DH1 ΔfadE/pACYCDuet-1*	3.05 g/L	0.13 g/L·h	(He et al., 2014)
fabF overexpression	1.4 fold	JESH1001 (*fabF*)	MG1655	0.18 g/L	0.007 g/L·h	(Lee et al., 2013)
fabF and *fabG* overexpression	1.68 fold	JESH1002 (*fabF* and *fabG*)	MG1655	0.21 g/L	0.009 g/L·h	

Strategy	Increase folds	Strain	Compared to	Yield	Productivity	Reference
fabF and *fabZ* overexpression	2 fold	JESH1003 (fabF and fabZ)	MG1655	0.25 g/L	0.010 g/L·h	(Lee et al., 2013)
fabZ and *fabI* overexpression	1.2 fold	JESH1004 (fabZ and fabI)	MG1655	0.15 g/L	0.006 g/L·h	
fabF, *fabG* and *fabZ* overexpression	2.8 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.35 g/L	0.014 g/L·h	
fabF, *fabZ* and *fabI* overexpression	3.2 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.40 g/L	0.016 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	3.8 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.48 g/L	0.02 g/L·h	
fabF overexpression	1.35 fold	JESH1001 (fabF)	MG1655	0.14 g/L	0.006 g/L·h	
fabF and *fabG* overexpression	1.5 fold	JESH1002 (fabF and fabG)	MG1655	0.15 g/L	0.006 g/L·h	
fabF and *fabZ* overexpression	1.2 fold	JESH1003 (fabF and fabZ)	MG1655	0.12 g/L	0.005 g/L·h	
fabZ and *fabI* overexpression	1.1 fold	JESH1004 (fabZ and fabI)	MG1655	0.11 g/L	0.004 g/L·h	
fabF, *fabG* and *fabZ* overexpression	1.75 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.18 g/L	0.008 g/L·h	
fabF, *fabZ* and *fabI* overexpression	1.75 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.18 g/L	0.008 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	1.75 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.18 g/L	0.008 g/L·h	
fabF overexpression	1.4 fold	JESH1001 (fabF)	MG1655	0.12 g/L	0.005 g/L·h	
fabF and *fabG* overexpression	1.4 fold	JESH1002 (fabF and fabG)	MG1655	0.13 g/L	0.005 g/L·h	
fabF and *fabZ* overexpression	1.2 fold	JESH1003 (fabF and fabZ)	MG1655	0.11 g/L	0.005 g/L·h	
fabZ and *fabI* overexpression	1.1 fold	JESH1004 (fabZ and fabI)	MG1655	0.1 g/L	0.004 g/L·h	
fabF, *fabG* and *fabZ* overexpression	1.8 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.16 g/L	0.007 g/L·h	
fabF, *fabZ* and *fabI* overexpression	1.7 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.15 g/L	0.006 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	1.8 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.16 g/L	0.007 g/L·h	
fabF overexpression	1.1 fold	JESH1001 (fabF)	MG1655	0.11 g/L	0.005 g/L·h	
fabF and *fabG* overexpression	0.8 fold	JESH1002 (fabF and fabG)	MG1655	0.08 g/L	0.003 g/L·h	
fabF and *fabZ* overexpression	1 fold	JESH1003 (fabF and fabZ)	MG1655	0.1 g/L	0.004 g/L·h	
fabZ and *fabI* overexpression	1.25 fold	JESH1004 (fabZ and fabI)	MG1655	0.125 g/L	0.005 g/L·h	
fabF, *fabG* and *fabZ* overexpression	0.8 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.08 g/L	0.003 g/L·h	
fabF, *fabZ* and *fabI* overexpression	1.4 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.14 g/L	0.006 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	1.5 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.15 g/L	0.006 g/L·h	
fabF overexpression	2 fold	JESH1001 (fabF)	MG1655	0.2 g/L	0.008 g/L·h	

Strategy	Increase folds	Strain	Compared to	Yield	Productivity	Reference
fabF and *fabG* overexpression	2.3 fold	JESH1002 (fabF and fabG)	MG1655	0.23 g/L	0.010 g/L·h	
fabF and *fabZ* overexpression	1.75 fold	JESH1003 (fabF and fabZ)	MG1655	0.175 g/L	0.007 g/L·h	
fabZ and *fabI* overexpression	1.75 fold	JESH1004 (fabZ and fabI)	MG1655	0.175 g/L	0.007 g/L·h	
fabF, *fabG* and *fabZ* overexpression	3.3 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.33 g/L	0.013 g/L·h	
fabF, *fabZ* and *fabI* overexpression	3.25 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.325 g/L	0.013 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	3.3 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.33 g/L	0.014 g/L·h	
fabF overexpression	2.22 fold	JESH1001 (fabF)	MG1655	0.2 g/L	0.008 g/L·h	(Lee et al., 2013)
fabF and *fabG* overexpression	2.5 fold	JESH1002 (fabF and fabG)	MG1655	0.23 g/L	0.010 g/L·h	
fabF and *fabZ* overexpression	2.8 fold	JESH1003 (fabF and fabZ)	MG1655	0.25 g/L	0.01 g/L·h	
fabZ and *fabI* overexpression	2.5 fold	JESH1004 (fabZ and fabI)	MG1655	0.23 g/L	0.010 g/L·h	
fabF, *fabG* and *fabZ* overexpression	4 fold	JESH1005 (fabF, fabG and FabZ)	MG1655	0.36 g/L	0.015 g/L·h	
fabF, *fabZ* and *fabI* overexpression	3.9 fold	JESH1006 (fabF, fabZ and fabI)	MG1655	0.35 g/L	0.014 g/L·h	
fabF, *fabG*, *fabZ* and *fabI* overexpression	4.5 fold	JESH1007 (fabF, fabG, fabZ and fabI)	MG1655	0.41 g/L	0.017 g/L·h	

To obtain fatty acids (FA), everything starts with acetyl-CoA carboxylase which is composed of four subunits. This enzyme directs acetyl-CoA towards de *novo* fatty acid biosynthesis and chain elongation (Janßen and Steinbüchel, 2014). In *Escherichia coli* the complex is formed by four subunits that are strictly transcriptionally regulated. This means that in order to reach a functional protein, the subunits need to be synthetized in equimolar amounts. As it can be seen in Table 1, this enzyme is a common target in metabolic engineering strategies due to its involvement in the availability of malonyl-CoA (Zha et al., 2009). The overexpression of this carboxylase was able to report a 6-fold increase in the fatty acid synthesis (Davis et al., 2000). In addition, several studies have reported the effect of overexpressing this enzyme in combination with other modifications. As an example, Lu et al reported a 20-fold increase when combining the overexpression of this carboxylase with a knockout of *fadD,* gene involved in the transport and activation of exogenous fatty acids prior to their subsequent degradation or incorporation to phospholipids (Lu et al., 2008).
S-malonyl transferase, encoded by *fabD*, is the enzyme responsible to transfer the malonyl moiety to ACP to form malonyl-ACP which is the first step of the fatty acid production. Modifications in the active site of this enzyme allowed the obtaining of new varieties of the acyl carrier protein that could allow to obtain a variety in the fatty acid biosynthesis (Marcella and Barb, 2017). X. Zhang et al, reported that an overexpression of this enzyme also led to an increase in the production of fatty acids (X. Zhang et al., 2012).
The formation of 3-ketoacyl-ACP is catalyzed by the 3-ketoacyl-ACP synthase I, II and III, encoded by *fabB, fabF* and *fabH*. This enzyme is able to condense malonyl-ACP with fatty acyl-ACP. Depending on the synthase, the substrates are different. For example, *fabH* uses malonyl-ACP and acetyl-CoA as substrates initiating this way the first cycle of chain elongation. The following elongation steps are performed by *fabF* and *fabB*. *fabH* is part of the *fabHDG* operon making more complex the control of its expression. Reports of using the overexpression of these genes showed a slight increase in the FA production (My et al., 2015, 2013).
Other authors have tried more complex modifications. Xu et al for example, elaborated a modular optimization of multigene pathways with several genes in different copy number plasmids into the own modules, and working at the same time with plant fatty acyl- ACP thioesterases (Xu et al., 2013). Modules GLY, ACA and FAS, to modulate glycerol, malonyl-CoA-acetyl-CoA and fatty acid concentrations respectively, were studied. Best results were obtained with medium copy number for GLY, low copy number for ACA and high copy number for FAS plasmids, obtaining a 20-fold increase in terms of fatty acid production.
In 2017, Tan et al reported an improvement of the *Escherichia coli* membrane integrity and fatty acid production by the expression of OmpF, an outer membrane porin that mediates the non-specific diffusion of small solutes such as sugars or aminoacids (Tan et al., 2017).
Another strategy that has been studied consisted in the accumulation of FA by removing genes related with fatty acid degradation pathways. The most common targeted genes are the ones that catalyze the firsts steps of the ß-oxidation, specifically *fadD*. Different modifications have been tested with different results (Lu et al., 2008; San and Li, 2017; Xu et al., 2013). Moreover, Dellomonaco et al (Dellomonaco et al., 2011) demonstrated the functional reversal of the fatty acid

ß-oxidation by using CoA thioesters intermediates and acetyl-CoA for acyl chain elongation instead of malonyl-CoA.

1.4.2. Glycosidic metabolism

Living cells use carbohydrate metabolism to obtain the required energy to maintain cellular functions and allow cell growing and division. Carbohydrate metabolism, which is well known, is very complex and provides to the cell a high number of intermediates that are latter involved in different metabolic reactions. One of the key molecules produced is pyruvate which can be involved in the Krebs cycle, gluconeogenesis or FA synthesis depending on the cellular circumstances (if it requires energy or storage) (Figure 12). Another example of an important intermediate would be UDP-Glc which after being synthetized is used in a high number of reactions as a building block of the sugars for the cells via glycosylation.

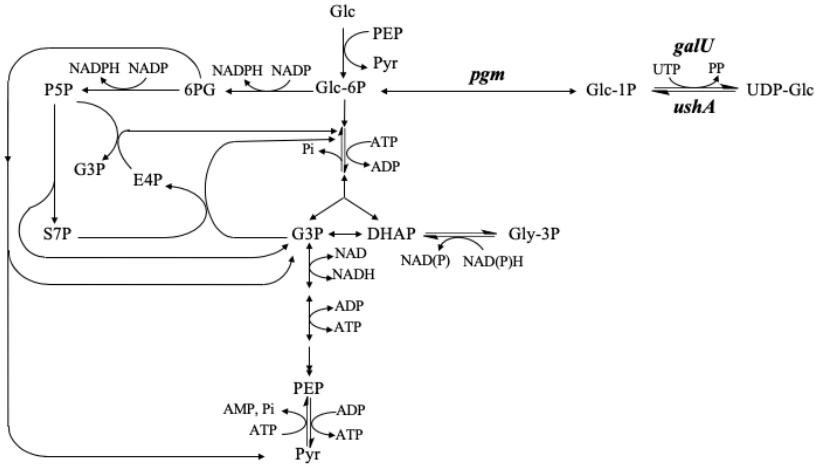

Figure 12. Escherichia coli UDP-Glc biosynthesis pathway

Glycosylation reactions use these nucleotides sugars to create new bioactive molecules. These reactions are based on allowing the formation of a covalent bond between a glycosidic donor (i.e., UDP-Glc) to other molecules such lipids or proteins among others catalyzed by glycosyltransferases enzymes.

Glycosylation is an important modification in nature as it plays a role in many metabolic pathways. It has several purposes such allowing the correct folding of proteins or providing stability to a certain molecule. It can also be used to activate compounds and provide them with certain characteristics that can later be used to perform specific reactions. One example of this can be found in cytokines or hormones. When these molecules are glycosylated, this has a direct impact in their specificity and activity. Moreover, glycosylation of small molecules (i.e., secondary metabolites) play an important role in their solubility, activity or stability making these products very interesting for different industries such as pharmaceutics, food additives or therapeutics.

It is known that many diseases are caused by alterations in the glycosylation patterns either caused by genetic disorders or by other alterations, being some examples Gauche disease, cancer or Wiskott-Aldrich syndrome (Durand and Seta, 2000; Freeze et al., 2015).

UDP-glucose (UDP-Glc) acts as an intermediate in the glycosidic catabolism and is a major reactive in carbohydrate metabolism, specially, to synthetize oligomers and polymers of glucose along with other complex carbohydrates since it can act as a glucose donor and also act as an intermediate in the galactose metabolism. Therefore, the availability of UDP-Glc had been studied thoroughly in living cells, enzymatic synthesis or multienzyme biocatalysis (De Bruyn et al., 2015a; Weyler and Heinzle, 2015; Woo et al., 2019).

In the current project, since UDP-glucose is a precursor for the biosynthesis of glycoglycerolipids, a special interest has been focused on the enzymes responsible for the synthesis of this compound being those GalU uridilytransferase or UshA UDP-sugar hydrolase, responsible for the opposite reaction of GalU. UshA has several substrates besides UDP-Glc but studies provided by De Bruyn et al reported that by knocking out this gene it was possible to increase UDP-Glc pool in *Escherichia coli* (De Bruyn et al., 2015b; Pandey et al., 2014, 2013)

OBJECTIVES

2. Objectives

The main objective of this project is to obtain a biological platform by metabolic engineering strategies using *E. coli* of BL21 Star (DE3) as a cell factory. Thus, this project is focused on obtaining a cell factory able to reproduce the complex structure of glycoglycerolipids and that is compatible with large-scale production. Moreover, the biological platform will provide a different profile of glycolipids either based on differences in the mono-/di-/tri-glucosyldiacylglycerol ratio or abundance of each fatty acid species (saturated vs unsaturated fatty acids).

Therefore, the specific objectives of this project are:

1) To design different strategies to improve GGL production by increasing the abundance of the precursors involved in the biosynthesis of glycoglycerolipids:
   a) By removing competing reactions enhancing the production of diacylglycerol to later be used to produce GGL.
   b) By removing competing reactions involved in the biosynthesis of the glycosidic donor, UDP-Glc
   c) By studying the role that different enzymes involved in the biosynthetic pathways of phosphatidic acid, DAG and UDP-Glc might have in the production of GGL by the:
      i) Overexpression of genes involved in the synthesis of diacylglycerol
      ii) Overexpression of enzymes involved in the synthesis of UDP-glucose

2) To characterize the different strains obtained in the specific objective 1 in order to determine the different parameters related to the GGL production, stability and type of product produced. This objective was reached by:
   a) Characterizing the growth rate of each strain obtained in this project to analyze the effect that metabolic modifications may have.
   b) Analyzing the production of glycoglycerolipids.
   c) Determining the enzymatic activities of the targeted genes related to the production of the precursors.
   d) Determining the fatty acid profile in order to analyze if the removal of certain genes or overexpression of others led to changes in the lipidic profile.
   e) Analyzing the effect that glycoglycerolipids may have in the phospholipid abundance.

3) To obtain a promoter and RBS library of a targeted gene related to the biosynthesis of phosphatidylethanolamine, major phospholipid found in *E. coli* membranes, to increase the production of glycoglycerolipids.

PROPOSED STRATEGIES TO INCREASE THE PRODUCTION OF GGL

3. Proposed strategies to increase the production of GGL

3.1. Designed strategies

The first generation of engineered strains contained different combinations of MG517 with GalU and PlsC enzymes, both enzymes involved in the biosynthesis of UDP-Glc and DAG precursors (Mora-Buyé et al., 2012) (Figure 13). As previously described, our group reported that when GGL were produced in *E. coli* from diacylglycerol and UDP-glucose, the bottleneck of its synthesis was diacylglycerol availability. This was concluded when the overexpression of MG517 glycosyltransferase in combination with acyltransferase PlsC afforded an increase of the GGL production.

Considering the prior information, this work was focused on trying to increase the diacylglycerol and UDP-glucose pools, decreasing the bottleneck in the production of GGL and therefore obtaining more GGL. To achieve so, five different strategies were proposed to impact in the availability of DAG, lipidic precursors, and UDP-Glc, glycosidic donor.

The **first strategy** was focused on increasing the acyl donor availability by blocking competing reactions. Acyltransferases PlsB and PlsC, responsible in *E. coli* for the transfer of the acyl groups to *sn*-glycerol-3-phosphate (G3P) to give lysophosphatidic and phosphatidic acid, respectively, use the acyl-acyl carrier protein (acyl-ACP) as well as acyl-CoA thioesters generated by the activation of exogenous fatty acids as acyl donors (Yao and Rock, 2013). Since these acyl donors can be degraded via fatty acid β-oxidation or hydrolyzed to free fatty acids, we proposed to increase of acyl donor pool through the deletion of the *fadE* and/or *tesA* genes involved in these competing reactions (Figure 13).

Acyl-CoA dehydrogenase (FadE) is a key enzyme that catalyzes the oxidation of acyl-CoA to 2-enoyl-CoA as the first step of β-oxidation (He et al., 2014). Previous studies in biofuels research reported the importance of knocking out the gene encoding this enzyme to increase the production of free fatty acids (FFA) (He et al., 2014; Janßen and Steinbüchel, 2014; Lu et al., 2008; Steen et al., 2010). Moreover, when combining the *fadE* knockout with overexpression of thioesterases such as *tesA*, FFA productivity was further increased (Bentley et al., 2016; Steen et al., 2010).

Thioesterase I encoded by *tesA* on the other hand, is responsible of catalyzing the hydrolysis of acylated carriers during fatty acid synthesis in order to release FFA. Specifically, this gene encodes a thioesterase type I that is specific for long chain fatty acids (C_{12}-C_{18}) (Lee et al., 1997). It was reported that redirecting the enzyme from the periplasm to the cytosol by removing the leader sequence had an positive effect on the FFA production (Cho and Cronan, 1995). Moreover, when TesA was overexpressed alone or in combination with a *fadD* knockout or ACC overexpression, an increase of FFA was observed (San and Li, 2017; Steen et al., 2010; Xu et al., 2013). Therefore, we proposed that the elimination of the thioesterase activity on acyl donors (Δ*tesA*) together with the acyl-CoA dehydrogenase (Δ*fadE*), which initiates the β-oxidation

pathway after activation of fatty acids by the acyl-CoA synthetase FadD, would cause an increase in the acyl donor pool and, therefore, the phosphatidic acid precursor to produce DAG.

The **second strategy** aimed to increase fatty acid availability by modulating important transcriptional regulators that control both the expression of fatty acid degradation and synthesis genes. The *fabR* gene is known to repress *fabA* and *fabB* genes regulating expression of type II fatty acid synthases and, therefore, the synthesis of unsaturated fatty acids (Marrakchi et al., 2002; Zhang et al., 2002). *fadR* on the other hand is an activator of fatty acids biosynthesis (*fab* operon) and a repressor of their degradation (*fad* operon) (Fujita et al., 2007; Karp et al., 2014). Different studies showed that *fadR* overexpression resulted in a significant increase of FFA production in comparison to overexpression of individual genes such as *fabA* or *fabB* (He et al., 2014; F. Zhang et al., 2012b). Moreover, *fadR* overexpression in combination with *fabR* deletion produced a significant improvement in fatty acid titer (San and Li, 2017). In this context, in this project we proposed knocking out *fabR* and overexpress *fadR* with the aim of increasing the acyl donor availability by transcriptional modulation in the ΔtesA/ mg517 background.

The **third strategy** aimed at increasing the conversion of acyl donors to phosphatidic acid (PA) through the overexpression of PlsC and PlsB acyltransferases. The first reaction to produce PA is catalyzed by PlsB, which converts glycerol-3-phosphate into lysophosphatidic acid. Once this product is obtained, acyltransferase PlsC is able to add a second acyl moiety to form PA, which is the precursor of DAG and the initial precursor of phospholipids biosynthesis (Parsons and Rock, 2013; Yao and Rock, 2013) (Figure 13). Our group had previously reported an increase in the production of GGL when PlsC acyltransferase was overexpressed along with MG517 glycosyltransferase (Mora-Buyé et al., 2012). These acyltransferases have also shown to be key targets to increase the production of other lipid derivatives such as triglycerides and polydroxyalkanoates in *Yarrowia lipolytica* and *Rhodosporidium toruloides* (Celińska et al., 2019; Czerwiec et al., 2019; Park et al., 2018; Rigouin et al., 2019). In this approach, both high and low copy number plasmids expressing PlsC and PlsB were used in order to evaluate the effect of the expression levels in the ΔtesA/ mg517 background.

The **fourth strategy** intended to increase phosphatidic acid production from phospholipids and used it for DAG synthesis. The overexpression of the hydrolase CDP-diacylglycerol pyrophosphatase (CDH), enzyme that catalyzes the conversion of CDP-DAG to PA, could increase PA availability for conversion to DAG at expenses of decreasing CDP-DAG for the synthesis of phosphatidylethanolamine and other essential phospholipids (Figure 13). This is especially interesting since our group and others have shown that it is possible to interchange phosphatidylethanolamine and monoglucosyldiacylglycerol in the membranes (Mora-Buyé et al., 2012; Xie et al., 2006) thus being a potential strategy to further increase the production of GGL. Alternatively, PA can be pushed towards DAG by the overexpression of the phosphatidic phosphatase PgpB, a key regulatory enzyme in lipid metabolism, responsible for the conversion of phosphatidic acid to DAG. However, PgpB is also involved in the last step of phosphatidylglycerol biosynthesis, which is a competing reaction for our purpose (Wikstrom et

al., 2009). Therefore, in this project the overexpression of a fusion protein plsCxPgpB in order to redirect the flux towards DAG and not to phospholipids biosynthesis was proposed.

The **fifth strategy** proposed was based on increasing the availability of UDP-Glc which is the glycosidic donor of GGL. The enzymatic reaction involved in the synthesis of this precursor is based on the conversion of Glc-1-P by GalU into UDP-Glc while the reverse reaction, catalyzed by UshA is responsible for the conversion of UDP-Glc into Glc-1-P. Two different approaches could be followed in this strategy. The first one was based on overexpressing GalU while the second one was based on knocking out *ushA*.

Overall, strategies can be summarized as follows:

1. Increase of acyl donor by removing competing reactions
2. Increase FA availability by modulating transcription factors
3. Increase PA availability by overexpressing acyltransferases
4. Increase PA availability from phospholipids
5. Increase UDP-Glc availability

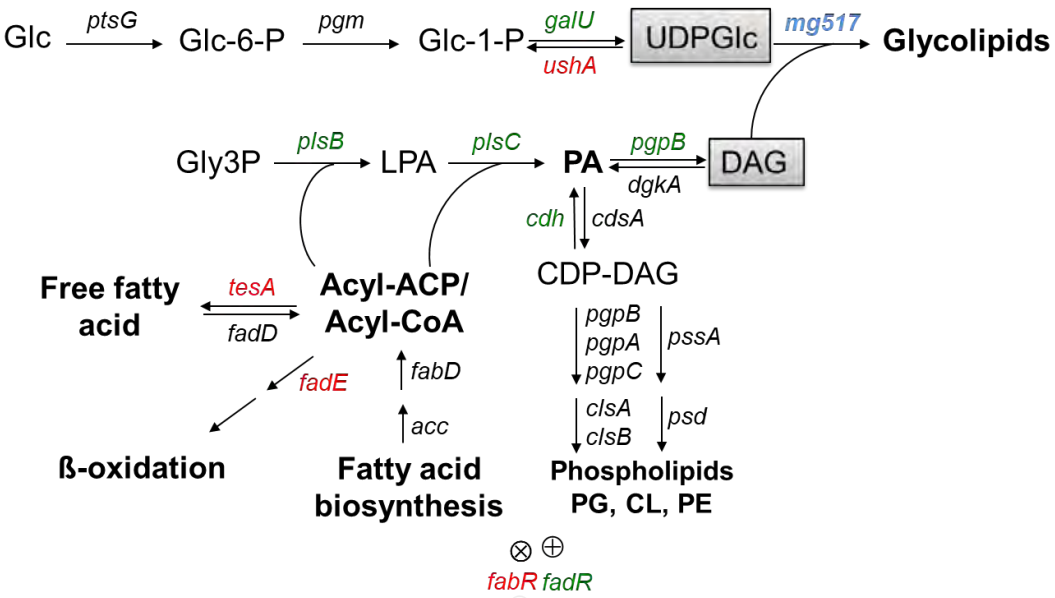

Figure 13. Metabolic engineering strategies to increase the production GGL from UDP-Glc and DAG precursors. In green and in red, overexpression and knockout of different genes, respectively. In blue, mg517, common to all strains. Enzymes encoded by the gene presented in the figure: ptsG: glucose-specific PTS enzyme IIBC component (EC 2.7.1.199); pgm: phosphoglucomutase (EC 5.4.2.2); galU: UTP-glucose-1-phosphate uridlytransferase (EC: 2.7.7.9); ushA: UDP-sugar diphosphatase (EC 3.1.3.5.3.6.1.45); plsB: glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15); plsC: 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51); pgpB: phosphatidylglycerophosphatase B (EC 3.1.3.27 3.1.3.81, 3.1.3.4, 3.6.1.27); dgkA: undecaprenol kinase (EC 2.7.1.66); cdsA: phosphatidate cytidyltransferase (EC 2.7.7.41); cdh: CDP-diacylglycerol diphosphatase (EC 3.6.1.26); pgpA: phosphatidylglycerophosphatase A (EC 3.1.3.27); pgpC: phosphatidylglycerophosphatase C (EC 3.1.3.27); clsA: cardiolipin synthase A (EC 2.7.8.-); clsB: cardiolipin synthase B (EC 2.7.8.-); pssA: phosphatidylserine synthase (EC 2.7.8.8); psd: phosphatidylserine decarboxylase (EC 4.1.1.65); tesA: thioesterase I (EC 3.1.2.2); fadD: fatty acyl-CoA synthetase (EC 6.2.1.3); fadE: acylCoA dehydrogenase (EC 1.3.8.7); fabD: malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39); acc: acetyl-CoA carboxyltransferase (EC 2.1.3.15); fabR: DNA-binding transcriptional repressor FabR; and *fadR*: DNA-binding transcriptional dual regulator FadR.

3.2. Molecular biology to follow the strategies

To remove competing reactions several genes were proposed to knockout. To obtain the strains where these genes were removed, a method based on Datsenko and Wanner was followed. This method is based on using helper plasmids to insert lineal DNA and recombine it with the gene to be extracted (Datsenko and Wanner, 2000).

In total, three different plasmids are required to obtain the KO. First one, pKD46, contains a λ red recombinase, inducible by arabinose that is responsible for allowing the homologous recombination event between the targeted chromosomal DNA locus and linear introduced DNA. This plasmid is sensitive to temperature and is not able to be replicated at 37ºC.

Plasmids pKD3 or pKD4 are required to insert the antibiotic resistance cassette and allow consequently the selection of the cells which exchanged the gene of interest for the resistance cassette. This cassette is flanked by FRT zones that, once pCP20 (last plasmid to be expressed) is added and the flipase that contains is expressed, allows the removal of the resistance cassette (Figure 14).

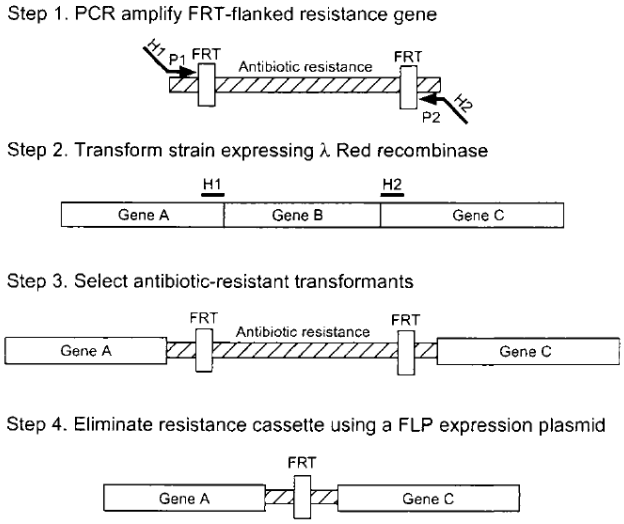

Figure 14. Procedure to obtain knockouts (Datsenko and Wanner, 2000)

Regarding to pKD3 and pKD4, the main difference between these two relies on the antibiotic resistance. In pKD3 chloramphenicol resistance would be added while when using pKD4 the cells that recombined the DNA would present resistance to kanamycin.

In order to design the linear DNA fragments to knockout the genes it is necessary to create homologous sequences of, at least, 50 bp for each flanking zone of the gene to be removed. The method used to obtain this linear fragment is based on a PCR where the template used could be either pKD3 or pKD4 and the primers would have at 5' the homologous sequence. For the *in silico* design of all the knockouts obtained in this project Clone Manager® software was used.

Other strategies involved the overexpression of several genes, alone or in polycistronic plasmids to analyze the effect that overexpressing these enzymes, related to the biosynthesis of the DAG and UDP-Glc precursors, may have in the overall production of GGL. To obtain the plasmids

different approaches were followed in this project. The first approach was based on Gibson et al work which uses the capacity of DNA nucleotides to complement each other (Gibson et al., 2009). They designed a method whereby linearizing the vector and the insert(s), the extreme of which were complementary to each other, it was possible by using exonucleases with 5'-3' activity, that allowed the exposure of the extremes and binding of the different fragments, and polymerase to obtain the desired plasmids. Another approach followed was CPEC which was also based on the DNA ability of nucleotides to complement each other and where it was only needed to have both linear fragments of DNA with, at least, 16 bp of homology between each other (Quan and Tian, 2014). The third approach followed in this project was based on using the classical restriction/ligation procedure while the fourth one, Golden Gate Assembly, was based on the ability of type II restriction enzymes to cut after the recognition sequence and the design of specific overlapping sequences of 4 nucleotides (Engler et al., 2008).

3.3. Analytical methods used to characterize the different strains obtained

3.3.1 Glycolipid quantification

The new engineered strains were characterized to determine the GGL production. Cultures were grown overnight in minimal medium and pellets were rinsed with 0.9% NaCl, lyophilized and finally a Folch extraction was performed to extract GGL from the cells (Folch et al., 1957). After this extraction two procedures were required to determine the amount of glycoglycerolipids. First of all, the glucose content in the membranes, which is related to the glycolipids, was quantified by using anthrone reaction (Bailey, 1958; Leyva et al., 2008; Shetlar, 1952; Yemm and Willis, 1954). As it can be seen in Figure 15, in this assay glucose from glycoconjugates can be hydrolyzed with sulphuric acid at 100ºC. Glucose is then dehydrated to 5-hydroxymethyl furfural which is able to react with anthrone reagent in acid conditions to produce a colorimetric complex that is read at 625 nm. Since samples came from lyophilized cells, it was necessary to resuspended them in one mL of water and treat them with phosphoric and sulphuric acid as protocol indicates (*see 13.3.2. Anthrone reaction section*).

One of the main advantages of this method is that allows the detection of total sugars, reduced and non-reduced. Moreover, this technique does not seem to interfere with any possible proteins placed in the samples (Percival Zhang et al., 2006).

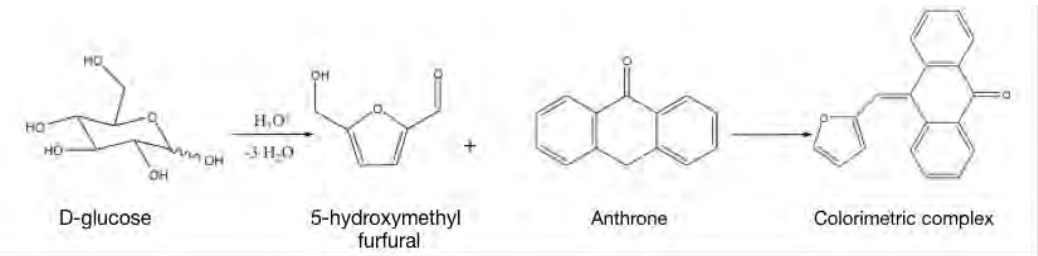

Figure 15. Anthrone reaction

To relate the amount of glucose quantified with the different forms of glycolipids produced by the cells, the distribution of mono-, di-, tri- and tetra-glycosyldiacylglycerols of the lipid fraction was analyzed by TLC. Spots were quantified by densitometry using ImageJ software (Schneider et al., 2012).

The total amount of GGL was determined considering both the glucose content (anthrone assay) and the GGL distribution (TLC analysis) of the lipid fraction according to equation 1:

$$[GGL]_T = \frac{[Glc]_T}{\%\text{MGDAG}+2\cdot\%\text{DGDAG}+3\cdot\%\text{TriGDAG}+4\cdot\%\text{TetraGDAG}} \quad (1)$$

where $[GGL]_T$ is the total amount of glycoglycerolipids in mmol/mg dry cells and %MGDAG, %DGDAG, %TriGDAG and %TetraGDAG are the percentage distribution of mono-, di-, tri and tetra-glycosyldiacylglycerols multiplied by the number of sugars of each form.

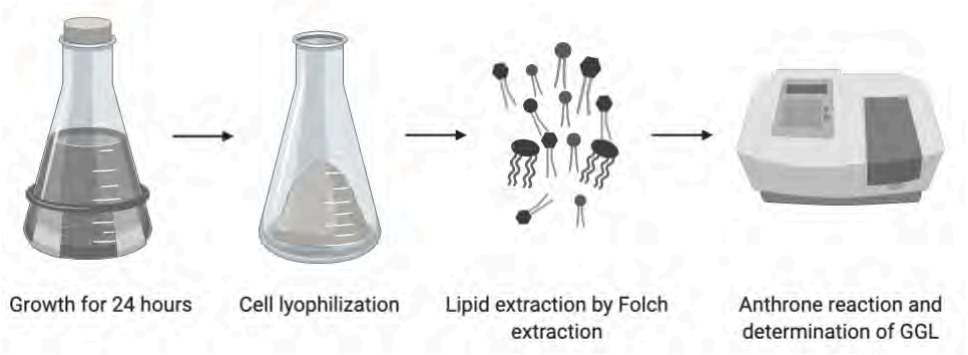

Figure 16. Summary of the analytical procedure followed to quantify GGL

3.3.1.1 Setting-up the analytical method

In our group, the anthrone methodology for glucose determination of GGL was set up. It was reported the linearity with a R^2 of 0.99 between 0 and 150 µg of glucose. The quantification limit reported was below 10 mg of cells as Table 2 shows.

Table 2. Data of the assay to correlate biomass with glucose

Samples (mg of cells)	OD $_{625\ nm}$	µg glc	µg glc/mg cells
5.75	0.005	0.09	0.020
12.1	0.057	9.53	1.260
21.96	0.110	18.24	1.410
53.76	0.280	49.65	1.080

3.3.2 Glycosyltransferase activity determination

In all the engineered strains, the MG517 glycosyltransferase was determined to evaluate its overexpression and amount. The glycosyltransferase activity present in the cell extracts was determined following the reaction between UDP-Gal and C6-NBD-ceramide (Figure 16).

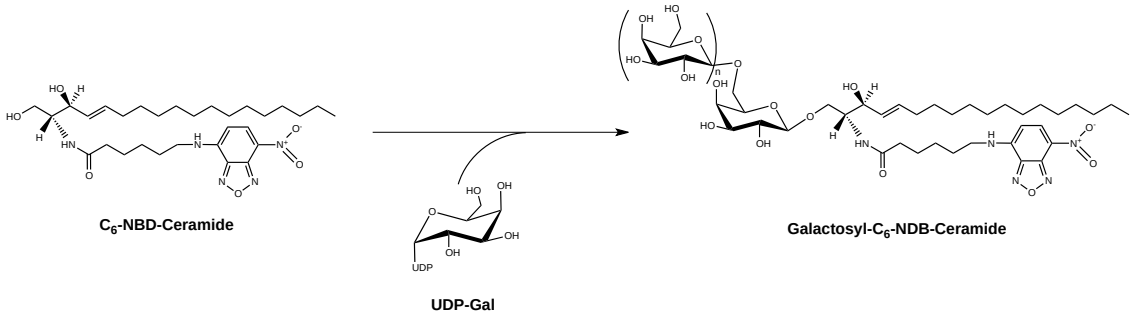

Figure 17. Reaction between C6-NBD-ceramide and UDP-Gal by the catalysis of MG517

This assay was based on the fluorescence of NBD moiety attached to the ceramide. By using a C18 column and a mixture of acetonitrile and water (75:25) as a mobile phase, it was possible to distinguish between the monogalactosylceramide (MGalCer), digalactosylceramide (DGalCer) and ceramide (Cer) that did not react since all of them presented different retention times (Figure 18).

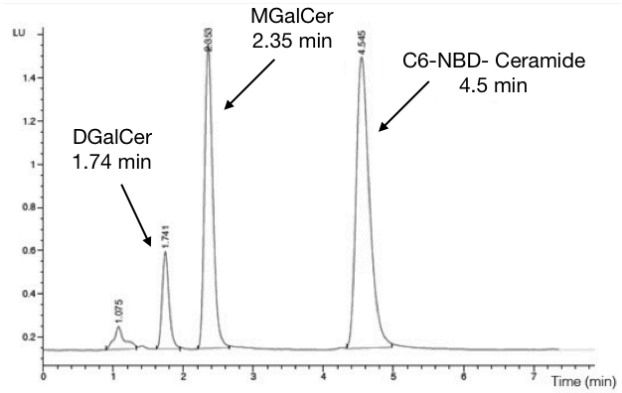

Figure 18. Chromatogram of MG517 reaction

Initial rates were determined following the progress curve of the reaction for 14 minutes. The specific activity was expressed as enzymatic activity per mg of total protein (protein of cell extract measured by BCA method) (Figures 19).

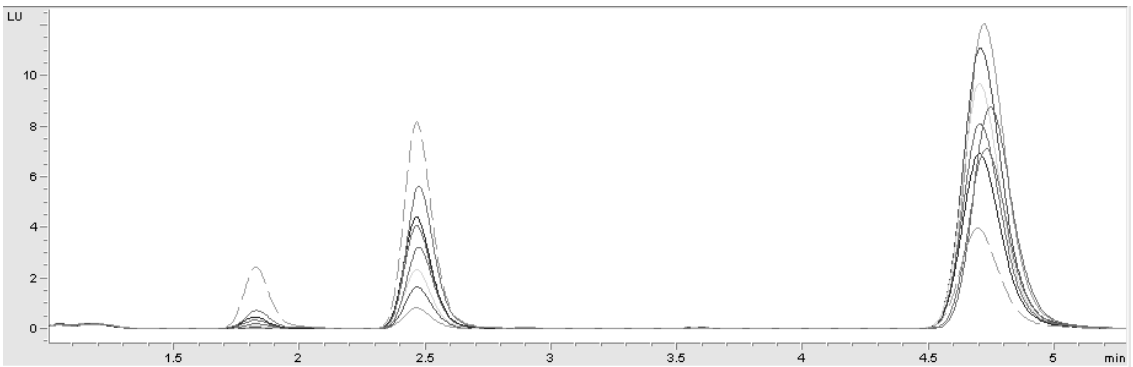

Figure 19. Overlap of the different chromatograms obtained during 2, 4, 6, 8, 10, 12, 14 and 25 minutes of reactions. As it can be seen, C6-NBD-ceramide decreases while MGalCer and DGalCer increase over time

Equation 2 (2) was used to calculate the amount of GGL produced:

$$[\mu M\ GGL] = \frac{(A_{MGalCer}+A_{DGalCer})}{(A_{MGalCer}+A_{DGalCer}+\ A_{Cer})} \times\ [Cer_0] \qquad (2)$$

Considering:

$A_{MGalCer}$: Area of monogalactosylceramide

$A_{DGalCer}$: Area of digalactosylceramide

A_{Cer}: Area of C6-NBD-Ceramide

Cer_0: Initial concentration of C6-NBD-Ceramide in µM

To determine the enzymatic activity of the glycosyltransferase in these engineered strains, cultures of 50 mL were set up and grown for 24 hours. An incubation with a buffer containing CHAPS, glycerol, HEPES, NaCl and $MgCl_2$ was performed for 24 more hours at 4ºC and slow agitation. Cells were then disrupted and the supernatant was kept being used for the reactions (*see 13.4.1. Glycosyltransferase activity section for more details*).

STRATEGY 1: INCREASE OF THE DAG AVAILABILITY BY REMOVING COMPETING REACTIONS

4. Strategy 1: Increase of the DAG availability by removing competing reactions

4.1. Study basis

The first strategy was focused on increasing the availability of diacylglycerol to later be used to produce GGL. To do so, a removal of *fad*E and *tes*A genes, both involved in the fatty acid β-oxidation and the hydrolysis of acyl-ACP were proposed (Figure 20).

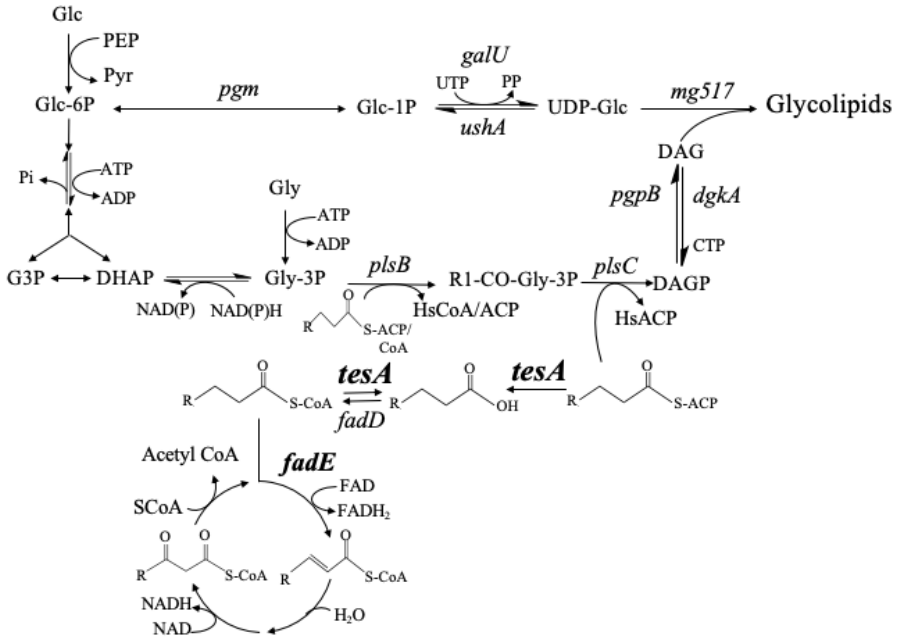

Figure 20. Escherichia coli's metabolic pathways to synthetize the glycoglycerolipids precursors, UDP-Glc and DAG.

AcylCoA dehydrogenase (**FadE**), which appears to be the only acyl-CoA dehydrogenase in *E. coli*, is the key enzyme to catalyze the oxidation of acyl-CoA into 2-enoyl-CoA being extremely important for the β-oxidation to take place (Campbell and Cronan, 2002; He et al., 2014). This enzyme was previously known as *yafH* open reading frame and was characterized by Campbell and Cronan in 2002.

Previous studies have reported the importance of knocking out this enzyme in order to increase the production of free fatty acids. An important field where this have been studied is in the biofuel production where alternatives to fossil fuels are being currently studied (Campbell and Cronan, 2002; He et al., 2014; Janßen and Steinbüchel, 2014; Lu et al., 2008; Steen et al., 2010). In some of these studies, knockouts of different genes involved in the degradation of fatty acids along with overexpression of thioesterases such as *tes*A reported an increase in the amount of FFA (Beld et al., 2015; Bentley et al., 2016). This combination generated 1g/L of FA in *Escherichia coli* DH1, representing an increase of the production between 35 and 50 folds (Steen et al., 2010). Furthermore, Zhang et al reported in 2012 that combining a strain with a deletion on this

gene with a overexpression of FadR fatty acid biosynthesis activator resulted in an increase of 7.5 folds of the desired product compared to the same strain without the overexpression of FadR (X. Zhang et al., 2012).

Since this enzyme is known to be the rate-limiting step of any cycle of oxidation of acyl-CoA (O'Brien and Frerman, 1977), which is the first step of ß-oxidation, it was proposed to remove it from the genome trying to block this way the degradation of fatty acids so these could impact on increasing acyl groups availability and affect DAG and GGL.

On the other hand, thioesterase I encoded by ***tesA*** is responsible to catalyze the deacylation of the FA during fatty acid synthesis pathway in order to release free fatty acids and form acyl carrier protein (ACP). Thioesterase I (TesA) is one of the thioesterases found in *Escherichia coli*, this enzyme is responsible for the generation of free fatty acids and it has been widely used to produce biofuels. It is specific for fatty acyl thioesters of greater than 10 carbons with a highest activity on palmitoyl, and palmitoleyl fatty acids, which are common lipids in *Escherichia coli* (Barnes and Wakil, 1968; Bonner and Bloch, 1972).

This enzyme is able to release acyl-ACP and acyl-CoA depending on the origin of the lipid (endogenous or exogenous respectively).

Other thioesterases found in *E. coli* would be Thioesterase III (FadM), which is involved in the ß-oxidation of oleic acid, and TesB that has a relatively broad substrate specificity cleaving medium and long-chain acyl-CoA substrates. Strains overexpressing or lacking this last enzyme do not show obvious defects but relieving inhibition of fatty acid synthesis by long-chain acyl-ACP.

TesA has been widely used in several metabolic engineering studies to produce free fatty acids for biofuel production due the main reaction that catalyzes. It was reported by Cronan and Cho that the amount of FFA could be increased by removing the leader sequence of *tesA* redirecting the enzyme to the cytosol instead of periplasm ('TesA) and showing an increase in the accumulation of FA (Cho and Cronan, 1995; Davis et al., 2000; Marella et al., 2018; Pan et al., 2017). The overexpression of this enzyme to produce FFA, has been combined with the deletion and overexpression of other genes reporting different levels of production (Cao et al., 2016; Liu and Khosla, 2010; Meng et al., 2013; F. Zhang et al., 2012a). Recently, it has been reported that by changing Arg^{64} for a Cys^{64} it was possible to increase the specific activity by 2 fold (Shin et al., 2016). Moreover, it was also reported the optimization of TesA to hydrolyze selectively octanoyl-ACP (Hernández Lozada et al., 2018). Other groups reported that by using Iterative Redesign and Optimization (IPRO) method it was possible to predict mutants of 'TesA with greater specificity towards C_{12}- or C_8- fatty acids that are interesting compounds for biofuel production (Grisewood et al., 2017). Using this software, it was possible to detect three mutants that significantly enhanced the production of C_{12} and 27 more mutants that were able to produce C_8 fatty acids. This software supposed an important tool for designing the enzyme to increase the production of a desired product.

The reason why this gene was selected to be removed was related to the need of having activated fatty acids, which are the substrate of this enzyme, to produce diacylglycerol and, therefore, glycoglycerolipids.

4.2. Obtaining knockout strains

The strain used to obtain the knockout strains and analyze the glycolipid production was *E. coli* BL21 Star (DE3). This strain has been widely used to overexpress recombinant proteins.

4.2.1. Δ*fadE*

To remove *fadE* from *E. coli* genome first it was analyzed the locus where this gene was placed. In there, it was seen that it was possible to remove the entire gene since there were no shared parts of the sequence with other ORFs. Two sets of primers were designed to remove this enzyme using Datsenko and Wanner method previously described (*see full protocol in 13.1.10. Datsenko and Wanner protocol to obtain KO strains).*

The first pair of primers were designed to create a homologous region that would allow the recombination event exchanging the genomic locus of *fad*E for the resistance cassette (oMEMO4468 and oMEMO4480). Figure 21 shows the *in silico* design of the primers to create the required homology at 5', both primers contained a sequence that bound to pKD3 plasmid in order to be able to add chloramphenicol resistance to the strains and select those where the recombination event was correctly performed *(full sequence in* 15.4. Primers used in this project*)*.

Figure 21. Designed primers to knockout of fadE in BL21 Star (DE3)

The second pair of primers, oMEMO4508 and oMEMO4509, were designed and used to check if the knockout was correctly obtained and the resistance cassette was correctly inserted into the genome removing *fade* sequence. If the procedure was correctly performed, the expected band would be around 2 kb (Figure 22A). Colony 3, showing the band at 2 kb, was transformed with pCP20 in order to remove the resistance cassette.

To confirm the correct removal of the gene, an aligning between the theoretical sequence and the sequence amplified from this strain was performed. Results showed that even some differences in the sequence appeared, since this region was no longer codifying for any protein and did not contain promoters or RBS it was concluded that BL21 Star (DE3) Δ*fadE* (Δ*fadE* from now on) was successfully obtained (Figure 22B).

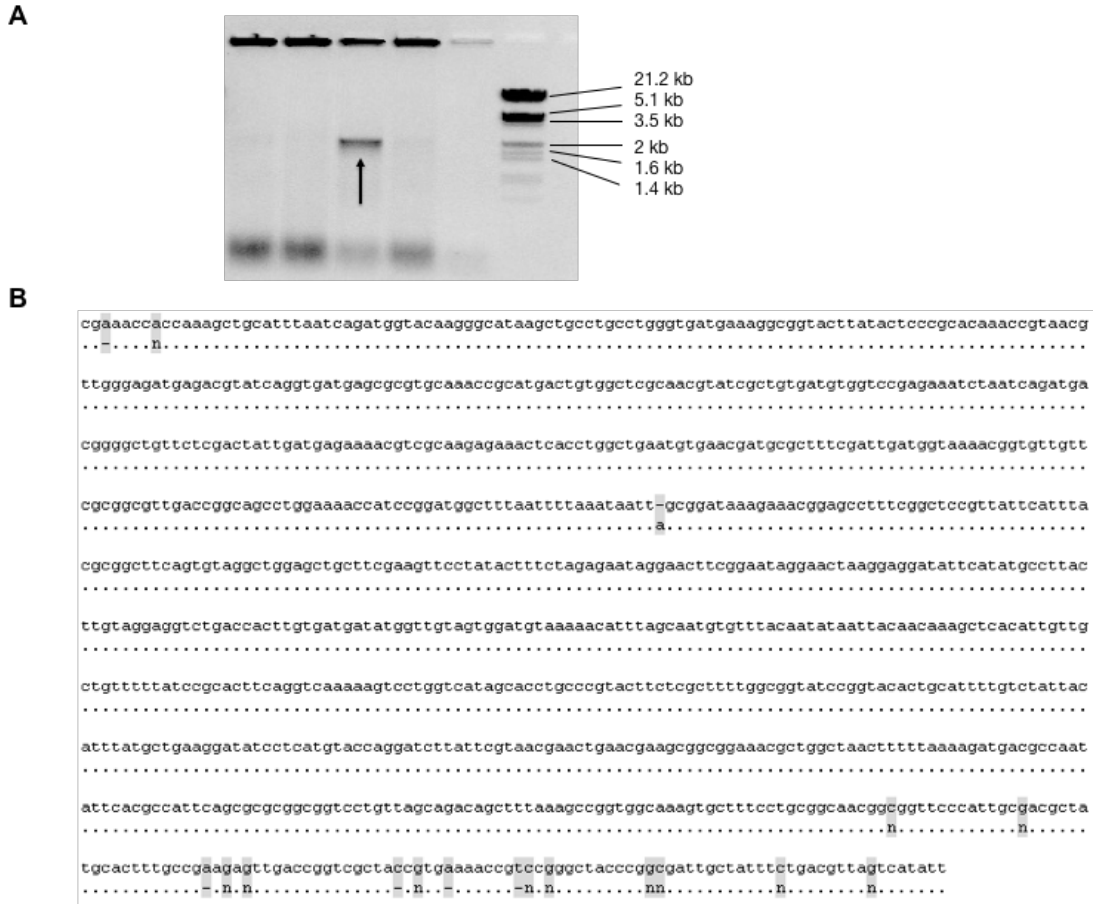

Figure 22. (A) 1% agarose gel where colony 3 shows the insertion of chloramphenicol resistance using Marker III (Roche®). (B) Sequencing of the obtained knockout strain.

4.2.2. Δ*tesA*

To obtain this knockout, pKD4 plasmid was used. This plasmid contained kanamycin resistance to be recombined with the genomic DNA of *E. coli*. As with Δ*fadE*, two sets of primers were designed. The first set was composed by oMEMO4564 and oMEMO4465 primers. These primers were used to create a homologous region while oMEMO4466 and oMEMO4467 were used to check the locus in each step of the process *(full sequence in* 15.4. Primers used in this project*)*. In this case, there were parts of the codifying sequence of *tesA* gene that were also codifying for other genes. Thus, the primers that were designed only annealed in the middle part of the gene so it did not affect to the expression of other genes. The expected length for the resistance cassette was around 2.5 kb while the *tes*A gene was around 750 bp (figure 23A). Colony 11 (arrow) was selected to continue with the procedure by transforming it with pCP20 and remove the resistance cassette. If kanamycin resistance gene was correctly removed, the expected band obtained by PCR colony would be around 600 bp (figure 23B). First colony (arrow) was sent for sequencing confirming that BL21 Star (DE3) Δ*tesA* (Δ*tesA* from now on) strain was successfully obtained (figure 23C).

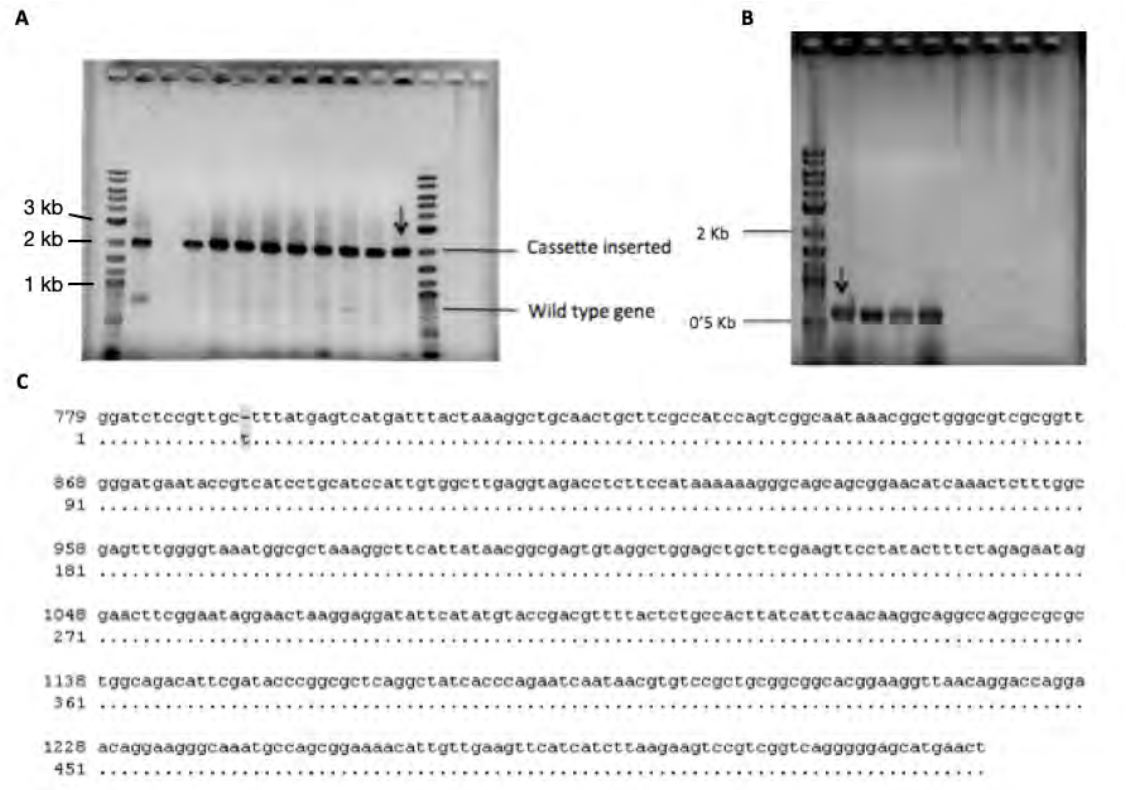

Figure 23. (A) 1% agarose gel using 2 log DNA ladder (NEB) showing the absence of the wild type gene in the genomic DNA of E. coli. Arrow indicates the selected colony to continue with the KO procedure (B) 1% agarose gel using 2 log ladder (NEB) showing the removal of the resistance cassette into the genomic DNA, the arrow indicates the selected colony for sequencing (C) Sequencing of the genomic locus where tesA gene was placed

4.2.3. Δ*tesA* Δ*fadE*

In order to obtain the double knockout, Δ*tesA* strain that was already obtained was transformed with pKD46 to initiate the entire procedure and be able to knockout *fad*E gene. The same primers and linear knockout fragments used to obtain Δ*fadE* were used. The first set of plasmids used to insert the linear DNA with the resistance cassette and remove *fade* were oMEMO4468 and oMEMO4480 (if correct a band of 2 kb should be seen) (Figure 24A) while oMEMO4508 and oMEMO4509 were used to check that the resistance cassette was correctly removed from the genome after using pCP20 plasmid. If the resistance cassette was correctly removed, a band of around 1 kb should be seen.

Colony 8 (Figure 24B arrow) was used for preservation and send for sequencing confirming that the *fad*E was correctly removed from the genome and strain BL21 Star (DE3) Δ*tesA* Δ*fadE* (Δ*fadE* Δ*tesA* from now on) was successfully obtained *(sequencing results of this knockout can be seen in section 15.1.4. BL21 Star (DE3) ΔtesA ΔfadE)*.

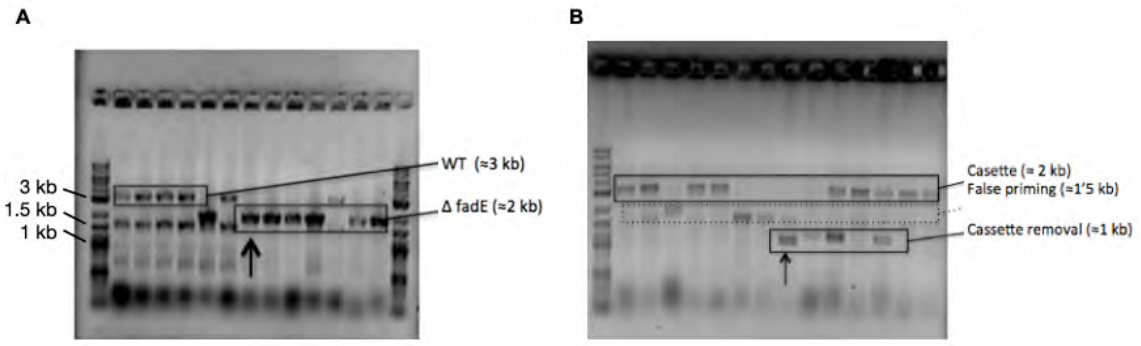

Figure 24. 1% agarose gels using 2-log ladder (NEB) as marker(A) Removal of FadE gene (B) Removal of the resistance cassette

4.3. Strain characterization

The knockout strains, Δ*tesA,* Δ*fadE* and Δ*tesA* Δ*fadE previously obtained were characterized co-expressing mg517-plsC*^{H}, a combination that was assigned as important in the first engineered strain generation (Mora-Buyé et al., 2012).

Cultures of 50 mL (minimal medium with ampicillin and kanamycin) were prepared and grown overnight. OD_{600} were measured every hour to calculate growth rate when cultures were not induced. Grofit software was used to determine growth rate (μ_g).

Table 3. Growth rate in the engineered strains when cells were grown in minimal medium

Strain BL21 Star (DE3)	Final OD	Growth rate
(#1) WT/mg517-plsC^{H}	2.45	0.210 ± 0.010 h^{-1}
(#2) ΔtesA/ mg517-plsC^{H}	2.52	0.260 ± 0.010 h^{-1}
(#3) ΔfadE/mg517-plsC^{H}	2.07	0.227 ± 0.006 h^{-1}
(#4) ΔtesA ΔfadE/ mg517-plsC^{H}	2.74	0.430 ± 0.140 h^{-1}

Table 3 shows there were no differences in the growth despite the genomic modifications. Only the double knockout of Δ*tesA* Δ*fadE* (#4) seems to grow faster. The final OD without induction after 24 hours of growth is above 2.

Moreover, characterization of the growth with and without induction was also performed. Three cultures of ΔtesA/mg517-plsC^{H} were grown in minimal medium for 24 hours. Among these cultures, two were induced and one was used as a control. Since glycolipids are compounds that are accumulated in the membranes, the IPTG induction was performed at the beginning of the exponential phase when new cells were formed. The OD was measured every hour (Figure 25). For the control strain, not induced, the growth started when the culture was initialized. The exponential growth phase started at the 2nd hour and lasted until the fifth hour post initiating the culture with a μ_g of 0.24 h^{-1}. For the induced strains, these cultures were induced when they reached an OD of 0.14. A larger lag phase was observed when compared to the control strain. This lag phase lasted approximately 8 hours and after this, an exponential phase of 6 hours with a μ_g of 0.18 ± 0.01 h^{-1} was observed. This slight decrease in the growth rate could be caused by

the metabolic burden the overexpression of two different proteins might suppose (Figure 25 B).

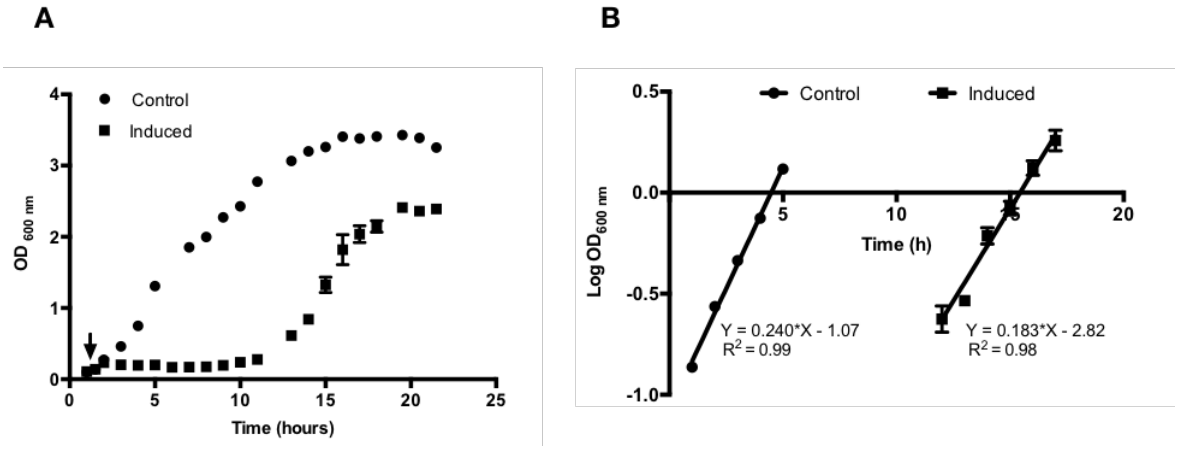

Figure 25. Growth of ΔtesA/mg517-plsC^{H} (A) Growth curve (B) Growth rate calculated

Furthermore, a microscopical analysis of these new engineered strains was performed showing that for all the strains expressing both enzymes (MG517 and PlsC), cells were larger when compared to the control strain. This result suggests that cells that produce GGL not only show changes in the membranes (i.e. larger, elongated in a rod-like shape) due the presence of these compounds but also promote other processes such as hampered separation. (Figure 26).

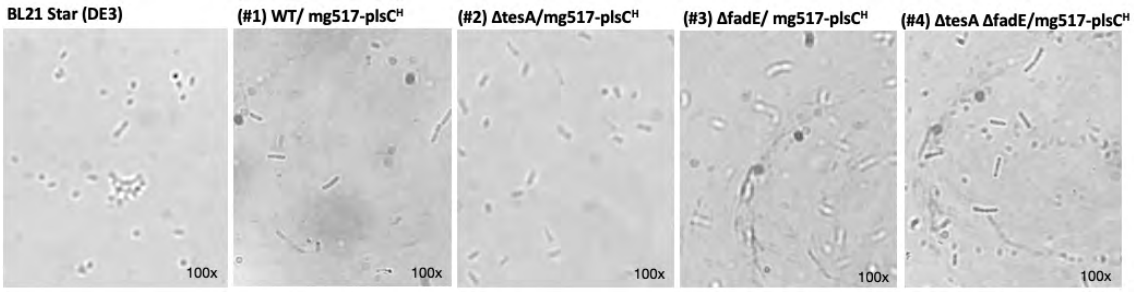

Figure 26. Microscopic analysis of the engineered strains (strains: control, #1, #2, #3 and #4)

4.4. GGL production

The effect of removing *fadE* and *tes*A on GGL production was analyzed in the new engineered strains co-expressing MG517 and PlsC. To quantify the amount of GGL, all the cultures were harvested, rinsed twice with 0.9% NaCl and lyophilized. After this, Folch extraction was performed for each 40 mg of dry cell mass (Folch et al., 1957). The lipid extract was dried with N_2 and GGL production was calculated by analyzing the sugar moieties of this lipid extract and the distribution of glycolipid as commented before *(for more details see section 13.3. GGL analysis)*.

Figure 27 includes the results of this quantification. The WT strain co-expressing MG517 and PlsC (reference strain #1) produced 3.2 nmol/mg of GGL, a higher production than that obtained when MG517 was expressed alone as previously reported by Mora et al (strain #0) (Mora-Buyé et al., 2012). A significant 1.8-fold increase was observed in the Δ*tesA* strain (#2) compared to the WT strain (#1) (p-value 0.02), meaning that acyl donor availability has been significantly

increased by removing the thioesterase activity as a competing reaction (Figure 27). By contrast, Δ*fadE* strain (#3), with blocked aerobic fatty acids ß-oxidation, only gave a slight improvement of GGL production, which indicates that little fatty acid ß-oxidation occurs when using aerobic growth conditions in minimal medium with glucose as carbon source (p-value 0.12). However, the double knockout, Δ*tesA* Δ*fadE* strain (#4), had a lower productivity than the WT, suggesting that knocking out simultaneously the initial steps of two acyl-CoA-utilizing pathways perturbs the regulation of this central metabolic node with a highly detrimental effect on our target GGL productivity (p-value 0.06).

Table 4. Production of GGL in the engineered strains

Strain (genotype)	OD	$[Glc]_T$ (µg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/WT (#1)
			%M	%D	%Tri	%Tetra		
(#0) WT/ mg517	3.6 ± 0.1	0.84	36	31	15	18	2.17 ± 0.50	0.7
(#1) WT/ mg517-plsC^H	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^H	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#3) ΔfadE/ mg517-plsC^H	3.1 ± 0.3	1.51	20	47	20	13	3.72 ± 0.70	1.1
(#4) ΔtesA ΔfadE/ mg517-plsC^H	3.1 ± 0.7	0.92	31	39	8	22	2.31 ± 0.70	0.7

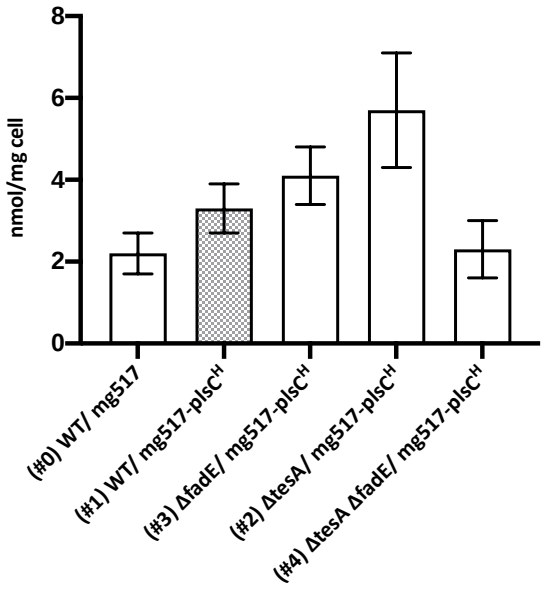

Figure 27. GGL production using MG517 and PlsC^H proteins in the different new engineered strains. Significant increase of the GGL production of strain #2 compared to #1 (p-value 0.023) (T-student analysis with a 95% of confidence, all statistical analyses can be found in annex 15.2 Production summary of the strains).

An example of a TLC and distribution of GGL can be seen in Figure 28. The glycolipid profile is represented in Figure 28B. This distribution suggested that the main GGL present in the highest producer strain is DGDAG, followed by TGDAG and TetraGDAG (Table 4). The glycolipid profile between the different strains is quite similar between them and only slight changes in the MGDAG are observed.

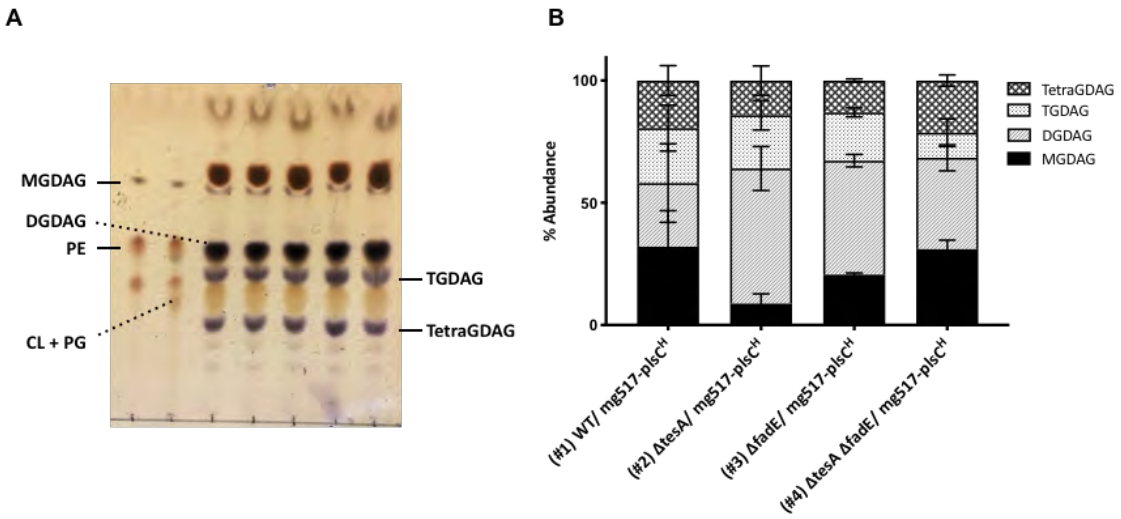

Figure 28. (A) Example of a TLC showing the lipidic fraction obtained from the different strains analyzed (B) Glycolipid profile representing the abundance of each glycolipid

As previously mentioned and in line with what Mora-Buyé reported in 2012, the production of glycoglycerolipids is increased when the glycolipid synthase MG517 is expressed in combination with acyltransferase $PlsC^{H}$ (#1) in comparison to the strain only expressing MG517 (#0). This result also indicates that diacylglycerol is the limiting step in the synthesis of GGL (Mora-Buyé et al., 2012). This bottleneck can be solved by removing the competing reaction catalyzed by TesA (#2). By removing the thioesterase activity it was possible to increase the production of GGL by 1.8-fold.

Interestingly, even though *fad*E (#3) has been widely used in different studies to produce free fatty acids for biofuel production when analyzing the production in these conditions only a 1.1-fold increase was observed compared to reference strain (#1 strain). Finally, the double knockout containing Δ*tesA* Δ*fadE* (#4), reported a decrease in the production when compared to strain #1. This suggested that knocking out simultaneously the initial steps of two Acyl-CoA-utilizing pathways perturbs the regulation of this central metabolic node with a detrimental effect on our target GGL productivity.

STRATEGY 2: INCREASE OF FA AVAILABILITY BY MODULATING TRANSCRIPTION FACTORS

5. Strategy 2: Increase of FA availability by modulating transcription factors

5.1. Strategy basis

Another strategy related with the fatty acid metabolism itself was focused on increasing the synthesis of these compounds. In order to do that, two different approaches were proposed. The first approach was focused on *fabR*, a monounsaturated fatty acid repressor, while the second approach was based on *fadR,* a transcription factor that acts as a global activator for the biosynthesis.

Specifically, the first approach was based on knocking out *fab*R, a gene that is known to repress *fab*A and *fab*B genes and, therefore, the synthesis of monounsaturated fatty acids which is also known to have an impact in the homeostasis of the membrane (Marrakchi et al., 2002; Zhang et al., 2002). FabA activity is restricted to fatty acids of 10 carbons while FabB is responsible for the elongation of the fatty acid product obtained by FabA and is known to be the primary factor determining the cellular levels of unsaturated fatty acids (Fujita et al., 2007). Both genes can be activated by FadR and inhibited by FabR.

On the other hand, *fadR* is a global activator of the biosynthesis of fatty acids and acts as a repressor of many genes and operons involved in the fatty acid degradation. It is a multifunctional dual regulator that is known to promote the expression of the four subunits of acetyl-CoA carboxylase, FabHDG and FabI while repressing the expression of FadE, FadM, FadD, FadI, FadJ, FadL, FadH, FadB and FadA (Fujita et al., 2007; He et al., 2014; Magnuson et al., 1993; Simons et al., 1980). It is also responsible for maximizing the production of unsaturated fatty acids by promoting the expression of FabA and FabB.

Different studies reported that by overexpressing this regulator it was possible to achieve a significant increase in the production of fatty acids (He et al., 2014; X. Zhang et al., 2012). Furthermore, San and Li reported that combining *fadR* overexpression with *fabR* knockout led to a significant increase in the fatty acid titer (San and Li, 2017). Kim et al reported in 2018 that it was possible to produce selectively omega-hydroxy palmitic acid from glucose in *Escherichia coli* by rewiring FadR regulon (Kim et al., 2018).

FadR repression acts by binding to the promoter sites in cells growing without fatty acids inhibiting therefore the transcription of acetyl-CoA carboxylase and other genes related with the biosynthesis while at the same time stops the inhibition of the expression of all the genes related with the ß-oxidation.

In this work, it was proposed to study the effect that these transcription factors may have in the production of glycoglycerolipids. Considering the previous results, and after confirming that DAG was limiting in BL21 Star (DE3) to produce glycoglycerolipids, the study was performed in Δ*tesA* strain, which showed to produce more GGL. On one hand a double knockout (Δ*tesA* Δ*fabR*) was proposed while on the other hand, the overexpression of *fadR* was also proposed to study the effect that both transcription factors may have in the production of GGL.

5.2. Molecular biology related to the strategy

5.2.1. Obtaining the Δ*tesA* Δ*fabR* strain

To obtain this KO Δ*tesA* strain was used as a template to remove *fabR* gene.

In this case, it was necessary to use homologous sequences of 100 bp to achieve the recombination since with the protocol used to obtain the previous KO (around 50-60 bp of homology) the recombination event was not achieved. To increase the specificity of the recombination, larger homologous sequences were used. To design these new regions, it was necessary to use two sets of primers. The strategy to obtain the extended regions can be found in Figure 29. The first set of primers, the ones annealing to pKD4 were named fabR_dats_fwC and fabR_dats_rvD while the ones used to extend the homologous sequence were named comp_C (to complement fabR_dats_fwC) and comp_D, which complemented fabR_dats_rvD *(full sequence in* 15.4. Primers used in this project*)*. By using this linear DNA and promoting the recombination event between this and *fabR* gene, the expected length using reverse_fabR_out and oMEMO2506_SthA_seq_rv primers was around 2.1 kb (Figure 30A).

Figure 29. Strategy used to increase the homologous sequences to recombine lineal DNA

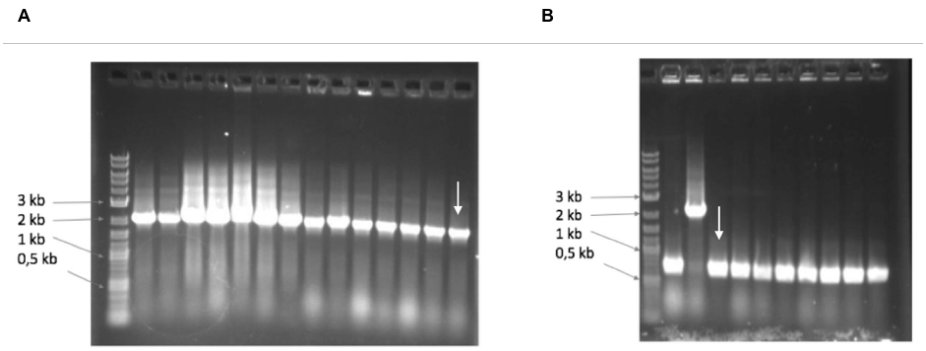

Figure 30. Agarose gels (A) Insertion of the resistance cassette (B) Removal of the resistance gene

Colony 14 (Figure 30A, arrow) was selected to be transformed with pCP20 to remove the resistance cassette. The expected band if the removal was achieved was around 600 bp. As figure 30B shows, only colony 2 did not remove the kanamycin resistance. Colony 3 (Figure 30B,

arrow) was sent for sequencing confirming the obtaining of Δ*tesA* Δ*fabR* double knockout strain *(see section 15.1.5. BL21 Star (DE3) ΔtesA ΔfabR)* (Santiveri, 2018).

5.2.2. FadR plasmid

fadR was amplified from genomic DNA of *E. coli* BL21 Star (DE3) Δ*tesA* using Q5 polymerase (NEB). A pair of primers of at least 20 bp flanked with homologous sequences of the p5T7 vector were designed (fw_fadR_p5T7fadR and rv_ fadR_p5T7fadR) to amplify the *fadR* insert.
The template used to obtain the backbone was galU^{L} and the primers used were fw_p5T7_p5T7fadR and rv_p5T7_p5T7fadR (Table 5).
A gel extraction was performed and after quantifying by Nanodrop®, a Gibson assembly procedure was performed obtaining one colony that was further tested by PCR colony confirming the expected length of *fadR*. The plasmid was confirmed by sequencing.

Table 5. Primers and template used to obtain fadR plasmid

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone (p5T7)	galU^{L}	fw_p5T7_p5T7fadR rv_p5T7_p5T7fadR	6137
fadR	Δ*tesA*	fw_fadR_p5T7fadR rv_ fadR_p5T7fadR	762

5.3. Strain characterization

The growth rate was analyzed and compared to the WT (#1) and Δ*tesA* (#2) strains with the same plasmids combination of mg517-plsC^{H}. No differences were observed by removing *fabR* from the genomic DNA of *E. coli* when compared to the WT and Δ*tesA* strains.

Table 6. Growth rate in the engineered strains when cells were grown in minimal medium

Strain (genotype)	Growth rate
(#1) WT/mg517-plsC^{H}	0.210 ± 0.010 h^{-1}
(#2) ΔtesA/ mg517-plsC^{H}	0.260 ± 0.010 h^{-1}
(#5) ΔtesA ΔfabR/ mg517-plsC^{H}	0.250 ± 0.040 h^{-1}

A microscopic analysis was also performed to check if there were physical differences between this new strain (#5) and #2 (ΔtesA/mg517-plsC^{H}) was performed. As it happened with strains #1 and #2, an elongation of the cells could also be seen (Figure 31).

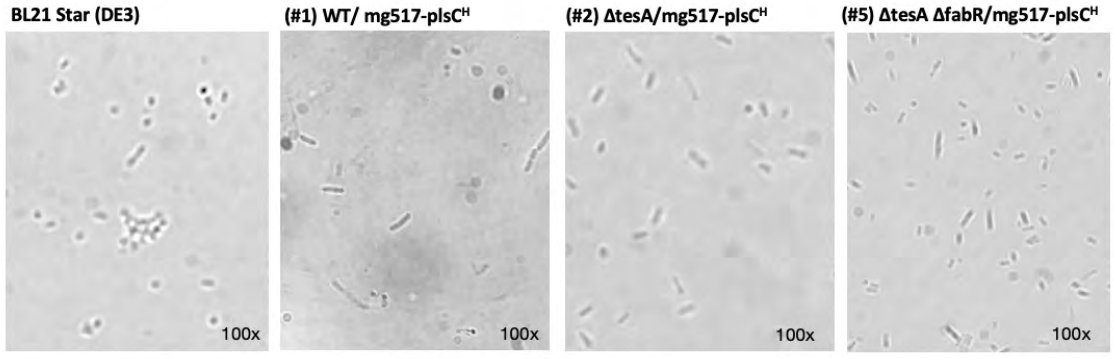

Figure 31. Microscopical analysis of strains: control (BL21 Star (DE3)), #1, #2 and #4

5.4. GGL Production

The GGL production of Δ*tesA* Δ*fabR* strain (#5) was analyzed and compared to the producers strains, WT (#1) and Δ*tesA* (#2) and, while an increase of 1.5 folds (4.84 ± 0.70 nmol/mg cell) in the production was observed when compared to #1 strain (p-value 0.02), no differences were observed when compared to #2 strain (p-value 0.14) (Table 7). All statistical analyses comparing the engineered strains with #1 strain (WT, reference strains) can be found in *Table 48. Overall glycoglycerolipid production of the engineered strains*. The significance of the production has been calculated using a T-student analysis with a 95% of confidence level.

Moreover, it was also analyzed the effect that the overexpression of *fadR* regulator may have, which is responsible for repressing the degradation of fatty acids while also activating its biosynthesis. As Figure 32 and Table 7 show, strains #6 and #7, which overexpress this regulator in Δ*tesA* and Δ*tesA* Δ*fabR* strains, produce a similar amount of GGL compared to strains #2 and #5. Maybe the effect of this regulator could not be seen in this study due the low copy number plasmid used (p5T7).

Table 7. Production of GGL modifying genetic regulators

Strain (genotype)	OD	$[Glc]_T$	%GGL composition				$[GGL]_T$	GGL
		(μg/mg)	%M	%D	%Tri	%Tetra	(nmol/mg)	strain/wt(#1)
(#0) WT/ mg517	3.6 ± 0.1	0.84	36	31	15	18	2.17 ± 0.50	0.7
(#1) WT/ mg517-plsC^H	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^H	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#5) ΔtesA ΔfabR/ mg517-plsC^H	2.1 ± 0.3	2.04	18	50	12	20	4.84 ± 0.70	1.5
(#6) ΔtesA/fadR-mg517-plsC^H	1.7 ± 0.2	2.40	12	60	12	16	5.75 ± 0.50	1.8
(#7) ΔtesA ΔfabR/ fadR-mg517-plsC^H	1.9 ± 0.1	2.29	14	64	17	6	5.86 ± 0.50	1.8

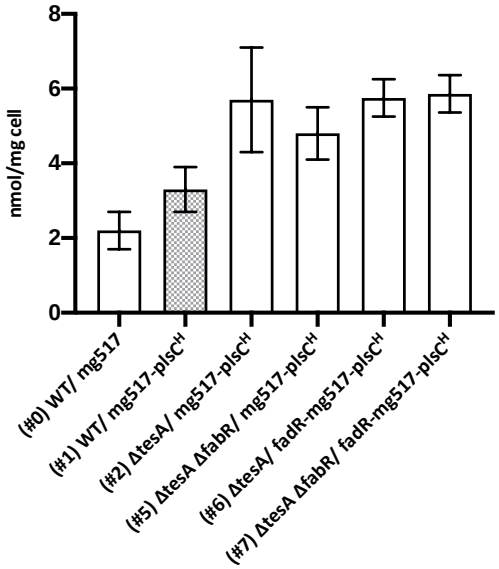

Figure 32. GGL production obtained by the modulation of the transcription factors fadR and fabR. Significant increase in the production in strains #2, #5,#6 and #7 compared to #1 (p-value 0.023, 0.024, 0.015 and 0.012 respectively)

Considering these results, it seems that the strategies proposed to increase fatty acid by modulating transcriptional factors did not have an impact on DAG and GGL levels. Nevertheless, since both genes are related to the synthesis of unsaturated fatty acids and analysis of the lipid profile was performed.

5.5. Lipid profile analysis

The major fatty acids of *E. coli* are palmitic acid (C16:0), 2-hexylcyclopropanoctanoyl (C17:0Δ) and oleic acid (C18:1). It was reported by our group that as the most common fatty acids, GGL contain the same abundance in their composition. To determine if knocking out genes related with the biosynthesis of DAG or unsaturated fatty acids had an effect on the composition of fatty acids, a methyl ester analysis of the fatty acids was performed in all the engineered strains.

To do this analysis, the lipidic fraction of 40 mg of dry cell mass, grown in minimal medium, were extracted by Folch extraction and evaporated with N_2 (Folch et al., 1957). Afterwards, these samples were resuspended with hexane and KOH 2N obtaining the desired methyl esters which latter were used to identify the lipids by Gas Chromatography *(more in detail in 13.5. Lipid analysis)*.

The fatty acid methyl esters of the lipid extracts were identified by comparing the retention times with a standard SUPELCO 37 FAME MIX (Sigma Aldrich®), which contained 37 different fatty acids from C4 to C22. Fatty acids were also confirmed by GC-MS (HRGC 6890N)*(all the chromatograms can be found in section 15.6. Gas chromatography chromatograms and confirmation).*

Figure 33 and table 8 present the percentage of abundance of each methyl ester found in the different knockout strains.

Table 8. Percentage abundance of the fatty acid methyl esters found in the different strains

	C14:0	**C16:0**	**C16:1**	**C17:0Δ**	**C18:0**	**C18:1**	**C19:0Δ**
BL21 Star (DE3)	7.7 ± 2.9	45.2 ± 2.1	11.5 ± 2.7	15.9 ± 1.2	1.2 ± 0.1	13.3 ± 3.0	5.8 ± 2.8
(#1) WT	3.1 ± 0.1	45.9 ± 0.2	4.3 ± 0.1	27.0 ± 0.01	1.2 ± 0.1	8.0 ± 0.1	10.6 ± 0.3
(#3) Δ*fadE*	3.7 ± 0.1	52.0 ± 0.5	4.9 ± 0.1	21.0 ± 0.5	7.7 ± 0.1	6.0 ± 0.2	4.7 ± 0.0
(#2) Δ*tesA*	2.8 ± 0.2	51.9 ± 0.8	7.2 ± 0.4	18.4 ± 1.3	9.5 ± 2.3	7.1 ± 0.7	3.1 ± 0.5
(#4) Δ*tesA* Δ*fadE*	4.3 ± 0.8	51.3 ± 0.6	3.3 ± 0.3	22.6 ± 2.1	9.2 ± 2.0	4.0 ± 0.4	5.2 ± 0.5
(#5) Δ*tesA* Δ*fabR*	2.1 ± 0.3	29.0 ± 1.0	18.1 ± 1.2	15.4 ± 0.1	3.0 ± 0.1	29.7 ± 0.3	2.8 ± 0.4

As it can be seen, the main difference between the lipid profiles take place in double knockout Δ*tesA* Δ*fabR* (#5) where an increase in the abundance of unsaturated fatty acids C16:1 and C18:1 can be observed (Figure 33, black). As a reminder, this gene is responsible for the repression of *fabA* and *fabB* genes, which are known to be responsible for the unsaturation of fatty acids. By knocking this gene out, *fabA* and *fabB* were expected to be more active producing higher titers

of unsaturated fatty acids. The Δ*tesA* Δ*fabR* strain did not report a significant increase in the production of GGL but regarding this data, it would be interesting to use this strain when a different lipidic profile is desired. This result is consistent with previous studies where it was reported that mutants on this gene were able to increase by 2-fold the expression of *fabA* and *fabB* genes, confirming the repression of these by *fabR*. In addition, this same strain showed an increase in the unsaturated fatty acids (Parsons and Rock, 2013; Zhang et al., 2002).

This Δ*tesA* Δ*fabR* strain could be interesting to be used in those assays where it was intended to use low temperatures since unsaturated fatty acids have lower melting temperatures and are able to enhance the fluidity of the membranes.

Regarding to the other engineered strains, only a slightly increase in the presence of C16:0 and C18:0 over C18:1 and C19:0Δ was observed in comparison to the BL21 Star (DE3) strain without co-expressing mg517-plsC^{H} (Figure 33, blue). These results suggested that by knocking out *fad*E and *tes*A no major changes in the lipid profile could be observed.

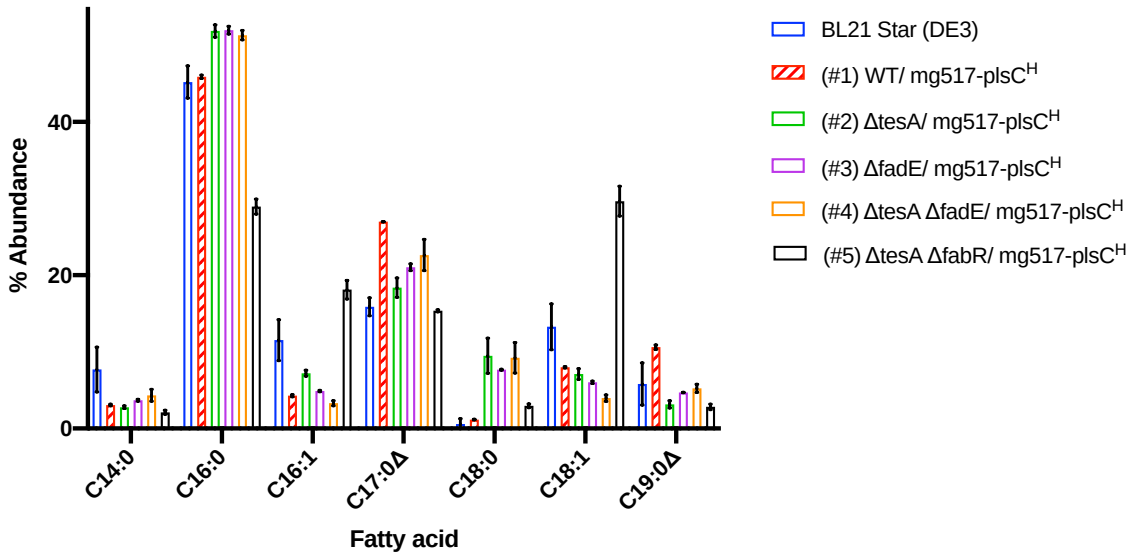

Figure 33. Lipid profile in the new engineered strains

STRATEGY 3: INCREASE THE CONVERSION OF ACYL DONORS TO PA BY OVEREXPRESSING PlsC AND PlsB

6. Strategy 3: Increase the conversion of acyl donors to PA by overexpressing PlsC and PlsB

6.1. Strategy basis

The third strategy proposed in this project was aimed at increasing the PA availability, precursor of DAG, by overexpressing the acyltransferases that are involved in its biosynthesis. Acyltransferases are enzymes placed in the cellular membranes that allow the formation of phospholipids. The biosynthesis starts with glycerol–3–phosphate (G3P) that is converted into lysophosphatidic acid (LPA) by the action of PlsB acyltransferase, an enzyme of 97 KDa that is specific for acylation at position 1 of G3P and can use either acyl-ACP or acyl-coenzyme A (acyl-CoA) thioesters as acyl donors. Fatty acids that are endogenously synthesized are activated using ACP while exogenously added fatty acids are attached to CoA (Green et al., 1981; Lightners et al., 1980; Ray and Cronan, 1987). This protein is encoded by *pls*B, which is an essential gene. After this reaction and being the second step in the phospholipid biosynthesis, PA is formed from LPA by PlsC, responsible for the acylation at *sn*-2 position (Figure 34). As PlsB, this enzyme is able to use both acyl-ACP and acyl-CoA as acyl donors. This acyltransferase is encoded by *pls*C gene, which is also an essential gene. Both PlsB and PlsC acyltransferases have been studied and used for *in vitro* studies showing that they maintained their activity and also were able to produce phospholipids when reconstituted in liposomes (Scott et al., 2016). Several reviews had also reported the role of PlsB and PlsC in the biosynthesis of phospholipids in bacteria (Parsons and Rock, 2013; Yao and Rock, 2013; Zhang and Rock, 2008).

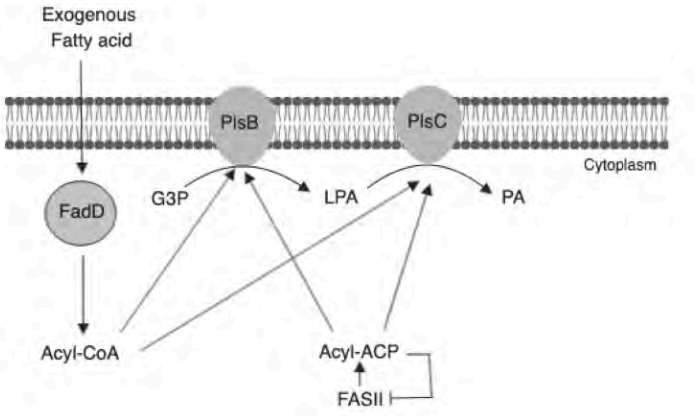

Figure 34. PA synthesis pathway. PlsB forms LPA from G3P by acylating it. Later this LPA is transformed into PA, precursor of phospholipids and DAG by a new acylation catalyzed by PlsC

6.2. Molecular biology related to the strategy

6.2.1. Study basis

To study the possible effect that acyltransferases may have in the production of GGL, PlsB and PlsC acyltransferases were subcloned into high and low copy number plasmids. Since all the studies were performed in the already metabolic engineered strains, the effect of using low copy

number plasmids was studied in order to know if their use was enough to push the production.

6.2.2. Low copy number plasmids

Both acyltransferases were subcloned into low copy number plasmids (p10T7) provided by Prof. Marjan de Mey's group at the University of Ghent. This plasmid contains p15a origin of replication allowing approximately the presence of 15 copies per cell.

6.2.2.1. PlsC^{L}

This subcloning was performed based on the work of Gibson et al (Gibson et al., 2009). *plsC* gene was obtained from pRSF-plsC (plsC^{H} from now on) plasmid that was previously constructed by our group (Mora-Buyé et al., 2012).

To get the required fragments as lineal fragments and with its overlapping sequences it was necessary to run two different PCRs. Firstly, primers oMEMO4562 and oMEMO4563 were used to linearize the backbone p10T7 while oMEMO4560 and oMEMO4561 primers were used to extract *plsC* from plsC^{H} plasmid (Table 9).

The expected lengths were 780 bp for PlsC acyltransferase and 3.6 kb for p10T7 backbone. Figure 35A shows the agarose gel where samples were loaded showing that both fragments had the desired length.

Table 9. Primers and template used to obtain the fragments for plsC^{L}

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone (p10T7)	p10T7	oMEMO4562 oMEMO4563	3600
plsC	plsC^{H}	oMEMO4560 oMEMO4661	780

An equimolar amount of PlsC and p10T7 was used for the Gibson assembly and after the reaction procedure, cells were transformed by electroporation and plated onto chloramphenicol plates. Grown colonies were checked by PCR colony where the expected length if PlsC was subcloned into the plasmid would be around 930 bp using oMEMO1568 and oMEMO3478 primers. Figure 35B shows the samples from the PCR colony. Colony 1, which had the desired length, was sent for sequencing confirming the correct subcloning.

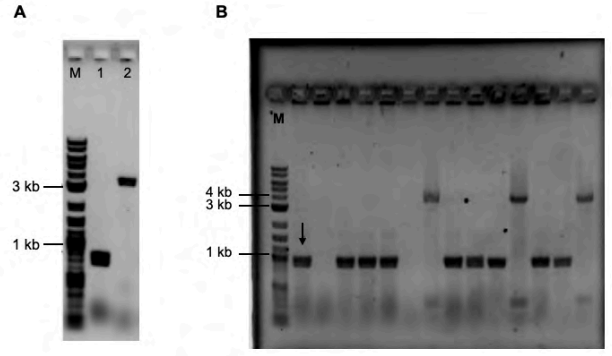

Figure 35. (A) Agarose gel of the fragments to do Gibson Assembly. Lane M is 2-log ladder (NEB), 1 is PlsC (insert) and 2 p10T7 (backbone) (B) Agarose gel of the PCR colony performed. Colony 1 (arrow) indicates the plasmid confirmed by sequencing

6.2.2.2. plsB^{L}

plsB gene was extracted from genomic DNA of *Escherichia coli* K-12 MG1655. The backbone was amplified from the p10T7 plasmid. Two PCRs were performed to obtain both linearized fragments. To linearize the backbone oMEMO4554 and oMEMO4555 were used while to obtain *plsB* gene from genomic DNA the primers were oMEMO4552 and oMEMO4553 (Table 10).

Table 10. Primers and template used to obtain the fragments for plsB^{L}

Fragment	Template used	Primers	Length of the fragment (bp)
p10T7	plsC^{L}	oMEMO4554	3600
		oMEMO4555	
plsB	*E. coli* K-12 MG1655	oMEMO4552	2500
		oMEMO4553	

The expected fragments were obtained and an equimolar amount of each fragment was used in a Gibson assembly reaction and transformed into DH5α electrocompetent cells (Figure 36). After the incubation time, these cells were plated onto chloramphenicol plates and grown overnight. The colonies obtained were checked by PCR colony using oMEMO1560 and oMEMO4383 where the expected fragment was of 2.9 kb. Only one of the colonies contained the desired construct confirmed by sequencing. Plasmid DNA was extracted from this colony and sent for sequencing reporting the correct construct.

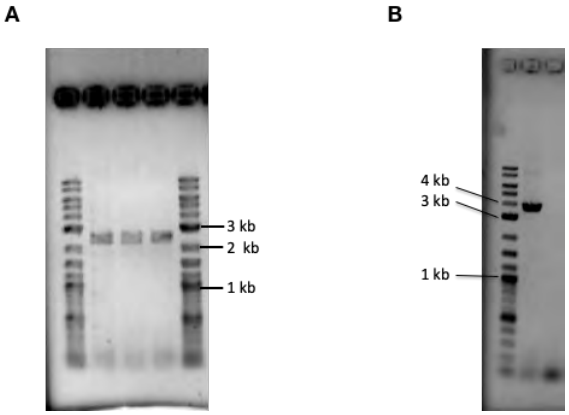

Figure 36. 1% agarose gels using 2-log ladder (NEB). (A) linearized plsB, three different replicates (B) Vector linearized, p10T7

6.2.2.3. plsC^{L}·plsB^{L}

The effect of overexpressing both acyltransferases was also explored by using the same plasmid to overexpress PlsB and PlsC as a polycistronic plasmid.

To get the backbone and *plsC*, plsC^{L} was used as a template for the backbone while to obtain *plsB* the template used was plsB^{L}. Between both genes, an intergenic region, which contained a restriction enzyme site (EcoRI, G'AATTC) and RBS was added (Figure 37). To do so, one of the primers of each set, contained the desired region into the 5' site. This new added part was used as an overlapping fragment to perform a Gibson assembly procedure. Table 11 contains a

summary of the primers and templates used to obtain this plasmid.

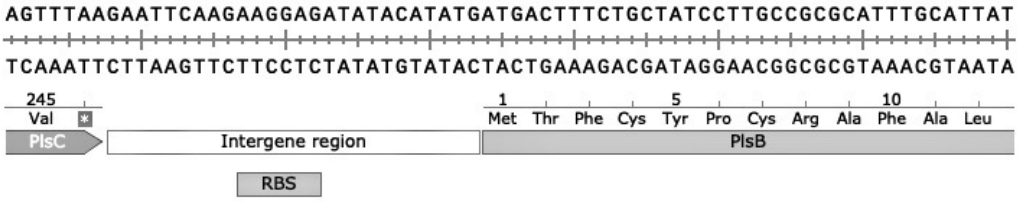

Figure 37. Design of the region between PlsC and PlsB proteins

Table 11. Summary of the required templates and primers used to obtain the desired fragments

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone (p10T7)	plsC^{L}	oMEMO4554 oMEMO4684	4421
plsB	plsB^{L}	oMEMO4553 oMEMO4685	2527

The agarose gel (Figure 38A) showed that in the backbone case there was a lot of false priming requiring a gel extraction procedure. For PlsB, in the first PCR, the gene was not correctly amplified, so a repetition of this procedure was performed. Figure 38B contains the purified fragment of backbone obtained after gel extraction (B) and PlsB bands obtained from the amplification of plsB^{L} plasmid (*plsB*).

An equimolar amount of each fragment was used for the Gibson assembly reaction and after transforming DH5α electrocompetent cells, cells were grown in LB medium supplemented with chloramphenicol.

After checking by PCR colony (T7 promoter and terminator primers, expected fragment of 3.4 kb), only colony 8 had the desired construct that was further confirmed by sequencing.

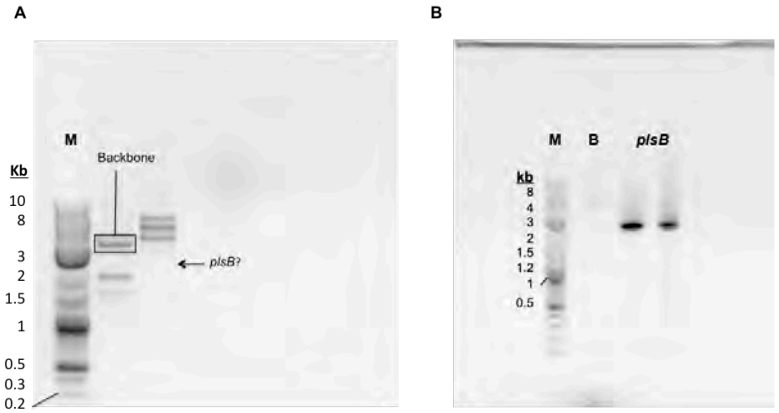

Figure 38. (A) Agarose gel of the first PCR performed to obtain the linear fragments (B) Second PCR to confirm the correct extraction by gel purification of the backbone and PlsB

6.2.3. High copy number plasmids

Both acyltransferases were subcloned into pRSF-1b plasmid. This vector contains a RSF1030 origin of replication that allows for a replication of over 100 copies per cell (Conrad et al., 1979). PlsC was already subcloned into this plasmid in previous studies by our group (Mora-Buyé et al.,

2012) where *plsC* was extracted from *E. coli* JM109. As indicated before, *plsB* gene was extracted from the genome of *Escherichia coli* MG1655.

6.2.3.1. plsB^{H}

To obtain this plasmid, a restriction-ligation approach was followed using EcoRI-HF and NcoI-HF restriction enzymes. To incorporate these restriction sites, it was required to perform a PCR where at 5' of the primers these sequences were added. Figure 39 shows the primers designed and annealed into the sequence of pRSF – plsB.

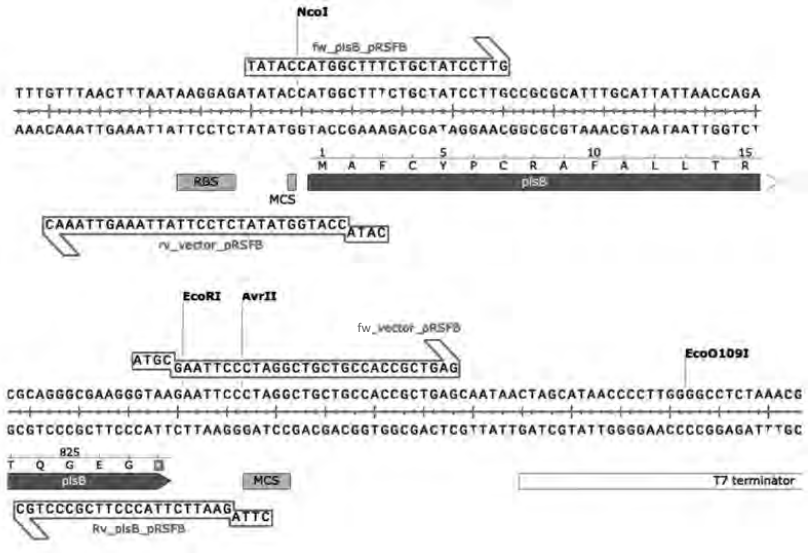

Figure 39. Primer design to subclone plsB into pRSF-1b plasmid

Table 12. Primers and templates used to obtain plsB^{H}

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone (pRSF-1b)	plsC^{H}	Fw_vector_pRSFB Rv_vector_pRSFB	3.5
plsB	*E. coli K-12 MG1655*	Fw_plsB_pRSFB Rv_plsB_pRSFB	2.5

The product obtained by PCR was purified and was digested for three hours at 37ºC by DpnI, so the template used for the amplification of the DNA could be eliminated. Obtained fragments were digested by EcoRI-HF and NcoI-HF restriction enzymes using Cutsmart buffer at 37ºC for three hours.

After inactivating the enzymes, samples were purified and quantified using Qubit HS (life technologies®). To ligate the obtained fragments, T4 ligase was used and ratios 1:3 and 1:6 (vector:insert) were used.

DH5α cells were transformed by heat shock with these products and grown overnight at 37ºC with LB kanamycin plates.

PCR colony using the primers of *plsB* mentioned above were performed in 13 colonies. The expected length of the fragment, if PlsB was correctly subcloned into pRSF – 1b plasmid was around 2.5 kb. As it can be seen in Figure 40, except for colonies 7, 8 and 14, all the tested

colonies seemed to contain the expected fragment. Colonies 4, 13 and 15 were sent for sequencing and finally, plasmid extracted from colony 13 was selected to be used in further studies.

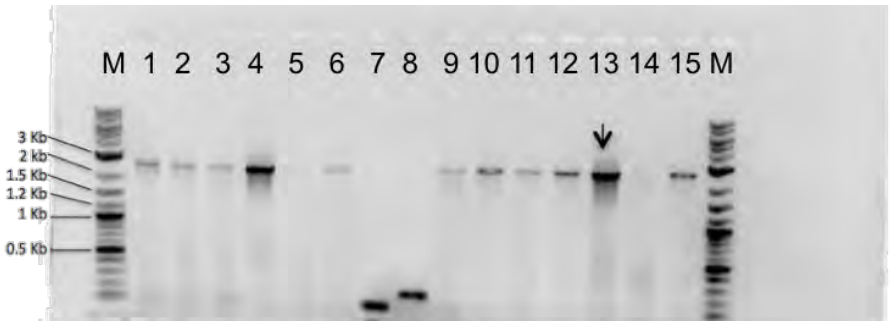

Figure 40. 1% agarose gel of PCR colony ($plsB^H$). If it correctly subcloned, the length of the fragment would be around 2.5 kb. Colony 13 (arrow) was selected to use after confirming the subcloning by sequencing. (M) 1 Kb plus ladder (NEB)

6.2.3.2. $plsC^H$ ·$plsB^H$

To obtain this plasmid a CPEC approach was followed (Quan and Tian, 2014). The template used as insert, plsC – intergenic region – plsB, was obtained from $plsC^L$·$plsB^L$. Primers used can be found in Table 13.

Table 13. Primers and templates used to obtain $plsC^H$-$plsB^H$

Fragment	Template used	Primers	Length of the fragment (kb)
Backbone (pRSF-1b)	$plsC^H$	fw_vector_Gibson CB rv_vector_Gibson CB	3.5
plsC-intergenic-*plsB*	$plsC^L$·$plsB^L$	fw_CB_Gibson rv_CB_Gibson	3.2

Obtained samples were purified, quantified and used in three different CPEC conditions where the main difference between them were the ratio of vector and insert (1:1 or 1:3) and the total amount of DNA (100 ng or 200 ng). In this sense, the first condition was based on 100 ng of total DNA with a ratio 1:1 (as recommended); second condition 200 ng of DNA with the ratio 1:1 and third condition contained 100 ng of DNA with a ratio 1:3.

Figure 41B shows the PCR colonies using T7 primers where it can be seen that almost all the fragments obtained did not correspond to the expected length (3.2 kb). Colonies 3, 11 and 18 seemed to have the desired length, were sent for sequencing. The sequencing results showed an insertion of 20 aminoacids in N-terminal of the protein. These 20 extra aminoacids correspond to the 20 first aminoacids of *plsB*. The active site of *plsB* is found in H(X)4D motif between 306 and 311 aminoacids which is far from this insertion. Furthermore, it was reported an autocatalysis event after the first methionine to provide the mature form of the protein. Considering these two facts we concluded that the obtained construct could be used without interference (Lewin et al., 1999; Yao and Rock, 2013).

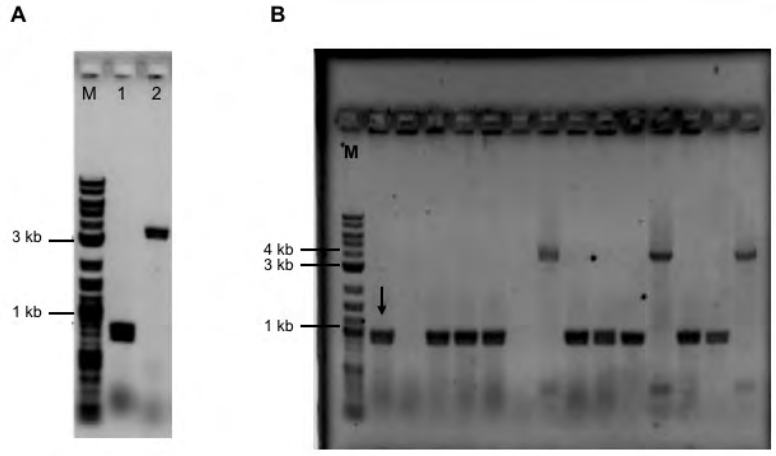

Figure 41. (A) Agarose gel containing the required fragments for the subcloning, M stands for 1 Kb plus ladder (NEB), 1 for plsC and 2 for pRSF-1b (B) Result of the PCR colony where colonies 3, 11 and 18 were sent for sequencing. Arrow indicates a negative control

6.3. Study of the GGL production using different acyltransferases

As previously studied in strategies 1 and 2, ΔtesA/mg517-plsC^{H} was the strain producing higher titer of GGL. Thus, this strain was used to study the effect of the different acyltransferases may have in the production of PA and DAG. To do so, *plsB* and *plsC* were subcloned in high and low copy number plasmids and were overexpressed in Δ*tesA* strain alone or as a polycistronic plasmid.

Table 14. Production of GGL in nmol/mg cell of engineered ΔtesA strain expressing MG517 along with PlsC, PlsB or PlsC·PlsB

Strain (genotype)	OD	$[Glc]_T$	%GGL composition				$[GGL]_T$	GGL
		(μg/mg)	%M	%D	%Tri	%Tetra	(nmol/mg)	strain/wt(#1)
(#1) WT/ mg517-plsC^{H}	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^{H}	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#9 ΔtesA/ mg517-plsC^{L}	2.7 ± 0.3	0.92	16	40	31	14	2.11 ± 0.50	0.6
(#10) ΔtesA/ mg517-plsB^{L}	2.0 ± 0.6	1.14	16	47	24	13	2.72 ± 0.70	0.8
(#11) ΔtesA/ mg517-plsC^{L}·plsB^{L}	1.5 ± 0.3	2.30	11	50	29	11	5.35 ± 0.50	1.6
(#12) ΔtesA/ mg517-plsB^{H}	1.7 ± 0.2	2.06	17	54	17	12	5.11 ± 1.20	1.6
(#13) ΔtesA/ mg517-plsC^{H}·plsB^{H}	1.9 ± 0.3	2.72	4	45	40	10	5.95 ± 0.60	1.8

As it can be seen, the production is higher for all the strains using high copy number plasmids than the strains using the low copy ones. This is especially important in those strains expressing only one acyltransferase. Comparing strains #2 and #10, where PlsC was expressed in high (#2) and low (#10) copy number plasmids, an almost 2.7-fold increase in the production was observed when using high copy number plasmids (p-value 0.01). Regarding to PlsB, comparing the production in low (#10) and high (#12) copy number plasmids, an increase of 1.9-fold of GGL production was observed using high copy number plasmid (p-value 0.05). This suggests that by using high copy number plasmids the amount of DAG produced seems to be more available to produce GGL. The same production of GGL can be obtained by co-expressing both

acyltransferases in both low and high copy number plasmids (p-value 0.17) (Figure 42 and Table 14).

Furthermore, to study the impact of using one acyltransferase or the other, the glycolipid profile was analyzed to detect the possible effect that it may have in the engineered strains (Figure 42B). In all cases, the major product produced was DGDAG followed by TGDAG.

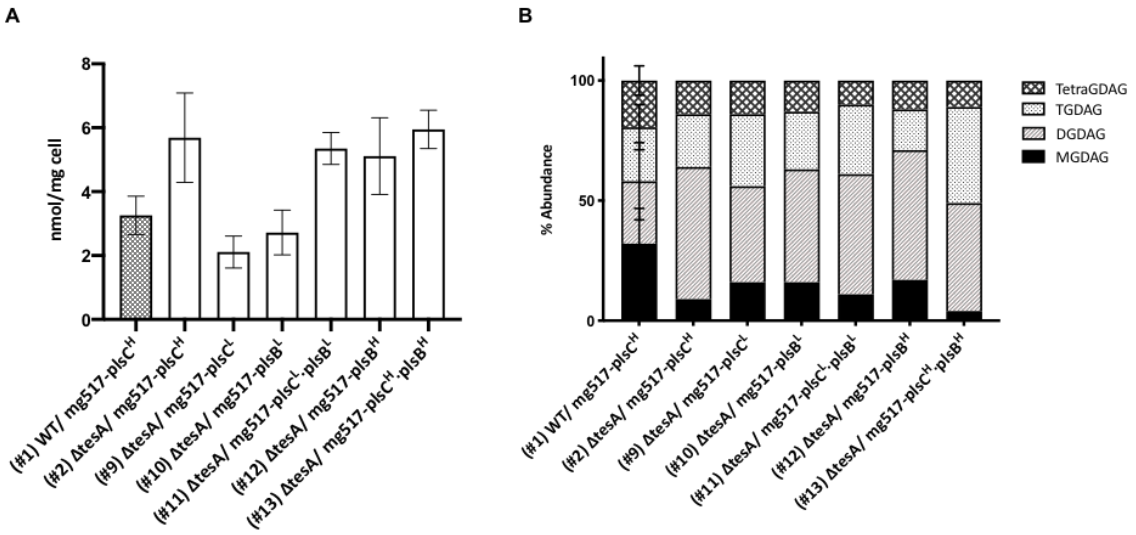

Figure 42. Production in BL21 Star (DE3) ΔtesA using MG517 and different acyltransferase in low (p10T7) and high (pRSF-1b) copy number plasmids. Significant increase of the GGL compared to WT strain (#1) in strains #2, #9, #11,#12 and #13 (p-value 0.023, 0.029, 0.019, 0.038 and 0.011 respectively)

6.4. Study of the acyltransferase enzymatic activity

To test if the enzymes were evenly expressed, the acyltransferase enzymatic activity was also determined in these engineered strains *(more details in 13.4.2. Acyltransferase activity)*. The enzymatic activity was measured *ex-vivo* using cell extract from cultures grown overnight in minimal medium and induced at the beginning of the exponential phase. The reaction was based on the availability of HsCoA to react with DTNB providing TNB. This reaction causes a change in the color of DTNB that can be measured by a microplate reader (Figure 43). Since PlsB and PlsC are natural enzymes of *E. coli*, the baseline expression of both enzymes was monitored (genomic) to be compared with the enzymatic activity achieved when the enzymes were overexpressed into low and high copy number plasmids. When PlsB was analyzed using the polycistronic plasmids the total activity of both enzymes was detected since PlsB is responsible for the synthesis of LPA which is the substrate of PlsC.

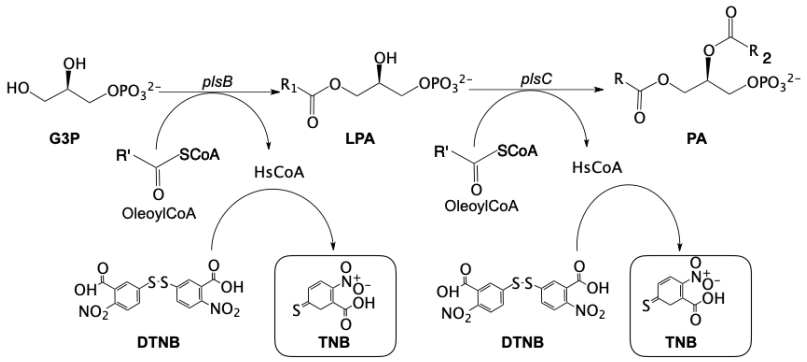

Figure 43. DTNB coupled assay

Results showed that when these acyltransferases were overexpressed its activity was increased regarding the genomic one (Table 15).

Table 15. Acyltransferase activity of the ΔtesA strains co-expressing MG517 in combination with PlsB, PlsC alone or together

		PlsC (mM · min^{-1} · mg protein^{-1})	**PlsB (mM · min^{-1} · mg protein^{-1})**
Genomic		0.04	0.5 ± 0.1
Low copy number	Alone	14.6 ± 1.8	1.2 ± 0.2
	Polycistronic	14.7 ± 0.7	2.4 ± 0.5
High copy number	Alone	10.9 ± 1.0	14.8 ± 2.2
	Polycistronic	12.3 ± 1.7	10.9 ± 1.5

The basal activity for PlsC acyltransferase was 0.04 mM · min^{-1} · mg protein^{-1} whereas it was around 15 mM·min^{-1}·mg protein^{-1}, when using plsC^{L}, and 11 mM·min^{-1}·mg protein^{-1} when plsC^{H} was used. This result supposed a mean increase of around 320 folds when comparing the activity of this enzyme to the genomic one. Furthermore, when both acyltransferases were co-expressed in a polycistronic plasmid the activity of PlsC was also around 15 mM·min^{-1}·mg protein^{-1} when plsC^{L}·plsB^{L} was used, supposing an increase of almost 370 folds, while an enzymatic activity of 12 mM·min^{-1}·mg protein^{-1} (around 310 folds increase) for plsC^{H}·plsB^{H}.

Regarding to PlsB activity, acyltransferase responsible for the synthesis of LPA, the basal one determined was 0.5 mM·min^{-1}·mg protein^{-1} whereas it was 1 mM·min^{-1}·mg protein^{-1} when using plsB^{L} and 15 mM·min^{-1}·mg protein^{-1} when plsB^{H} was used. This represented an increase of 2.5 (plsB^{L}) and 28-folds (plsB^{H}) in comparison to the basal activity. Analyzing the same enzymatic activity when the PlsB was expressed along with PlsC in the same plasmid, the activity found was 2 mM·min^{-1}·mg protein^{-1} and 11 mM·min^{-1}·mg protein^{-1} for plsC^{L}·plsB^{L} and plsC^{H}·plsB^{H} respectively. This supposed an increase of 4 folds when plsC^{L}·plsB^{L} was used and 20-folds when plsC^{H}·plsB^{H} plasmid was expressed. In this case, it seems that the overexpression with a high copy number plasmid is important in the activity of the enzyme.

Relating these results with the production of GGL, when PlsB overexpression was studied, the highest production and activity were achieved when high copy number plasmids were used. In PlsC overexpression, a higher production was observed when high copy number plasmids were used while the activity was similar between using one plasmid or the other. For the

overexpression of both enzymes, both the activity and production were similar. Only a decrease in the activity of PlsB in low copy number, compared to the high copy number plasmid, was observed (Figure 44).

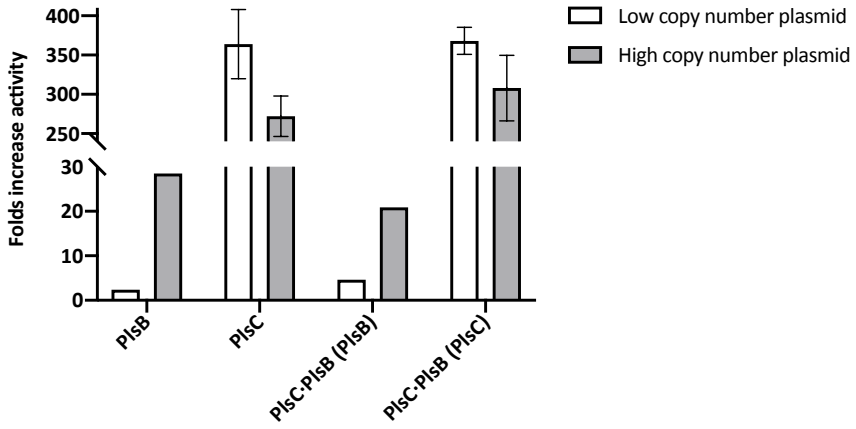

Figure 44. Folds increase of the enzymatic activity of each acyltransferase when using high and low copy number plasmids

Considering all the results, the overexpression of acyltransferases is a key point to increase GGL production. This suggest that PA availability from acyl donor and G3P is increased and has an effect on DAG when PlsC and PlsB enzymes are overexpressed, especially PlsC. The strains selected to continue the studies were ΔtesA/mg517-plsC^{H} (#2) and ΔtesA/mg517-plsC^{H}·plsB^{H} (#13) which are the most important GGL producers by now.

STRATEGY 4: INCREASE THE PRODUCTION OF PHOSPHATIDIC ACID FROM PHOSPHOLIPIDS

7. Strategy 4: Increase the production of phosphatidic acid from phospholipids

7.1. Strategy basis

As far as we know, there are no strategies that modulate the biosynthesis of phospholipids to increase the availability of PA. Following the metabolic pathway of phosphatidic acid in *E. coli* that ends with the production of phospholipids, different genes were considered. The first one was *cdh*, which is responsible for the production of CDP-diacylglycerol diphosphatase, an enzyme responsible to catalyze the reaction of CDP-DAG to PA. By overexpressing this gene, it was expected to decrease the availability of CDP-diacylglycerol while increasing phosphatidic acid to later be used to produce DAG (Figure 45).

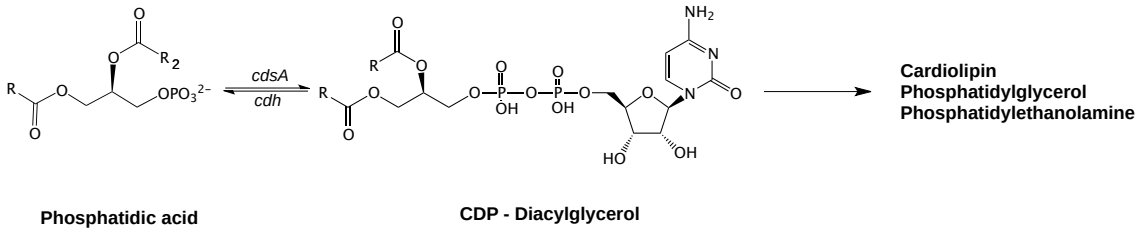

Figure 45. Phospholipid synthesis in E. coli

On the other hand, phosphatidylglycerophosphatase B (PgpB), along with phosphatidylglycerophosphatase A and C (PgpA and PgpC) are responsible for catalyzing the dephosphorylation of phosphatidylglycerol phosphate (PGP) to phosphatidylglycerol (PG), an essential phospholipid of *Escherichia coli* membranes. Whereas PgpA and PgpC seem to be specific for PGP, PgpB is a multifunctional enzyme that is also active in PA and undecaprenyl diphosphate (Dillon et al., 1996; Ghachi et al., 2005) (Figure 46). In this project, it was proposed to pull PA towards DAG by overexpressing the phosphatidic acid phosphatase B (PgpB) since it might have a positive impact on the DAG level or availability for GGL synthesis. This enzyme reported a Km value, using 1,2,-dioleoyl-*sn*-glycero-3-phosphate as substrate, of 0.3 mM (Fan et al., 2014). PgpB is also involved in the last step of phosphatidylglycerol biosynthesis, which is a competing reaction for our purpose (Wikstrom et al., 2009). To overcome this and considering the enzymatic Km, it was proposed to fuse this protein to acyltransferase PlsC so the product formed by this last enzyme, PA, could be used by PgpB to form DAG. By fusing these two enzymes it would be possible to increase the efficiency of PgpB to catalyze the formation of PA to DAG which could later be used in combination with UDP-Glc to produce GGL.

The fusion protein was designed by removing the TAA stop codon of the *pls*C gene, adding a sequence encoding the $(PT)_7P$ linker (Kavoosi et al., 2007) and the sequence of *pgp*B (with the ATG start codon).

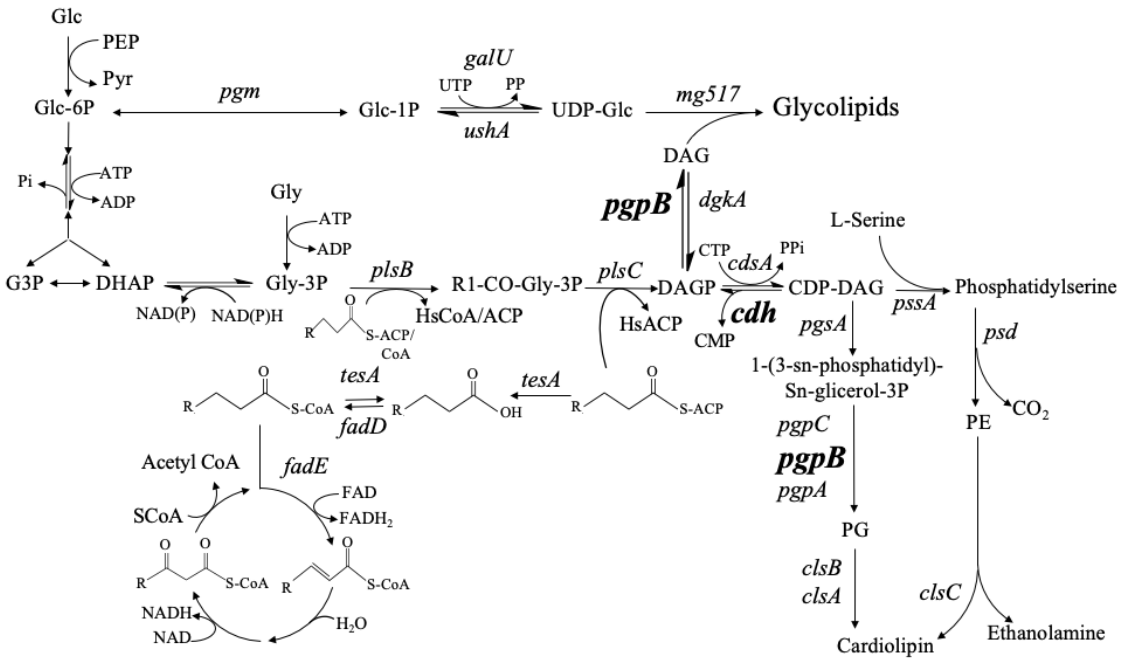

Figure 46. Reactions where pgpB and cdh are involved in the phospholipid biosynthesis metabolism

7.2. Molecular biology related to the strategy

7.2.1. mg517·cdh

In this project, we hypothesized that by overexpressing CDH there might be a decrease in the general production of phospholipids whereas an increase of DAG levels could be achieved at the same time. Thus, CDH was subcloned into the same plasmid as MG517 in order to reduce the number of plasmids per cell and try not to overstress the cells.

The design of this plasmid consisted in, after the full sequence of *mg517,* a spacer of 50 nucleotides and an RBS and, finally, the sequence of *cdh*.

cdh was extracted from BL21 Star (DE3) Δ*tesA* genomic DNA while for the backbone, the template used was mg517 plasmid. In Table 16 are detailed the template and primers used to obtain the fragments for the CPEC procedure.

Table 16. Primers and templates used to obtain mg517·cdh plasmid

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone	mg517	fw_pET44b_polycistronic rv_ pET44b_polycistronic	6400
plsC-intergenic-*plsB*	Δ*tesA* genome	fw_cdh_polycistronic rv_ cdh_polycistronic	809

A CPEC approximation was followed using 100 and 200 ng of backbone and equimolar amount of the insert using a ratio of 1:1 between backbone and the insert.

The obtained mixtures were used to electroporate DH5α cells and grown overnight into LB ampicillin plates at 37ºC.

Several colonies were obtained and tested by PCR colony. Only 4 plasmids obtained from these colonies were sent for sequencing reporting that all contained *cdh* gene.

7.2.2. plsCxpgpB^{H}

The fusion of both proteins was designed by removing TAA stop codon of PlsC, adding a linker followed by the sequence of *pgpB* (Figure 47). Our group reported that the linker proposed by Kavoosi et al, rich in prolines and threonine ($(PT)_7P$), presented high stability in front of *E. coli* hydrolases and allowed a correct folding for fused proteins due its rigid structure (Codera et al., 2015; Kavoosi et al., 2007). Considering this, this linker was selected to fuse PgpB to PlsC acyltransferase.

To obtain this plasmid, four primers were designed in order to create this protein by CPEC assembly. The backbone used to fuse the both enzymes was plsC^{H}. PgpB was extracted from the genomic DNA of *E. coli* BL21 Star (DE3) Δ*tesA* strain. Table 17 shows the design of the intergenic region used to fuse both proteins.

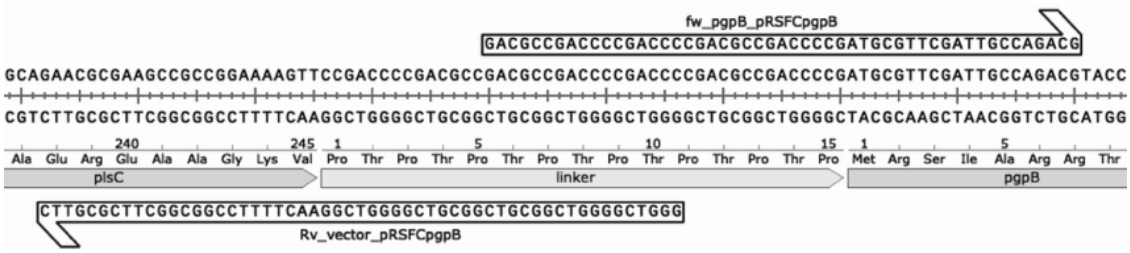

Figure 47. Design of the primers and linker zone to obtain PlsCxPgpB fusion protein

Table 17. Primers and templates used to obtain the desired fragments for plsCxpgpB^{H}

Fragment	Template used	Primers	Length of the fragment (bp)
plsC^{H}	plsC^{H}	fw_vector_pRSFCpgpB rv_ vector_pRSFCpgpB	4240
pgpB	Δ*tesA* genome	fw_pgpB_pRSFCpgpB rv_ pgpB_pRSFCpgpB	817

The obtained fragments from the DNA amplification previously described can be seen in Figure 48. After purifying these fragments, CPEC assembly was performed using 100, 200 and 300 ng of backbone and the equimolar amount of PgpB using 1:1 ratio. The new DNA was used to transform DH5α cells by electroporation. These cells were grown in kanamycin plates overnight at 37ºC obtaining several colonies. A PCR colony of these colonies was performed and 5 of them were selected for sequencing. One contained the desired construct and was used for further studies.

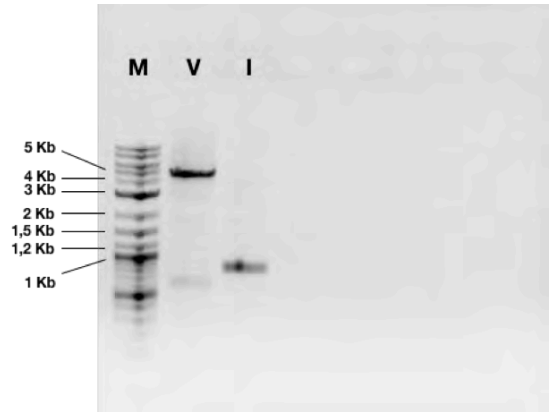

Figure 48. Agarose gel containing the fragments required for the subcloning. V stands for backbone pRSF-plsC while I stands for insert (PgpB) (M) 1 kb plus ladder (NEB)

7.3. Testing the enzymatic activities

7.3.1. CDH enzymatic activity

CDP-diacylglycerol diphosphatase is responsible for catalyzing the conversion of CDP-diacylglycerol to phosphatidic acid as Figure 49 shows.

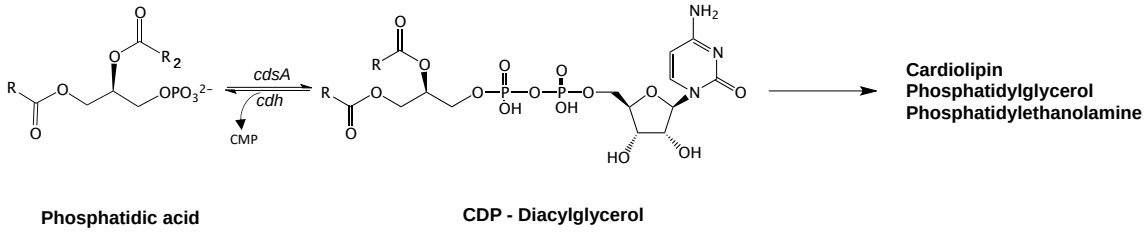

Figure 49. CDP-diacylglycerol synthesis reaction

An assay to detect the increase in the activity of CDH versus the genomic activity when this enzyme was overexpressed was performed. To do so, a coupled assay was designed using alkaline phosphatase (AP) to release phosphates groups, able to react with malachite green reagent. CDH is known to catalyze the following reaction:

Reaction 1:

$$\text{CDP-DAG} + H_20 \rightarrow \text{1,2-diacyl-}sn\text{-glycerol-3-phosphate} + \text{CMP} + 2\ H^+$$

This released CMP can be used in the following reaction (Reaction 2) catalyzed by AP obtaining a phosphate group (PO_3^{2-}), which is able to react with malachite green and be monitored.

Reaction 2:

$$\text{CMP} + H_2O \rightarrow \text{Cytidine} + PO_3^{2-}$$

The main consideration of this coupled assay was that the CDH reaction needs to be the limiting reaction and AP in excess. In addition, alkaline phosphatase is responsible for the catalysis of over 37 different reactions in the cells and CDH activity is measured in the cell extract. Thus, it was very important to use a blank where there was no overexpression of this CDH to determine the genomic activity. Moreover, to ensure that the activity that was monitored was catalyzed

by CDH, alkaline phosphatase was added in excess so it would not limit the reaction.
To obtain the cell extracts, cultures of 50 mL were grown and induced overnight. After disrupting the cells, pellets were resuspended in a buffer proposed by Bulawa et al that did not contain phosphates that could interfere with the malachite green reaction (Bulawa and Raetzs, 1984) Reactions were started when substrate, CDP-DOG was added to the reaction mixture which contained cell extract and alkaline phosphatase *(more details in13.4.4. CDH activity)*.
The control strain used to determine the genomic activity of CDH, without overexpressing the enzyme, was strain #2 (ΔtesA/mg517-plsC^{H}) while the strain that overexpressed the mg517·cdh was strain #24 (ΔtesA/mg517·cdh-plsC^{H}). Results showed that there was a slight increase of this activity in the strains expressing CDH (Figure 50).

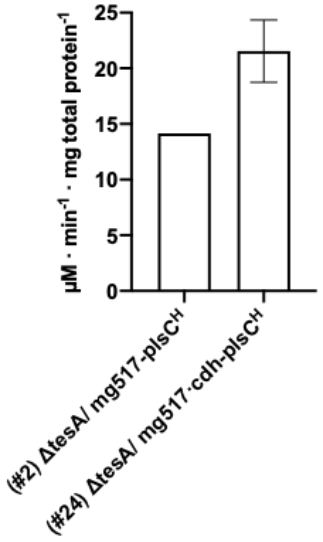

Figure 50. Effect of using CDH in BL21 Star (DE3) ΔtesA strain

7.3.2. Phosphatidylglycerophosphatase enzymatic activity

Δ*tesA* strain was transformed with plsCxpgpB^{H} and mg517 plasmids (#15 ΔtesA/mg517-plsCxpgpB^{H}). To measure the enzymatic activity, a protocol from Dillon et al was adapted to use malachite green reagent (Dillon et al., 1996). This assay was based on the ability of the malachite green to react phosphates groups.
The monitored reaction started with LPA, enzymatic substrate of PlsC acyltransferase, which was then converted by this enzyme into PA. This PA was substrate of PgpB, which when catalyzed the conversion of this to DAG, released phosphates. Those freed phosphates were able to react with the malachite green reagent allowing us to monitor the enzymatic activity (Figure 51) *(more details in 13.4.3. PgpB activity)*.

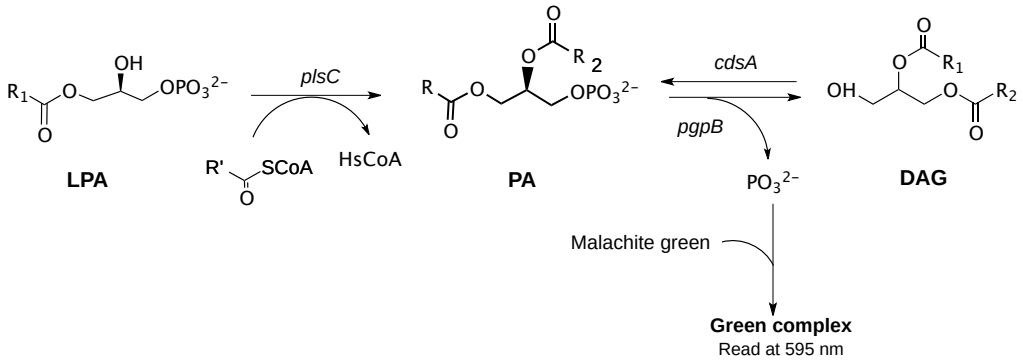

Figure 51. Assay principle to quantify by Malachite green PgpB enzymatic activity

As a control strain, BL21 Star (DE3) Δ*tesA* strain overexpressing mg517-plsC^{H} was used (strain #2) which provided the baseline activity of PgpB. Results in Figure 52 report an increase of 5-folds in the activity when this enzyme was compared to the genomic expression of this gene.

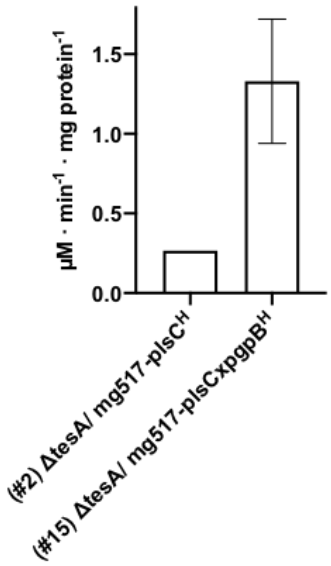

Figure 52. Enzymatic activity of PlsCxPgpB protein

7.4. Effect of CDH and PgpB on the GGL production

The production was studied in ΔtesA/mg517·cdh-plsC^{H}·plsB^{H} (#14) and ΔtesA/mg517-plsCxpgpB^{H} (#15) strains and compared to strains #2 (ΔtesA/mg517-plsC^{H}) and #13 (ΔtesA/mg517-plsC^{H}·plsB^{H}) as controls.

Table 18. Production of GGL produced by ΔtesA strain when expressing mg517·cdh-plsC^{H} and mg517-plsCxpgpB^{H}

Strain (genotype)	OD	[Glc]$_T$	%GGL composition				[GGL]$_T$	GGL
		(μg/mg)	%M	%D	%Tri	%Tetra	(nmol/mg)	strain/wt(#1)
(#2) ΔtesA/ mg517-plsC^{H}	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#13) ΔtesA/ mg517-plsC^{H}·plsB^{H}	1.9 ± 0.3	2.72	4	45	40	10	5.95 ± 0.60	1.8
(#14) ΔtesA/ mg517·cdh-plsC^{H}·plsB^{H}	1.7 ± 0.1	2.20	4	86	6	5	5.71 ± 1.40	1.8
(#15) ΔtesA/ mg517-plsCxpgpB^{H}	1.6 ± 0.2	4.83	4	45	29	23	9.93 ± 1.20	3.0

Δ*tes*A strain co-expressing CDH, MG517 and both acyltransferases did not have a significant effect in the total GGL production compared to the strain where CDH was not overexpressed (Table 18, #14 *vs.* #13 p-value 0.77 and #14 *vs.* #2, p-value 1), but presented a different GGL profile, with DGDAG accounting for 86% of all GGL products where the mean when this enzyme was not overexpressed was around 50% (Figure 53). On the other hand, the overexpression of the PlsCxPgpB fusion protein resulted in a significant increase of 1.7-fold GGL production compared to the parental strain (#2) (up to 9.9 nmol·mg^{-1}) (p-value 0.006), which also resulted in a 3-fold increase of the production compared to the WT strain (#1) (p-value 0.001). This #15 strain was the best GGL producer from all the engineered strains, indicating that the combined effect of the acyltransferase PlsC and the phosphatase PgpB on a ΔtesA/ mg517 background drives the pathway to DAG, which becomes more readily available as substrate for the glycosyltransferase MG517, thus increasing GGL production.

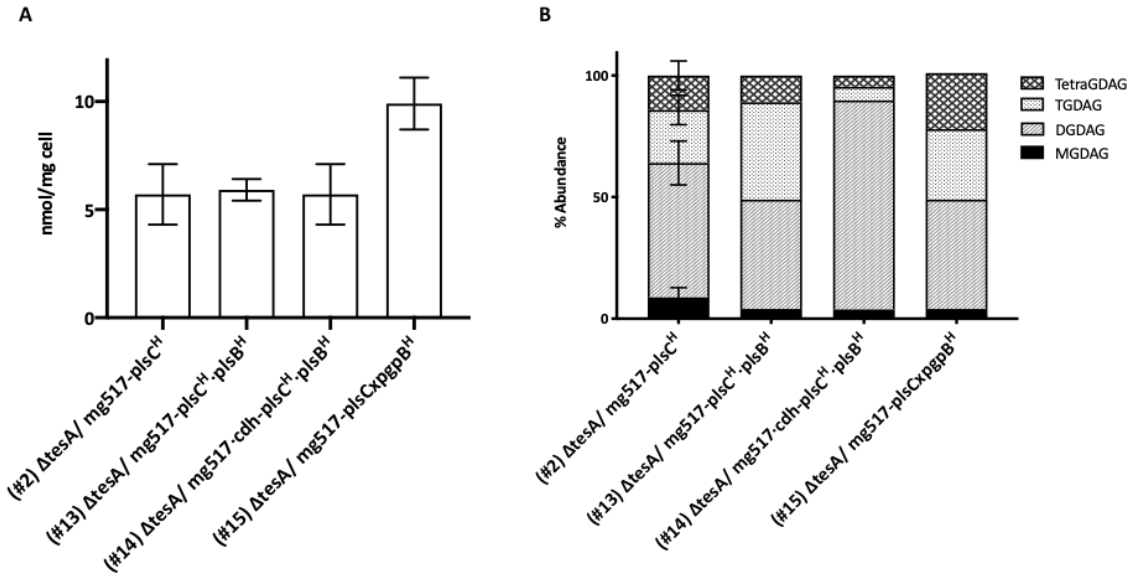

Figure 53. GGL production using CDH and fusion protein PlsCxpgpB (A) production in nmol/mg cell (B) glycolipid profile. Significant increase of the GGL compared to WT strain (#1) in strains #2, #13 and 15 (p-value 0.023, 0.011 and 0.001 respectively). No significant increase was observed comparing strains #14 and #1 (p-value 0.051)

STRATEGY 5: INCREASING UDP-Glc AVAILABILITY

8. Strategy 5: Increasing UDP-Glc availability

8.1. Strategy basis

DAG and UDP-Glc are the substrates of the glycosyltransferase GT MG517 to produce GGL. Once diacylglycerol availability was increased using a combination of the deletion of *tes*A gene, the overexpression of different acyltransferases and other genes related to the lipidic metabolic pathway it became an open question if the glycosidic donor had become limiting in these new-engineered strains.

Different enzymes that act sequentially are responsible to obtain the glycosidic donor required to produce GGL, UDP-glucose. The starting point is glucose-6-phosphate, obtained from glycolysis, which is transformed into glucose-1-phosphate by phosphoglucomutase (*pgm*) an enzyme responsible for the reversible transformation (Lu and Kleckner, 1994). Following this step, glucose-1-phosphate is then converted into UDP-glucose by the addition of the UDP moiety by UTP-glucose-1-phosphate uridylyltransferase (*galU*) (Hossain et al., 1994). This reaction is also reversible and the enzyme responsible for the opposite reaction is UDP-sugar hydrolase (*ushA*) (Glaser et al., 1967; Neu, 1967) (Figure 54). This enzyme has several substrates including UDP-galactose and UDP-N-acetylglucosamine among others (Glaser et al., 1967). This protein requires divalent cations which also give protection against heat inactivation. It can be inhibited by ADP-, CDP- and GDP-D-glucose in a competitive way.

In the first studies performed by our group, it was reported that when WT strain (#1) overexpressed GalU, no effect in the glycolipid production was observed. Considering the previous results obtained in the present project where an increase of GGL was observed by tuning the metabolic pathways to produce DAG, it was needed to test if UDP-Glc was now limiting (Mora-Buyé et al., 2012). Thus, two different approaches were followed. The first approach was based on overexpressing GalU. If UDP-Glc was now limiting the production of GGL, by overexpressing this enzyme an increase in the production should be observed. The second proposed approach was based on the results of the first approach. If an increase in the production was observed by overexpressing GalU, a knockout of *ushA* was proposed. UshA is the enzyme responsible for catalyzing the opposite reaction of GalU, meaning the conversion of UDP-Glc to Glc-1P. Several studies have reported the importance of *ushA* in the increase of the UDP-glucose pool (De Bruyn et al., 2015b; Pandey et al., 2014, 2013).

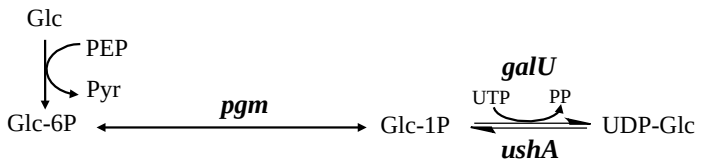

Figure 54. Biosynthesis of UDP-glucose in Escherichia coli

8.2. Molecular biology related to the strategy

8.2.1. GalU

This plasmid was already obtained in the work of Mora-Buyé et al. The vector used to express this protein was pCDF-1b which contained CloDF13 origin of replication and allowed between 20-40 copies per cell. GalU was obtained from *E. coli* JM109 strain (Mora-Buyé et al., 2012).

8.2.2. Obtaining Δ*tesA* Δ*ushA* strain

The strain used to remove *ushA* was Δ*tesA* strain.

As it happened with the double knockout Δ*tesA* Δ*fabR*, it was required to use large homologous sequences to be able to recombine the cassette which contained the kanamycin resistance with *ushA* gene.

Before removing the interest gene, the locus of this gene was studied in order to see if there were any interferences, meaning other open reading frames, promoters, etc. To do so, BioCyc database website was consulted and it was confirmed that *ushA* was an independent gene, without any promoters and open reading frames from other genes inside their CDS. It was also confirmed that it was not an essential gene so it could be removed without killing the strain.

Two sets of primers were designed to obtain the linear DNA required for the recombination event. The first set contained more than 50 bp for the genomic homology plus 20 bp more that annealed to pKD4 plasmid. The other set contained 50 bp more of homologous sequences and 20 bp that annealed to the previous set of primers (Figure 55).

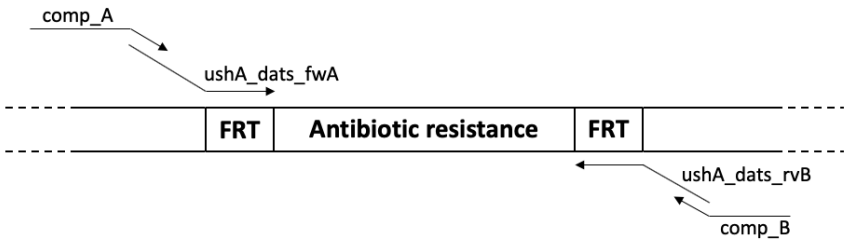

Figure 55. Primer design to obtain larger homologous sequences

The linear DNA fragment was created via PCR with primers ushA_dats_fwA, ushA_dats_rvB, comp_A and comp_B. To evaluate if the gene exchange process was successful, two PCRs were performed. One with the internal *ushA* primer (ushA_dins_fw), and the other one with the kanamycin internal primer (oMEMO3168). The control strain (BL21 Star (DE3)) was also analyzed as a reference.

In the first PCR using ushA_dins_fw and oMEMO2890 primers, it was expected to observe a band around 1509 bp if *ush*A gene was still present in the genome whereas no band was expected if the kanamycin cassette was successfully inserted (Figure 56). Only colonies 7, 16, 17, 19, 23 and 24 seemed to have removed *ushA*.

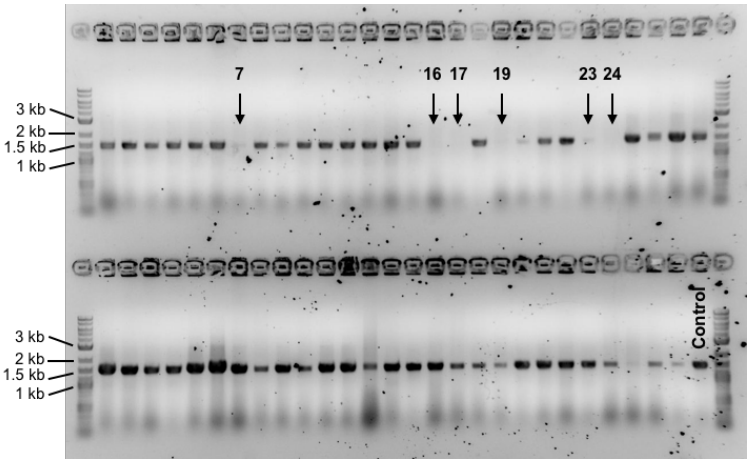

Figure 56. Agarose gel to confirm the presence of the kanamycin gene in the genome of the different ΔtesA strains tested. Those lanes without bands meant that ushA was removed from the genome

To confirm these results, a second PCR colony using an internal primer for the cassette (oMEMO3168) and an external primer of the region was performed. Figure 57 contains the agarose gel of these testes colonies where it can be seen false priming was present in the samples. If the cassette was correctly inserted, the band length should be around 1.2 kb while if *ush*A was in the genome, and the cassette did not get into the genome no band should have been seen. The colonies that were highlighted from the previous gel were taken into deep consideration and finally colonies 17, 19 and 24 were used for further steps.

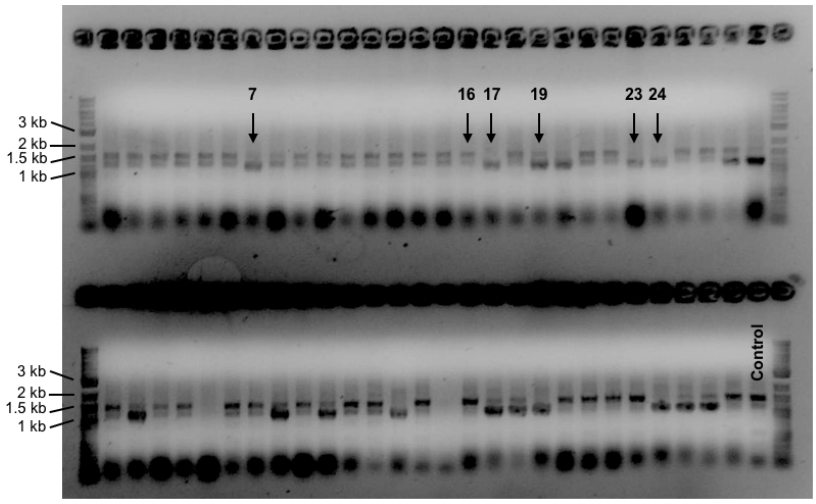

Figure 57. Agarose gel to confirm the presence of the resistance cassette using internal primers of the different colonies analyzed

The colonies previously selected were transformed with pCP20 to remove the resistance cassette and a PCR colony using oMEMO2890 and oMEMO2891 primers was performed to check the results. If the resistance cassette was correctly removed a band of 598 bp was expected (Figure 58, colony 29).

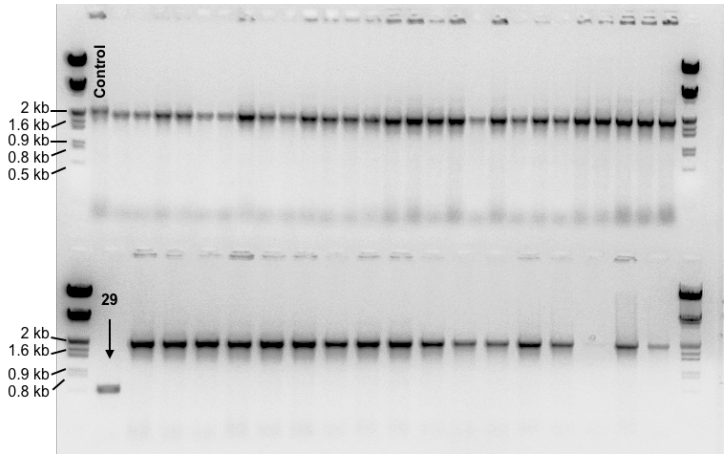

Figure 58. Agarose gel to confirm the removal of the resistance cassette

Analyzing the gel, only colony 29 (arrow) removed the resistance cassette so this one was picked and send for sequencing showing the correct removal of the gene *(more details in 15.1.6. BL21 Star (DE3) ΔtesA ΔushA)*.

8.3. GGL production

8.3.1. GalU overexpression

To study if UDP-Glc became limiting in the new engineered strains the highest producer, Δ*tesA* strain, was transformed with GalU, the enzyme responsible for the catalysis of glucose-1-phosphate to UDP-glucose.

The GGL production obtained by ΔtesA/mg517-plsC^{H}-galU (#16) was compared to ΔtesA/mg517-plsC^{H} (#2) strain showing that when Δ*tesA* was overexpressing GalU, an almost 2-fold increase in the production was observed. When Δ*tesA* was only co-expressing MG517 and PlsC^{H} (#2) the amount of product produced was 5.7 ± 1.4 nmol/mg cell while when this same strain co-expressed GalU (#16) the amount of product obtained was 9.1 ± 0.3 nmol/mg cell reporting a significant increase in the glycoglycerolipid production (p-value 0.02) (Figure 59). When the production of Δ*tesA* strain overexpressing GalU (#16) is compared to the first generation of engineered strains reported by Mora-Buyé et al (#1), it can be seen that an almost 3-fold increase in the production is reached (p-value 0.001).

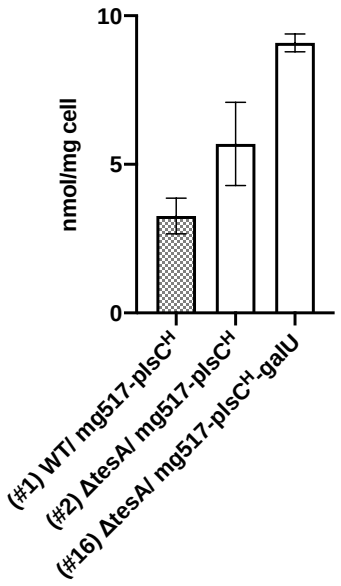

Figure 59. Production of ΔtesA with or without the expression of GalU enzyme

After analyzing these results, it became clear that in this new strain, by overexpressing GalU more UDP-glucose could be available and the production of GGL increased. Thus, in this project we proposed a knockout of *ush*A gene, the enzyme responsible for catalyzing the reverse reaction of GalU.

8.3.2. GGL production in Δ*tesA* Δ*ushA*

To test if GGL production was similar in this new engineered strain than in strain #16 (ΔtesA/mg517-plsC^{H}-galU) a GGL production analysis was performed.

As Figure 60 indicates, it is possible to increase the amount of glucose that can later be used to produce GGL by removing *ush*A gene when this new strain is compared to the parental (#2) and reference (#1) ones. The Δ*tesA* Δ*ushA* strains overexpressing MG517 and the acyltransferases (Table 19, strains #18, #19) provided a similar effect as the GalU overexpression in the Δ*tesA* strain (#16), obtaining around 9 nmol·mg^{-1} of GGL. This means that prevention of UDP-Glc hydrolysis also improves UDP-Glc availability for GGL synthesis. However, overexpression of GalU in this Δ*tesA* Δ*ushA* background (strain #20) did not further increase GGL production (p-value 0.12 comparing 18 *vs.* 20 strains). The lack of synergic effect of GalU overexpression and *ush*A knockout suggests that UDP-Glc reached a maximum level and by knocking out *ushA* or overexpressing of GalU, the UDP-glucose produced was already enough to not to be the bottleneck on the GGL production (Figure 60).

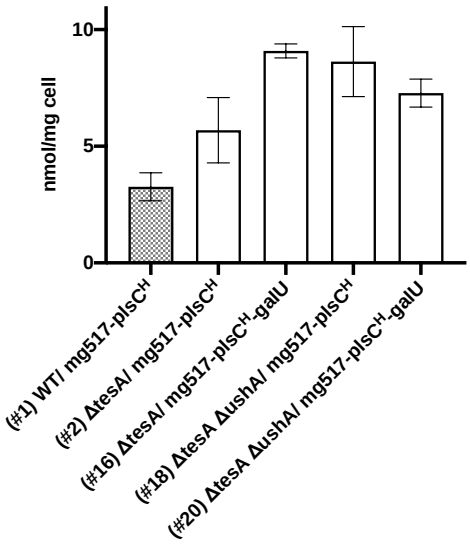

Figure 60. GGL production using ΔtesA and ΔtesA ΔushA engineered strains. Significant increase in the production compared to strain #1 of strains #2, #16, #18 and #20 (p-value 0.023, 0.002,0.006 and 0.004 respectively). Compared to strain #2, significant increase in the production in #16 and #18 strains (p-value 0.02 and 0.03 respectively)

Table 19. Summary of the production comparing BL21 Star (DE3) ΔtesA and BL21 Star (DE3) ΔtesA ΔushA strains

Strain (genotype)	OD	$[Glc]_T$ (μg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/wt(#1)
			%M	%D	%Tri	%Tetra		
(#1) WT/ mg517-plsC^H	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^H	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#16) ΔtesA/ mg517-plsC^H-galU	1.7 ± 0.2	4.30	25	21	20	34	9.08 ± 0.90	2.8
(#18) ΔtesA ΔushA/ mg517-plsC^H	2.2 ± 0.6	3.41	6	71	13	9	8.63 ± 1.50	2.6
(#20) ΔtesA ΔushA/ mg517-plsC^H-galU	1.8 ± 0.2	2.85	33	32	19	16	7.28 ± 0.60	2.2

The cellular phenotype was also analyzed by optical microscopy and as it happened with all the engineered strains producing glycoglycerolipids, Δ*tesA* Δ*ushA* strain (#18) had also filamentous shapes, sign of metabolic stress (Figure 61).

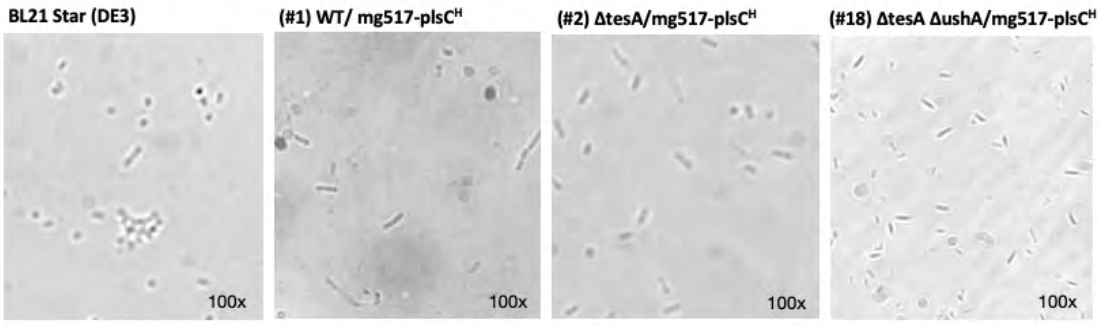

Figure 61. Microscopic analysis of the different engineered strains

8.5. Study of the different acyltransferases

Regarding to all the data obtained before, this new strain, which combines a removal of *tesA*

gene related to the DAG metabolic pathway and a removal of *ushA*, involved in the glycosidic pathway, was transformed with plsC^{H} and plsC^{H}·plsB^{H} plasmids to analyze the possible effect that it may have in the production of GGL. These new strains were compared to the control strains (#1 and #2) and also to the strain that showed that by overexpressing *galU* was possible to increase the GGL production. Figure 62 and table 20 contains the results of the production in the different strains.

Table 20. Glycoglycerolipid production using different acyltransferases in ΔtesA ΔushA strain

Strain (genotype)	OD	$[Glc]_T$ (µg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/wt(#1)
			%M	%D	%Tri	%Tetra		
(#1) WT/ mg517-plsC^{H}	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^{H}	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#16) ΔtesA/ mg517-plsC^{H}-galU	1.7 ± 0.2	4.30	25	21	20	34	9.08 ± 0.90	2.8
(#18) ΔtesA ΔushA/ mg517-plsC^{H}	2.2 ± 0.6	3.41	6	71	13	9	8.63 ± 1.50	2.6
(#19) ΔtesA ΔushA/ mg517-plsC^{H}·plsB^{H}	1.7 ± 0.1	4.02	2	63	15	18	9.13 ± 0.70	2.8

As it can be seen, by using plsC^{H} (#18) or plsC^{H}·plsB^{H} (#19) the same range of glycoglycerolipids can be produced as when using strain #16 (ΔtesA/mg517-plsC^{H}-galU) (p-value 0.45 and 1 respectively) and almost 3-fold higher than when compared to the WT strain (#1) (p-value 0.001 and 0.001 respectively). This result suggest that it is not required to use a plasmid containing *galU* to obtain similar amount of GGL decreasing this way the metabolic burden that supposes the overexpression of three different proteins and also its associated-resistance genes. Regarding to the Δ*tesA* Δ*ushA* strain, no differences between using one or both acyltransferases can be appreciated. Both strains were selected to be used in further studies.

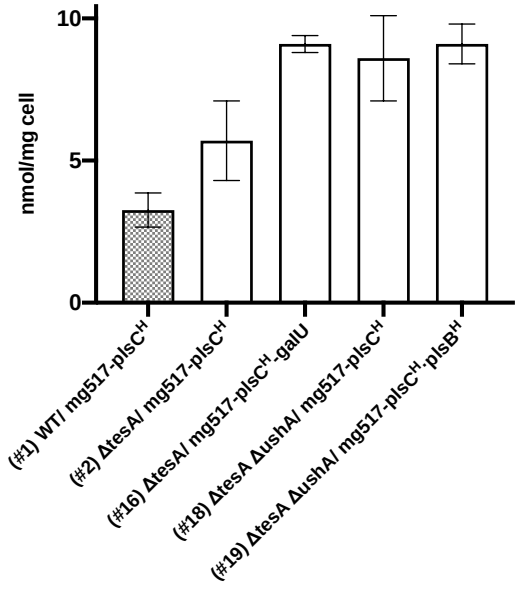

Figure 62. Production in ΔtesA ΔushA overexpressing different acyltransferases. Significant increase in the GGL production compared to WT (#1) strain (p-value 0.023, 0.002, 0.006 and 0.002 respectively) and strains #16, #18 and #19 compared to #2 strain (0.021, 0.033 and 0.011 respectively)

8.6. The effect of *fadR* overexpression in Δ*tesA* Δ*ushA* strain

The effect of the overexpression of FadR proposed in strategy 2 was also tested in Δ*tesA* Δ*ushA* strain co-expressing MG517 and PlsC·PlsB^{H}. Figure 63 shows that GGL production was lower when *fadR* was overexpressed (strain #21) compared to the strain #19 which does not overexpress this transcription factor (p-value 0.008) (Table 21). The level of GGL reached the value of 5.3 nmol/mg cell similar to the Δ*tesA* strain (p-value 0.93). This result indicates that FadR overexpression did not affect directly the DAG availability, no matter the UDP-Glc limitation.

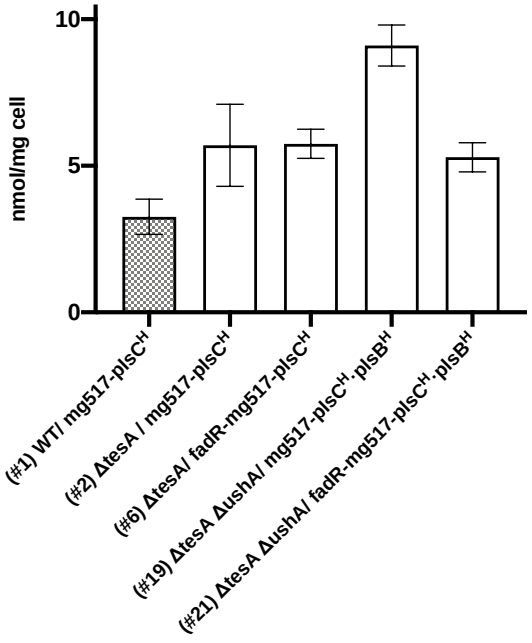

Figure 63. GGL production in ΔtesA and ΔtesA ΔushA engineered strains overexpressing fadR. A significant increase was observed in strains #2, #6, #19 and #21 compared to #1 strains (p-value 0.023, 0.015, 0,002 and 0.013 respectively) while a significant decrease in the production was observed comparing strains #21 and #19 (p-valua 0.009) which ment that fadR had a negative effect on the production

Table 21. Production in the different strains using fadR

Strain (genotype)	OD	$[Glc]_T$ (µg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/wt(#1)
			%M	%D	%Tri	%Tetra		
(#1) WT/ mg517-plsC^{H}	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^{H}	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#6) ΔtesA/fadR-mg517-plsC^{H}	1.7 ± 0.2	2.40	12	60	12	16	5.75 ± 0.50	1.8
(#19) ΔtesA ΔushA/ mg517-plsC^{H}·plsB^{H}	1.7 ± 0.1	4.02	2	63	15	18	9.13 ± 0.70	2.8
(#21) ΔtesA ΔushA/ fadR-mg517-plsC^{H}·plsB^{H}	2.9 ± 0.1	2.28	26	34	16	25	5.29 ± 0.50	1.6

As it happened with Δ*tesA* Δ*fabR* strain (#5), the main difference of overexpressing this protein relies on the content of unsaturated fatty acids produced. In this case, strain #21 showed an increase of C18:1 fatty acid over C18:0 (Figure 64, Table 22). This result suggests that the

overexpression of FadR causes more synthesis of unsaturated fatty acids in the different engineered strains. Thus, this plasmid could be interesting to be used when fatty acid profile needs to be modulated to affect cells (i.e., fluidity of cell membranes, growth conditions at low temperatures) or when products from them are obtained (i.e., rich unsaturated glycoglycerolipids).

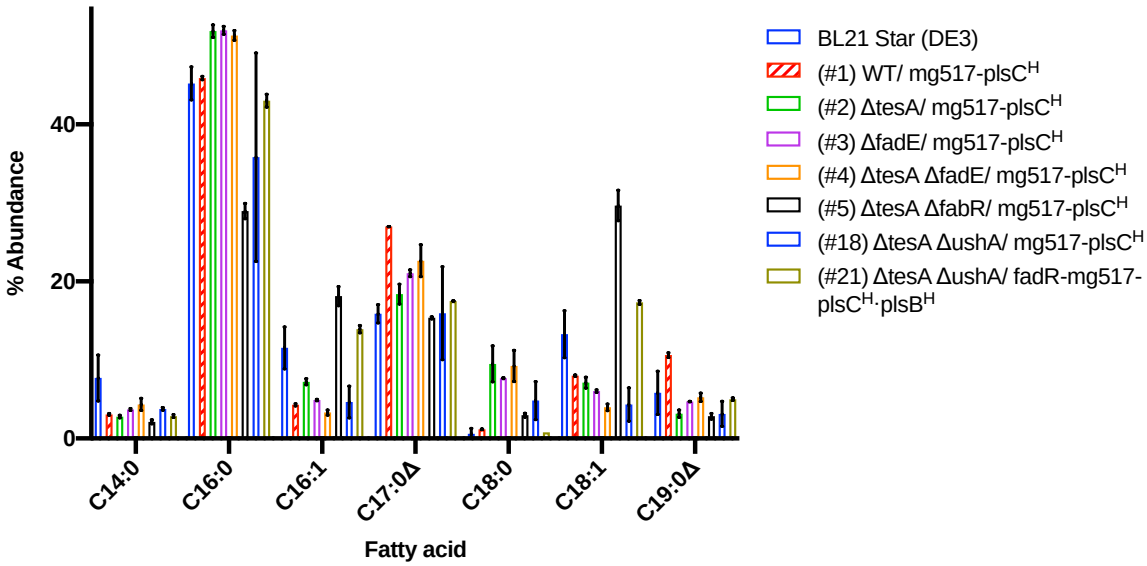

Figure 64. Fatty acid profile in the different engineered strains and using FadR in ΔtesA ΔushA strain to analyze its effects

Table 22. Fatty acid content in the different engineered strains

	C14:0	C16:0	C16:1	C17:0Δ	C18:0	C18:1	C19:0Δ
BL21 Star (DE3)	7.69 ± 2.92	45.22 ± 2.09	11.52 ± 2.68	15.87 ± 1.18	1.19 ± 0.05	13.28 ± 2.99	5.82 ± 2.76
(#1) WT	3.07 ± 0.12	45.89 ± 0.22	4.28 ± 0.14	26.98 ± 0.01	1.17 ± 0.05	8.00 ± 0.09	10.61 ± 0.31
(#2) Δ*tesA*	2.78 ± 0.17	51.87 ± 0.80	7.22 ± 0.39	18.37 ± 1.27	9.49 ± 2.29	7.11 ± 0.69	3.14 ± 0.48
(#5) Δ*tesA* Δ*fabR*	2.12 ± 0.28	28.97 ± 0.98	18.11 ± 1.22	15.37 ± 0.14	2.95 ± 0.14	29.66 ± 0.25	2.82 ± 0.36
(#18) Δ*tesA* Δ*ushA*	3.76 ± 0.18	35.81 ± 13.29	4.63 ± 2.03	15.94 ± 5.93	4.82 ± 2.43	4.32 ± 2.12	3.13 ± 1.61
(#21) Δ*tesA* Δ*ushA* FadR	2.84 ± 0.18	43.01 ± 0.82	13.91 ± 0.48	17.51 ± 0.03	0.79	17.33 ± 0.24	5.00 ± 0.18

8.7. CDH effect on the production

In chapter "*7. Strategy 4: Increase the production of phosphatidic acid from phospholipids*", it was demonstrated that the CDH overexpression did not increase the levels of GGL production in the Δ*tesA* strain (#13) but it did increase the proportion of DGDAG over TGDAG and TetraGDAG. The effect of the overexpression of CDH proposed in strategy 4 was also tested in

the new Δ*tesA* Δ*ushA* strain.

This new-engineered strain Δ*tesA* Δ*ushA* was transformed with the polycistronic plasmid mg517·cdh along with both acyltransferases (#22 strain). As it happened with strains #13 and #14, by overexpressing CDH in the strain #19, no increase in the DAG availability to be used to produce GGL was observed. Furthermore, when this enzyme is overexpressed in Δ*tesA* Δ*ushA* strain, a decrease in the production of GGL is observed (p-value 0.002). The reason why this happens could be due some genomic regulations and is something to be studied in further studies (Table 23).

Table 23. Production in nmol/mg cell in the different engineered strains overexpressing CDH

Strain (genotype)	OD	$[Glc]_T$ (µg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/wt(#1)
			%M	%D	%Tri	%Tetra		
(#1) WT/ mg517-plsC^H	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#13) ΔtesA/ mg517-plsC^H·plsB^H	1.9 ± 0.3	2.72	4	45	40	10	5.95 ± 0.60	1.8
(#14) ΔtesA/ mg517·cdh-plsC^H·plsB^H	1.7 ± 0.1	2.20	4	86	6	5	5.71 ± 1.40	1.8
(#19) ΔtesA ΔushA/ mg517-plsC^H·plsB^H	1.7 ± 0.1	4.02	2	63	15	18	9.13 ± 0.70	2.8
(#22) ΔtesA ΔushA/ mg517·cdh -plsC^H·plsB^H	1.8 ± 0.2	1.43	6	82	9	3	3.81 ± 0.30	1.2

When the glycolipid profile was analyzed, it was seen that when CDH was overexpressed in Δ*tesA* Δ*ushA* (strain #22), as it happened with strain #14, the proportion of DGDAG over TGDAG and TetraGDAG was increased. The proportion of this DGDAG glycolipid was increased supposing around 83% of the total abundance. These results suggest that CDH was active when was overexpressed in #14 and #22 strains and had an effect in the production of GGL even though this effect was not translated into an increase in the production (Figure 65).

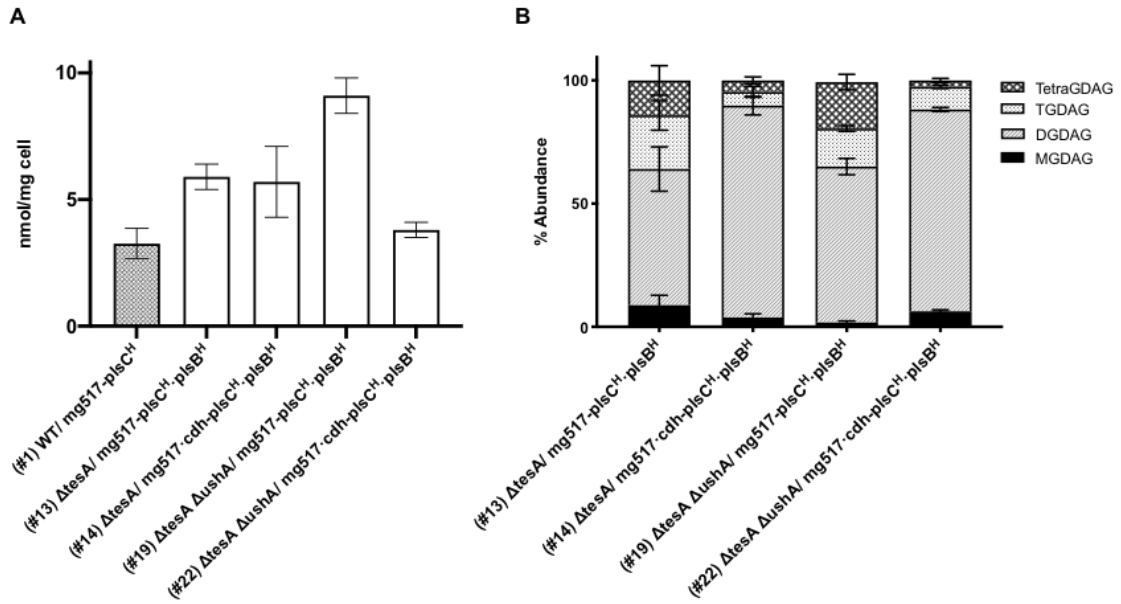

Figure 65. Effect of using CDHin ΔtesA and ΔtesA ΔushA engineered strains (A) Production of GGL in the different strains. A significant decrease in the production of GGL was observed comparing strains #22 and strain #19 (p-value 0.002) (B) Effect in the GGL profile

8.8. Study of the plsCxpgpB^{H} in Δ*tesA* Δ*ushA*

As reported in section "*7.4. Effect of CDH and PgpB on the GGL production*" it was seen that when PlsCxPgpB fusion protein was overexpressed it was possible to increase the production of GGL by 1.8 folds compared to ΔtesA/mg517-plsC^{H}. If this production is compared to the WT strain (#1) the production was increased by 3 folds.

Since UDP-Glc seemed to be a bottleneck in the Δ*tesA* engineered strain, as it can be seen in section "*8.3.1. GalU overexpression*", it was proposed to overexpress this same protein into Δ*tesA* Δ*ushA* strain to increase the production of GGL.

Surprisingly, whereas in ΔtesA/mg517-plsCxpgpB^{H} strain a 3-fold increase in the production was observed when plsCxpgpB (#15) was expressed, in Δ*tesA* Δ*ushA* (#23) the GGL production not only was not increased but it was slightly decreased when compared to ΔtesA ΔushA/mg517-plsC^{H} strain (#18) (p-value 0.07) (Table 24).

Table 24. GGL production in the different strains when plsCxpgpB protein was used

Strain (genotype)	OD	$[Glc]_T$ (μg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	GGL strain/wt(#1)
			%M	%D	%Tri	%Tetra		
(#1) WT/ mg517-plsC^{H}	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
(#2) ΔtesA/ mg517-plsC^{H}	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
(#15) ΔtesA/ mg517-plsCxpgpB^{H}	1.6 ± 0.2	4.83	4	45	29	23	9.93 ± 1.20	3.0
(#17) ΔtesA/mg517-plsCxpgpB^{H}-galU	1.9 ± 0.3	2.45	10	50	18	21	5.44 ± 1.50	1.7
(#18) ΔtesA ΔushA/ mg517-plsC^{H}	2.2 ± 0.6	3.41	6	71	13	9	8.63 ± 1.50	2.6
(#23) ΔtesA ΔushA/ mg517-plsCxpgpB^{H}	1.7 ± 0.1	3.09	6	63	17	13	7.22 ± 0.70	2.2

These results suggested that somehow the increase of the UDP-glucose levels might interfere in the effect that the fusion protein may have in the production of DAG. To confirm this hypothesis Δ*tesA* which contained PlsCxPgpB protein was also transformed with GalU enzyme (#17) in order to have similar conditions regarding the synthesis of UDP-Glc (Figure 66A). Surprisingly, when PlsCxPgpB was overexpressed along with GalU in Δ*tesA* strain, the production of GGL did not maintain the previously reported 9.9 nmol/mg cell (#15 strain) production and was decreased to 5.4 nmol/mg cell (#17 strain). While the overexpression of this fusion protein in Δ*tesA* strain led to an increase in the GGL production up to 9.9 nmol/mg cell by increasing the availability of DAG, this effect was not seen in Δ*tesA* Δ*ushA* strain. When *ushA* was removed from the genome of *E. coli* or even when GalU was overexpressed, the achieved production of GGL combining these strains with the overexpression of PlsCxPgpB protein was decreased. This suggests that the increase in the UDP-Glc availability is affecting somehow the regulation of the metabolic pathway to produce DAG compromising the production of GGL.

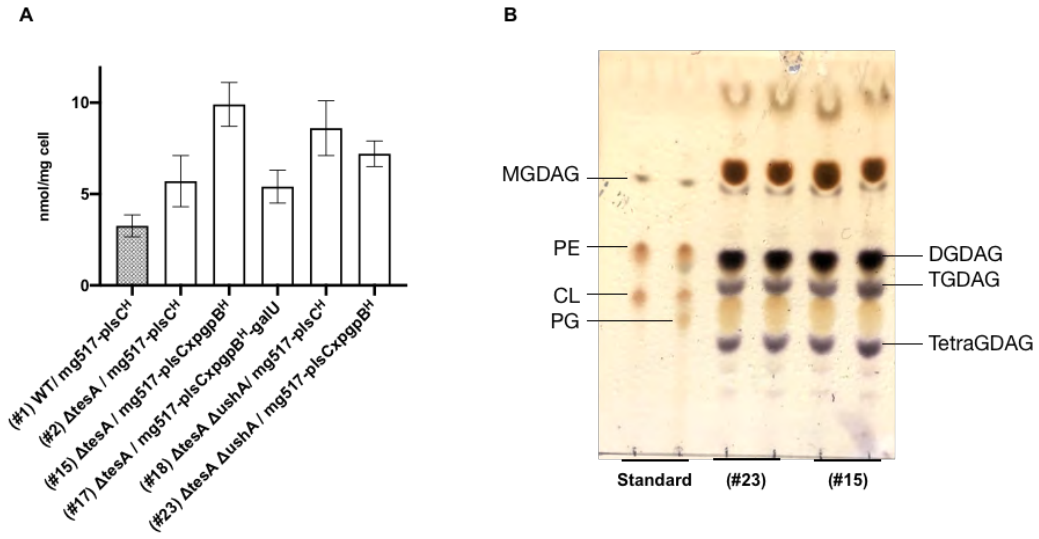

Figure 66. Production using PlsCxPgpB protein in different strains (A) Production. A significant increase in the production has been observed in strains #2, #15, #18 and #23 compared to #1 (p-value 0.023, 0.001, 0.006 and 0.017 respectively) and in strains #17 and #23 compared to #15 (p-value 0.009 and 0.011 respectively) (B) TLC

It has been reported that an accumulation of UDP-glucose may inhibit different genes such as glycogen phosphorylase and dTDP-glucose pyrophosphorylase. Both enzymes are related with the glycan pathways and this somehow could affect the metabolism of glycoglycerolipids (Bernstein and Robbins, 1965; Chen and Segel, 1968). Moreover, UDP-glucose has been also reported to be an activator of ADP-sugar pyrophosphatase which is involved in the production of glycogen (Moran-Zorzano et al., 2007).

To confirm that PgpB was active in Δ*tesA* Δ*ushA* strain (#23), an enzymatic assay was performed showing that, interestingly, when was PlsCxPgpB was overexpressed in Δ*tesA* strain (#15) the activity was increased by almost 5-folds while when this same protein was expressed in Δ*tesA* Δ*ushA* (#23) this enzyme decreased its activity almost in a half. This result could suggest that maybe UDP-Glc is affecting the activity of this enzyme (Figure 67).

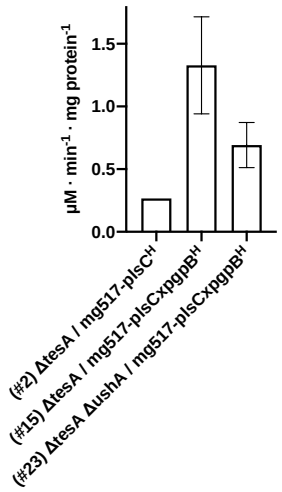

Figure 67. Enzymatic activity of PgpB enzyme in the different tested strains using BL21 Star (DE3) Δ*tesA* without overexpressing PlsC-PgpB protein as a genomic control

To explain why the production of GGL was decreased when PlsCxPgpB protein was used in those strains with higher UDP-Glc availability. We came with the following possible explanations:

On one hand, it was reported by Funk et al that PgpB enzyme can be inhibited by phosphates (Funk et al., 1992). In this sense, GalU which is responsible for catalyzing the conversion of Glc-1P to UDP-Glc, releases diphosphates groups that can lead to an inhibition of the PgpB protein. This might explain why when PlsCxPgpB fusion protein was studied in Δ*tesA* Δ*ushA* or Δ*tesA* overexpressing GalU strains (#23 and #17 respectively), which led to an increase in diphosphates groups, caused a decrease in the DAG availability by inhibiting PgpB enzyme. Overall, this inhibition could cause a decrease in the production of GGL.

On the other hand, it was also possible that the overexpression of this fused protein was affecting GalU and the production of UDP-Glc. In this sense, it was seen that in *Escherichia coli* this enzyme could be inhibited by different compounds such as TDP-glucose and TDP-rhamnose while in other species such as *Bos taurus* it was reported that this enzyme could be inhibited by phosphates (Turnquist and Hansen, 1973). The inhibition of GalU by phosphates could explain that when PgpB was overexpressed releasing phosphates, those could be inhibiting this uridylyltransferase decreasing the production of GGL. Even though this inhibition has not been reported in *E. coli* all the data obtained in these assays seem to indicate that it is possible that phosphates could also inhibit GalU. It can be that the overexpression of both enzymes would be freeing many phosphates leading to an inhibition of both enzymes.

Finally, another possible explanation would be that MG517 could be inhibited by product at certain point. This would explain that when PlsCxPgpB protein was used in Δ*tesA* Δ*ushA* (#23) or even in Δ*tesA* expressing GalU enzyme (#17), UDP-glucose and DAG pools would allow a great formation of GGL that inhibited MG517 and limit the production. This hypothesis cannot be supported by any literature since no inhibitors of this protein have been reported.

FURTHER CHARACTERIZATION OF THE ENGINEERED STRAINS

9. Further characterization of the engineered strains

The new engineered strains obtained throughout this project were further studied determining the MG517 glycosyltransferase activity and phospholipid analysis.

9.1. Glycosyltransferase characterization

9.1.1. Obtaining mg517xmCherry

To quantify *in vivo* the amount of glycosyltransferase present in the engineered cells, a fusion of *mg517* and *mCherry* was designed. *mCherry,* which was extracted from a fusion protein designed by our group, was fused to the C-terminal of MG517, which was cloned into pET44b(+) plasmid. The strategy followed to obtain this protein was based in restriction and ligation where the restriction enzymes chosen to perform the construct were EcoRI-HF (G'AATTC) and BamHI-HF (G'GATCC). To introduce these restriction sites in the insert (*mCherry*) and backbone (mg517) fragments, a PCR using primers that contained them in 5' was performed (Table 25).

Table 25. Primers and template used to obtain mg517xmCherry plasmid

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone	mg517	fw_mg517_mcherry rv_mg517_mcherry	6500
mCherry	pET16b-ß-glucanase-mCherry	fw_mcherry_mcherry rv_mcherry_mcherry	750

As Figure 68 shows, for the vector, where the expected band was of 6.5 kb, non-specific bands were also obtained. To purify the specific band of the vector, a gel extraction was required using GenElute gel extraction kit. For the insert, only the specific band of 750 bp was obtained.

After these procedures, a digestion was performed using the restriction enzymes EcoRI-HF and BamHI-HF and CutSmart as a reaction buffer ensuring a 100% activity for both enzymes. After three hours of digestion, an inactivation at 65 °C for 20 minutes was needed so the enzymes could be inactivated.

Later on, a PCR clean-up was performed in order to get rid of all the unnecessary compounds and, more importantly, to concentrate the samples. A Qubit HS (Life technologies®) quantification was done to determine the required amount of vector and insert needed to perform a ligation with T4 DNA ligase with a 1:3 (vector:insert) ratio. DH5α cells were transformed by heat shock and finally plated onto LB ampicillin plates and grown overnight at 37ºC.

Of all the colonies grown, eight were selected to perform a PCR colony (Figure 68B), where the expected length using T7 promoter and terminator primers was around 2.3 kb and 1.6 kb if this construct did not contain *mCherry*. Colonies 1 to 4 showed a band at 2.4 kb and all of them were sent to sequencing confirming the construct. Plasmid from colony 2 was extracted and selected to be used in further studies.

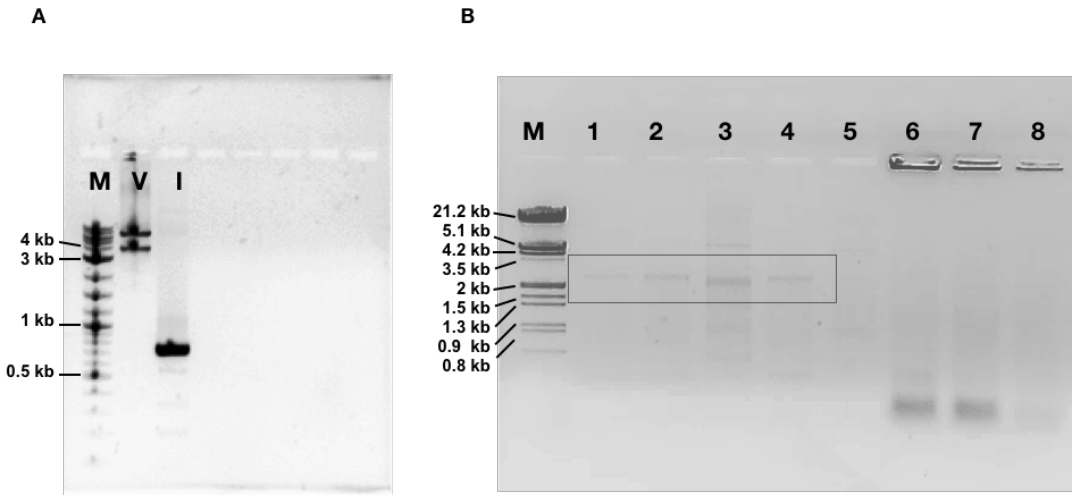

Figure 68. (A) Agarose gels containing the required fragments for the subcloning. M stands for 1 kb plus ladder (NEB), V for vector (mg517 plasmid) and I for insert, mCherry (B) Agarose gel of the resulting PCR colonies using Marker III (Roche). Highlighted the band expected for the amplification of 2.4 kb mCherry fragment

9.2.2. MG517 characterization

The GT activity of these new engineered strains was characterized in the cell extracts after 24h of growth as it is explained in section "*3.3.2 Glycosyltransferase activity determination*". Table 26 shows the MG517 activities where it can be seen that there were differences among them. The highest activities were found in Δ*tesA* (#2) and Δ*tesA* Δ*fabR* (#5) with values higher than 1.5 $\mu M \cdot min^{-1} \cdot mg\ protein^{-1}$ whereas Δ*fadE* (#3), Δ*tesA* Δ*ushA* (#18) and Δ*tesA* Δ*fadE* (#4) presented lower activities with values lower than the WT (#1) one (0.6 $\mu M \cdot min^{-1} \cdot mg\ protein^{-1}$). This result was surprising due to the fact that the expression of this protein is depending on the T7 promoter and it should not be related to the expression level of the protein.

Table 26. Glycosyltransferase activity of the different studied strains coexpressing mg517-plsC^H

Strain (genotype)	Specific activity ($\mu M \cdot min^{-1} \cdot mg\ protein^{-1}$)
(#1) WT/ mg517 - plsC^H	0.60 ± 0.11
(#2) ΔtesA/ mg517 - plsC^H	1.86 ± 0.38
(#3) ΔfadE/ mg517 - plsC^H	0.31 ± 0.10
(#4) ΔtesA ΔfadE/ mg517 - plsC^H	0.25 ± 0.08
(#5) ΔtesA ΔfabR/ mg517 - plsC^H	1.57 ± 0.30
(#18) ΔtesA ΔushA/ mg517 - plsC^H	0.31 ± 0.14

Figure 69 presents the glycosyltransferase (GT) activity and GGL production of these strains. As it can be seen, the GT activity does not seem to be related with the production of GGL since in strain #18, which presents the highest production of GGL, presents a lower activity than in strain #2 (Δ*tesA*) and #5 (Δ*tesA* Δ*fabR*) which are the second and third producer strains. Furthermore, for similar activities the level of production is different (i.e., strains #3 and #18 present the same activity but have very different GGL production). This suggests that the production seems not to be limited by the activity of MG517.

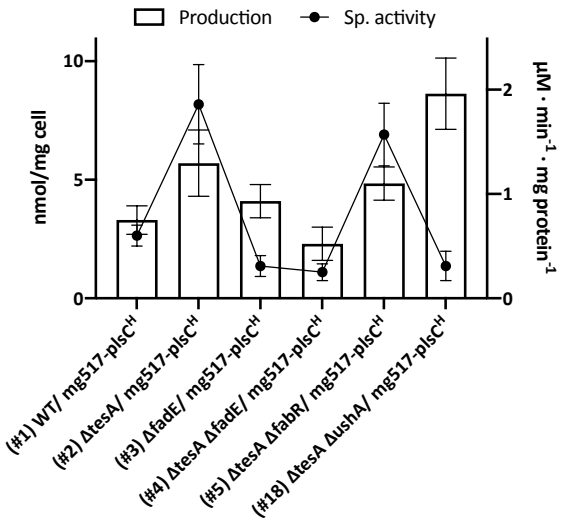

Figure 69. Relation between production and enzymatic activity of MG517 in the different engineered strains coexpressing the same plasmids. A significant increase in the GGL production was observed in strains #2, #5 and #18 compared to strain #1 (p-value 0.023, 0.024 and 0.006 respectively).

To investigate whether these significant variations of enzyme activity were caused by the different protein expressions levels or by changes in the specific activity that depends on the enzyme regulation by the lipidic environment, MG517 was expressed as a fusion protein with the fluorescent mCherry protein in the engineered strains.

All the new-engineered strains were transformed with the plasmid containing this new construct (mg517xmCherry) alone and also in combination with plsC^H.

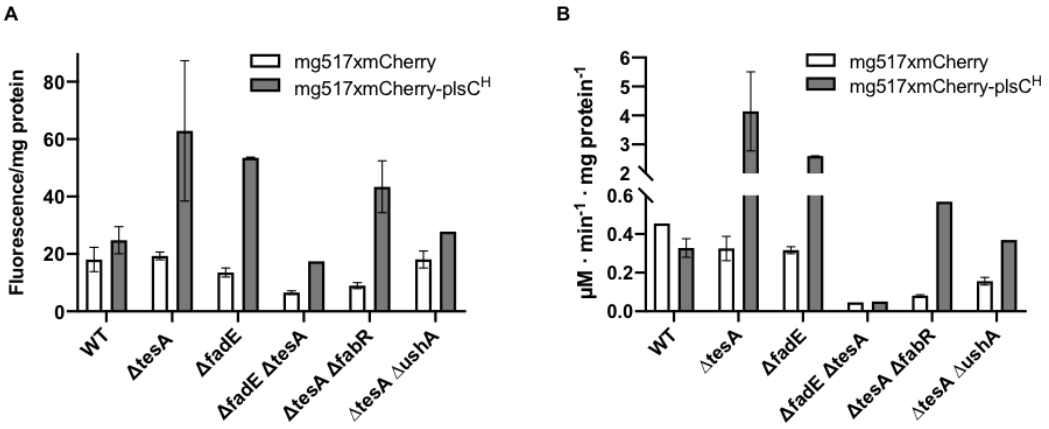

Figure 70. MG517 quantification by using mCherry as a reporter gene. (A) Fluorescence per mg of total protein when the strains were only expressing glycosyltransferase mg517xmcherry (B)Fluorescence per mg of total protein when the strains were expressing both GT mg517xmcherry and acyltransferase PlsC^H

Figure 70 shows the fluorescence and, therefore, the amount of MG517xmCherry protein with mg517 plasmid alone (white) or both in combination with plsC^H (grey). As it can be seen, when only MG517xmCherry was expressed, similar levels of fluorescence were found and only a slight decrease in Δ*tesA* Δ*fadE* and Δ*tesA* Δ*fabR* fluorescence was found.

On the other hand, when PlsC acyltransferase was expressed (grey bars) the amount of MG517xmCherry protein was higher when compared to the expression of only the

MG%17xmCherry fusion protein which indicated that more presence of phosphatidic acid or diacylglycerol make the glycosyltransferase more available.

The activity of this new fusion protein was also measured by the same activity assay with UDP-Gal and C6-NBD-ceramide described in section *"3.3.2 Glycosyltransferase activity determination"*.

Figure 70B shows that as it had happened with fluorescence, when only GT was expressed the enzymatic activity, even though it was similar for WT, Δ*fadE* and Δ*tesA* it was different from the other tested strains suggesting that there could be some interferences at transcription/translation levels that could interfere with the activity. Furthermore, it was also possible that some of the genomic modifications performed in the different strains favored a cellular microenvironment that provided a higher stability of MG517 and, therefore, activity. Moreover, these activities were lower compared to the same strains expressing plsC^{H} which strengthens the hypothesis of the cell microenvironment. Maybe by overexpressing PlsC acyltransferase, more anionic lipids, which are known to be activators of this GT, were produced (Andrés et al., 2011).

Table 27. Specific activity using MG517-mCherry fusion protein expressed alone or in combination with PlsC acyltransferase in a high copy number plasmid (PlsC^{H})

Strain (genotype)	Specific activity mg517xmCherry (μM · min^{-1} · mg protein^{-1})	Specific activity mg517xmCherry-plsC^{H} (μM · min^{-1} · mg protein^{-1})
WT	0.45 ± 0.00	0.32 ± 0.05
Δ*tesA*	0.33 ± 0.06	4.15 ± 1.37
Δ*fadE*	0.32 ± 0.02	2.60 ± 0.01
Δ*tesA* Δ*fadE*	0.05 ± 0.00	0.05 ± 0.00
Δ*tesA* Δ*fabR*	0.08 ± 0.00	0.57 ± 0.00
Δ*tesA* Δ*ushA*	0.16 ± 0.02	0.37 ± 0.00

To confirm that this effect of PlsC was also present in MG517 and not only in MG517xmCherry protein, this principle was also tested using mg517 and mg517-plsC^{H} combination of plasmids in order to check if this effect in the activity was also reproducible when the fusion protein was not used (Figure 71).

Results showed that the expression of PlsC acyltransferase had the same effect in the activity as it happened when mg517xmCherry was used. This result suggested that the overexpression of the acyltransferases may play a role in the cellular environment that makes the GT more stable and active.

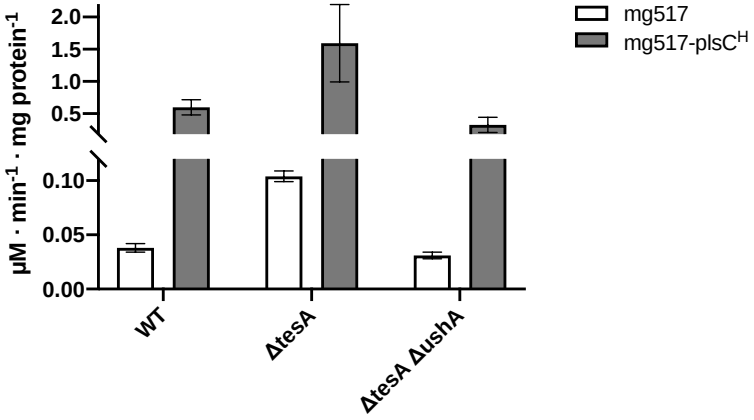

Figure 71. MG517 enzymatic activity when expressing PlsC or not

Table 28. Enzymatic activity in BL21 Star (DE3) when only using MG517 or MG517 and PlsC

Strain (genotype)	Sp. Activity (μM · mg protein^{-1} · min^{-1})	
	mg517	mg517-plsC^{H}
WT	0.038 ± 0.004	0.599 ± 0.117
Δ*tesA*	0.104 ± 0.005	1.593 ± 0.601
Δ*tesA* Δ*ushA*	0.031 ± 0.003	0.325 ± 0.118

Although the variability of MG517 expression levels needs to be further studied, if we consider the MG517 activity and the GGL production we can conclude that the enzymatic activity of this glycosyltransferase is enough in all the cases to detect the effect of the genomic modifications performed in the different strains. This can be seen for example, in Δ*tesA* Δ*ushA* strain, that presents lower enzyme amount and activity but reports higher amounts of GGL. Other strains even producing the same range of glycoglycerolipids present different GT activities (e.g., Δ*tesA* and Δ*tesA* Δ*fabR*).

9.3. Phospholipid analysis DAG

The main lipids in the *E. coli* membrane are phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL). PE (neutral zwitterionic lipid) can reach up to 75% of the total phospholipids depending on the growth conditions, whereas PG and CL (anionic lipids) account for 20% and 5% of the phospholipids composition (Wikstrom et al., 2009, 2004). GGL are neutral polar lipids that perturb the membrane properties. Our group previously reported for the 1st generation strains expressing MG517 and PlsC (strains #0 and #1) that GGL replaced PE, which was reduced between 10-20% relative to the control strain that does not present GGL to compensate membrane charge (Mora-Buyé et al., 2012). In this project it was analyzed the membrane lipid composition of the engineered strains to evaluate the effect of higher amounts of GGL.

To do so, cells were grown in minimal medium supplemented with ^{14}C-acetate for 24h. The lipid fraction, which now contained ^{14}C, was extracted. To quantify the phospholipids and GGL,

different conditions for TLC and exposure time were analyzed to select the option were the lipids could be quantified. Figure 72 shows the different TLC performed for each culture analyzed showing the correct separation of phospholipids and GGL.

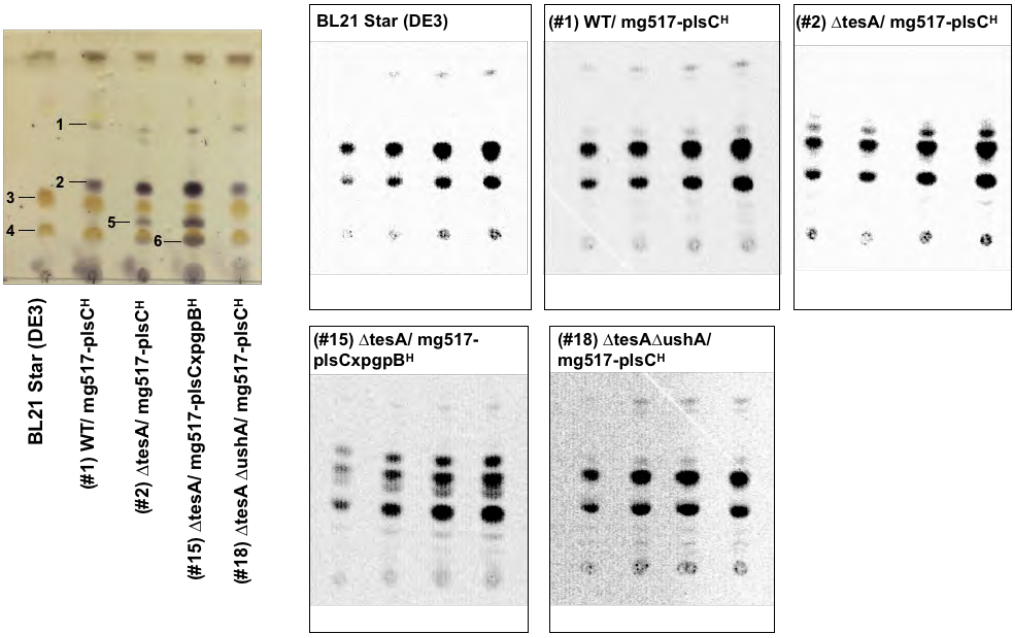

Figure 72. One-dimension TLC for radioactivity measure. (1) MGDAG (2) DGDAG (3) PE (4) CL + PG (5) TGDAG (6) TetraGDAG

As TLC shows, a correct separation of glycoglycerolipids and phospholipids was obtained since all the species could be seen independently as purple and brownish dots when stained with α-naphthol. The analyzed cultures were:

- BL21 Star (DE3) as a control. This strain did not express the GT from *Mycoplasma genitalium,* MG517 so no GGL were expected.
- (#1) WT/mg517-plsC^H. This strain was used as a reference from previous studies (Mora-Buyé et al., 2012).
- (#2) ΔtesA/mg517-plsC^H as one of the engineered strains that showed higher production of GGL.
- (#15) ΔtesA/mg517-plsCxpgpB^H as one of the highest producer strains.
- (#18) ΔtesA ΔushA/mg517-plsC^H as other of the highest producer strains.

To further separate all the lipidic species, a 2-dimension TLC was performed (*see 13.3.3.2. Two-dimension TLC*) and expose these TLC for 16 hours. The radioactivity that was measured was analyzed by Quantity One® software so the percentage and abundance of each specie was determined. Final conditions used to quantify the phospholipids and GGL can be seen in the following table.

Table 29. Conditions and equipment used to reveal the radiolabeled TLC

Radiometric equipment	Personal Molecular Imager (PMI) (Bio-Rad®)
TLC exposure time	16 hours
Software analysis	Quantity One (Bio-Rad®)
Silica used for the TLC	TLC silica gel 60 F_{254} (Merk®)
TLC Mobile phase	1st Dimension: Ethyl acetate, isopropanol, chloroform, methanol, 0.25% KCl in water (25:25:25:11:9) 2nd Dimension: Chloroform, methanol and water (65:25:4)

The main lipids in the *E. coli* membrane are phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL). PE , that can reach up to 75% of the total phospholipids depending on the growth conditions, has been previously reported by our group to be exchangeable by MGDAG. In the 1st generation of strains expressing MG517 and PlsC (#1 strain) GGL replaced PE, which was reduced between 10-20% relative to the control (no GGL) strain to compensate the membrane charge (Mora-Buyé et al., 2012). Figure 73 shows that the WT strain (#1) contained around 31% GGL, whereas the engineered strains expressing MG517 and PlsC, Δ*tes*A, Δt*es*A Δ*ush*A or Δ*tes*A in combination with PgpB reached GGL levels between 40 and 50%. Compared to a control *E. coli* strain, the amount of PE declined from 63 to 12% in the best producer strain, ΔtesA/ mg517-pgpBxplsC^{H} (#15). Therefore, GGL is replacing PE for membrane charge compensation.

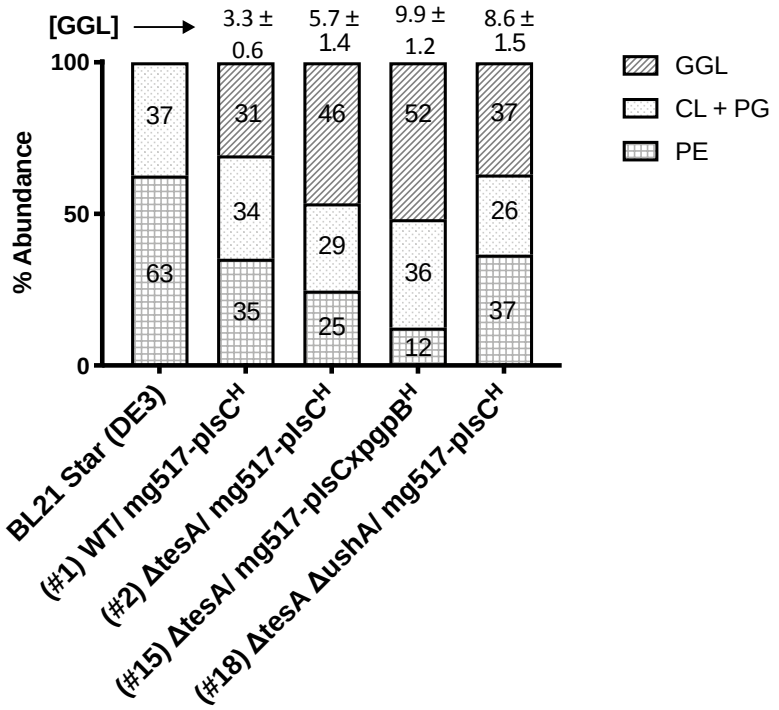

Figure 73. Membrane composition of the different engineered strains

Every phospholipid plays a role in the maintenance of the membrane. Phosphatidylethanolamine is the major phospholipid found in *E. coli* cell membranes. Its main

biological role is to spread out the negative charge caused by other anionic membrane phospholipids and is especially important for its role in supporting lactose permeases and work as a chaperone to ensure the correct formation of the membrane proteins. Along with cardiolipin, the structure of this phospholipid tends to form a monolayer instead of a bilayer like phosphatidylglycerol.

The main role of cardiolipin in the membranes is to give negative charge while the role of phosphatidylglycerol is, as it happens with cardiolipin, give negative charge and form a bilayer.

Regarding to glycoglycerolipids functions, its structure also has an impact in the function of membranes (Brandenburg et al., 2003; Heinz and Christie, 1996; Klement et al., 2007). For example, MGDAG it is known to be a neutral and non-bilayer lipid while for DGDAG and TGDAG, even though having neutral charge, are bilayer-forming lipids. Xie and coworkers reported that monogalactosyldiacylglycerol and phosphatidylethanolamine were interchangeable in the membranes and could perform the same functions (Xie et al., 2006). Interestingly, the most abundant GGL in all the engineered strains was DGDAG, which is known to be a bilayer-forming lipid and therefore should not be interchangeable by PE. This DGDAG should be exchanged by PG since is the natural bilayer-forming phospholipid of *Escherichia coli*. Even though there was a slight decrease in the levels of PG and CL, the main decrease was observed in PE which is the neutral phospholipid. This could indicate that, whereas DGDAG and TGDAG are the bilayer-forming lipids and will confer more rigidity to the membranes the fact that they are neutral make them potential exchangeable for PE.

DISCUSSION

10. Discussion

Regarding all the previous results a general discussion of the results is presented.

To produce the product of interest, glycoglycerolipids, it was necessary to have diacylglycerol and UDP-glucose, which are common metabolites in *Escherichia coli*. Also, it required an enzyme that was not present in this strain but in *Mycoplasma genitalium*, being this one a diacylglycerol ß-glycosyltransferase (MG517). This enzyme was able to transfer UDP-glucose to diacylglycerol.

In previous studies performed by our group, it was reported that by expressing this enzyme into *E. coli* it was possible to obtain GGL (Mora-Buyé et al., 2012). In this project, genetic modifications of the biosynthetic pathways related to GGL precursors were explored in order to detect which potential genes could be removed or increased.

Different strategies were proposed to increase both precursors: UDP-Glc and DAG. On one hand, since GGL are glycoconjugates with diacylglycerol as the lipidic part, we faced the complexity of modulating the biosynthetic pathway of diacylglycerol, whose precursor is phosphatidic acid. This PA is highly since it is the precursor of phospholipids. On the other hand, the biosynthetic pathway of UDP-Glc was also studied in order to increase it.

To do so, several knockout strains were proposed removing different metabolic pathways such as ß-oxidation (*fad*E), free fatty acid hydrolysis (*tes*A), general regulation of the biosynthesis of fatty acids (*fab*R) and hydrolysis of the glycoside donor UDP-glucose (*ush*A). All these knockouts were also combined with the overexpression of, firstly acyltransferases and afterwards different enzymes related to the biosynthesis of phosphatidic acid, diacylglycerol or glycosidic donor (Figure 74).

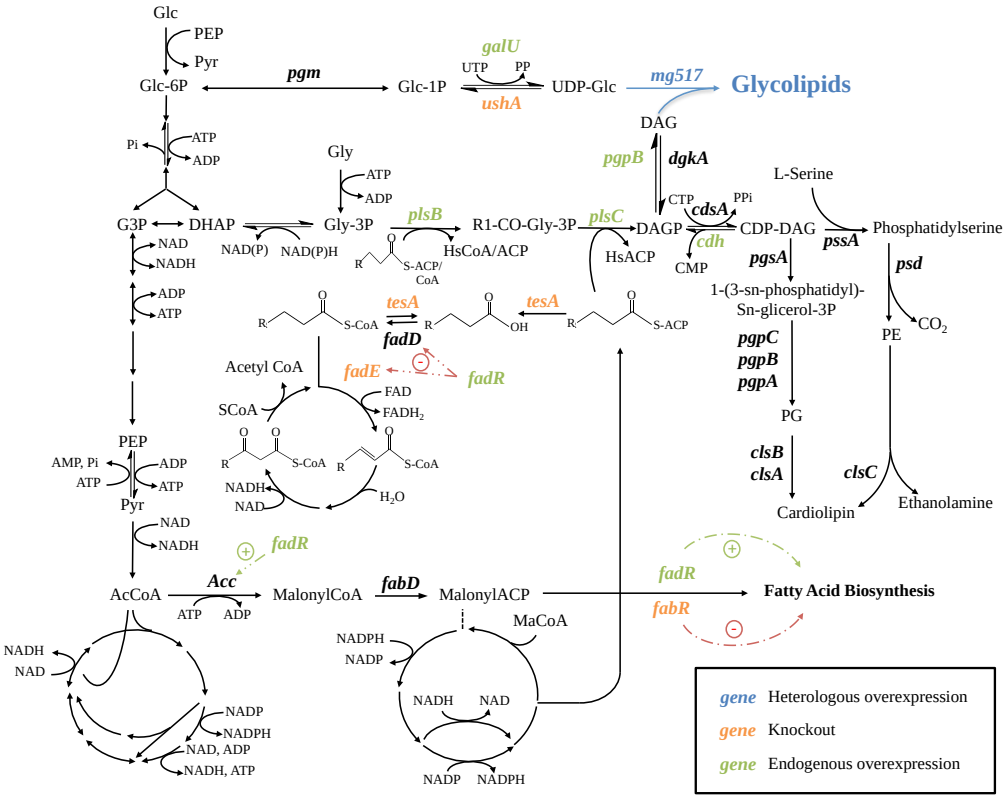

Figure 74. Metabolic pathway for the synthesis of glycoglycerolipids in Escherichia coli

To analyze the effect of each KO in the production, all the strains were tested using the same combination of plasmids (mg517-plsC^{H}) which were reported to be the best producer among the first engineering strains by our group in previous studies (Mora-Buyé et al., 2012). The highest production was found in the strains where *tesA* was removed. This result indicated to us that by removing this gene from the genomic DNA of *Escherichia coli,* it was possible to decrease free fatty acid production and use these fatty acids instead to enhance phosphatidic acid availability for phospholipid or glycolipid synthesis.

When this first knockout was combined with a removal of *fabR* gene, the global transcription regulator that is responsible for the inhibition of the fatty acid synthesis, it was seen that although the production was not increased, the lipid profile was changed to synthesize more unsaturated fatty acids. These fatty acids could be interesting when some properties like more fluidity and lower melting points in the membranes are sought. Having more unsaturated fatty acids could be interesting due the possibility of growing cultures at lower temperatures since the membrane fluidity would be increased. Furthermore, it was seen that in Δ*tesA* new engineered strains, the glycoside donor became limiting. This was seen by overexpressing GalU enzyme (#16 strain), showing an increase in the production of GGL, and by the removing *ushA* (#18). *ushA is* a gene implied in the hydrolysis of UDP-glucose to Glc-1-P. The production achieved by removing *ushA* was comparable to the production obtained in ΔtesA/mg517-plsC^{H}-galU.

These two strains, Δ*tesA* and Δ*tesA* Δ*ushA*, were selected to perform the different studies due their higher productions in comparison to the other ones. Both strains were transformed with different acyltransferases showing that the highest production was achieved when any of the acyltransferases were expressed into high copy number plasmids.

To study the effect that increasing the availability of PA may have in the production of GGL, these strains overexpressed CDP-diacylglycerol diphosphatase (CDH). This enzyme is responsible for catalyzing the reaction of CDP-DAG to PA. Thus, this PA could be used to produce DAG, precursor of GGL. Even though the overexpression of CDH did not significantly contribute to increase GGL levels, it changed membrane glycolipids distribution in the engineering strains by increasing the abundance of DGDAG over TGDAG and TetraGDAG. As far as we know, no previous studies have evaluated the effects of CDH overexpression in engineered cells. The closest example is the overexpression of a related enzyme involved in ceramide biosynthesis in yeast, the inositol phosphosphingolipid-phospholipase Isc1, which showed a 4-fold increase in ceramide levels (Murakami et al., 2015).

Finally, phosphatidylglycerophosphatase B (PgpB) was fused to PlsC acyltransferase (plCxpgpB^{H}). The idea of obtaining this fusion protein was based in facilitating the formation of DAG by proximity effect since both enzymes catalyze sequentially the formation of this compound. To do so, it was necessary that PgpB had a mM Km so it could quickly transform PA to DAG once it was obtained. The reason why both enzymes were fused was due to PgpB having several substrates being the most important one PGP so it was hypothesized that this reaction could be facilitated by proximity effect. This new protein was subcloned into high copy number plasmid and the activity was analyzed showing an increase compared to the genomic gene. Δ*tesA* strain coexpressing mg517-

plsCxpgpB^{H} showing an almost 3-fold increase in the production but, surprisingly, it did not have the same effect in Δ*tesA* Δ*ushA*. To confirm that this strain was actually expressing the enzyme, the enzymatic activity was tested showing that its activity was increased when compared to the genomic one but it was not as high as when this same protein was used in Δ*tesA*. In order to know if this difference was due to the genomic modifications of Δ*tesA* Δ*ushA* strain or it was a metabolic regulation, Δ*tesA* was transformed with this fused protein along with MG517 and GalU, the enzyme which catalyzes the opposite reaction of *ushA* (#17 strain). Results showed that when GalU enzyme was also expressed, the production of GGL, which was previously 3-fold higher, decreased to the same level as if PgpB was not expressed (#2). This data was consistent with the results obtained in Δ*tesA* Δ*ushA* indicating that there was a metabolic regulation that somehow interfered with the production of glycoglycerolipids. Different hypotheses why this was happening were proposed. The first one was based on an inhibition of PgpB enzyme by phosphate released during the formation of UDP-Glc by GalU. On the other hand, another possible explanation was the inhibition of GalU by the release of phosphates produced by PgpB. Another explanation could be that a product inhibition was happening when PlsCxPgpB was overexpressed in combination with a removal of *ushA* or overexpression of *galU* genes. The product obtained, glycoglycerolipids, could be inhibiting *Mycoplasma genitalium* glycosyltransferase MG517.

Moreover, the fatty acid profile of all the engineered strains obtained was analyzed and it was seen that the major lipid found in almost all the strains was C16:0 (palmitic acid) which is the major FA found in *E. coli* followed by C17:0Δ, C18:0 and C19:0Δ. Only Δ*tesA* Δ*fabR* and Δ*tesA* Δ*ushA* overexpressing FadR presented a different profile where the major lipid species found were oleic acid (C18:1), followed by palmitic (C16:0) and palmitoleic (C16:1) acids. These results suggested that by overexpressing *fadR* or knocking out *fabR* the titer of unsaturated fatty acids could be increased.

Nowadays, is still unclear which composition of GGL is the best as bioactive compounds for biomedical applications. It is known that monogalactosylglycerolipid and derivatives present antitumoral activities (Akasaka et al., 2016, 2013; Colombo et al., 2011; Maeda et al., 2013, 2011) and that GGL from *Ficus microcarpa* (mainly mono- and digalactosylglycerolipid) and *Meithermus taiwanensis* (tetrasaccharide derivative) had shown inhibitory effect on TNF-α-induced IL-8 secretion and cytokine production of monocytes respectively (Ghosh et al., 2013; Kiem et al., 2012). In addition, GGL are analogs of mono- and diglycosylceramides which are considered interesting marine bioactive compounds because of their immunostimulatory properties (Blunt et al., 2018; Costantino et al., 2005; Rocha-Martin et al., 2014). The different GGL composition of these engineered strains could be assayed as vaccine adjuvants to increase the immune response.

The effect that GGL production may have in the phospholipids cellular levels was also studied. To do so, it was necessary to supplement the minimal medium with ^{14}C- acetate during cell growth so cells could use it as a carbon source to produce GGL and phospholipids. Results showed that when glycolipid synthase MG517 was expressed, the levels of phospholipids, especially PE, were decreased. In the most prominent case, the levels of PE decreased from 63% (control strain) to 12% (in ΔtesA/mg517-plsCxpgpB^{H}). There was a relation between the strains. Those strains that

produced more GGL presented lower amount of PE whereas those who produced less product had higher proportion of phospholipids. It was also seen that even though the less abundant glycolipid produced was MGDAG, which was structurally closer to PE than DGDAG and TGDAG, the reduction was present in PE lipid. This could be explained due the produced GGL were neutral, as PE, which would cause less cellular effect than change it for CL or PG which had negative charge. These results were very interesting and were the base for the next chapter.

All these strategies were explained in the article published in 2020 by our group in Metabolic Engineering (*annex 16. Publication*).

PHOSPHATIDYLSERINE DECARBOXYLASE LIBRARY

11. Phosphatidylserine decarboxylase library

11.1. Introduction

This work was done with the collaboration of Prof. Marjan de Mey at the Ghent University.

The biosynthesis of phospholipids is a well-known metabolic pathway. The general precursor to synthetize each phospholipid starts with phosphatidic acid (PA) which is then converted by *cds*A into CDP-diacylglycerol. This enzyme is responsible for catalyzing the conversion of phosphatidic acid to CDP-DAG, opposite reaction of the one performed by the previous described enzyme CDH. This CDP-DAG can then follow two different pathways to obtain the different phospholipids. On one hand, to synthetize phosphatidylglycerol, one of the phospholipids that confer negative charge to the membrane, PgsA followed by the complex PgpA, PgpB and PgpC catalyze the formation of 1-phosphatidylglycerol which can then be converted into cardiolipin by *cls*A and *cls*B enzymes. On the other hand, to obtain phosphatidylethanolamine, which is the major phospholipid in *E. coli*, *pssA* catalyzes the formation of phosphatidylserine from CDP-diacylglycerol. Once this compound is produced a decarboxylation catalyzed by phosphatidylserine decarboxylase (*psd*) produces phosphatidylethanolamine (PE). Cardiolipin can also be obtained from PE by the catalysis of ClsC (Figure 75).

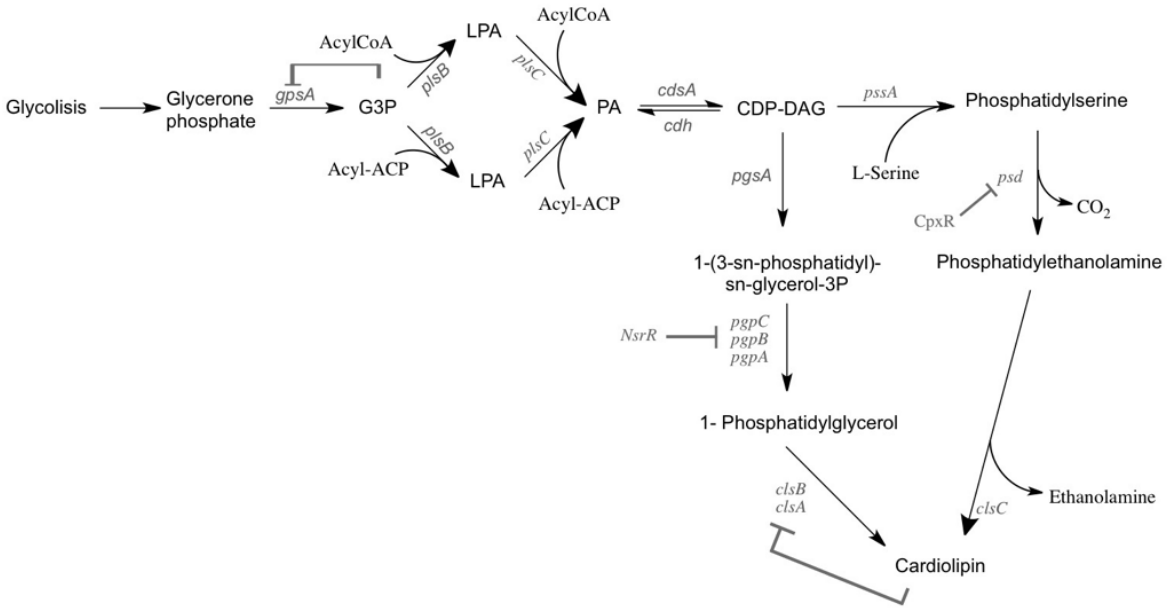

Figure 75. Phospholipid biosynthesis in Escherichia coli

In this chapter, we were focused in the biosynthesis of phosphatidylethanolamine and its relationship with the glycoglycerolipid synthesis. It was reported by Matsumoto et al that when *psd* enzyme was mutated, decreasing its activity, it was possible to increase the amount of glucosyldiacylglycerol, naturally produced by *B. subtilis,* two to four-folds (Matsumoto et al., 1998). In addition, Dowhan and coworkers that it was possible to interchange PE for MGDAG in *Escherichia coli* when the MGlcDAG synthase from *Acholeplasma laidlawii* was expressed (Wikstrom et al., 2004; Xie et al., 2006) In this project, we proposed a library of promoters and RBS for *psd* gene,

which is an essential gene and, therefore, could not be removed from the genome. This gene codifies for the last responsible enzyme involved in the biosynthesis of PE. The idea was to obtain a strain where the expression of this gene was as minimum as possible so when it was combined with glycosyltransferase MG517 and the desired genetic modifications that produced the highest amount of GGL, this production could be increased.

Before starting to explain how the library was designed and performed, it is interesting to discuss a little more about the role of PE in the cell envelope. Dowhan et al reviewed the biological functions of the three major phospholipids found in *E. coli* (PE, PG and CL) (Dowhan, 1997). Phosphatidylethanolamine as the major phospholipid found in the cell envelope represents between 70 and 80% of the total phospholipids. It is a non-bilayer prone lipid and its presence affects the activity of different proteins *in vitro.* To study the effect of this phospholipid may have *in vivo* in 1991, DeChavigny et al reported that it was possible to suppress the biosynthesis of phosphatidylethanolamine by removing *pssA* gene in *E. coli* and conserve cell viability by adding metal divalent cations which were able to interact with cardiolipin (DeChavigny et al., 1991; Killian et al., 1994). The reduction of the PE caused several phenotypic alterations as they were not able to grow without the presence of divalent metal cations at millimolar concentrations being the most important one's calcium followed by magnesium and strontium. Individual cells tended to form long filaments, indicating cell stress that grew in number as culture reached stationary phase. These alterations where decreased when a glycolipid synthase of *Acholeplasma laidlawii* was expressed in these cells (Wikstrom et al., 2004)

Dowhan and coworkers reported that when the MGlcDAG synthase from *Acholeplasma laidlawii*, PE was interchanged by MGDAG the dependence on the divalent metals was reduced and it was possible to restore a part of the cellular phenotype (Wikstrom et al., 2004; Xie et al., 2006).

Here, we focused on trying to obtain a new strain by modifying the promoters and RBS of *psd* to maintain the minimal required production of PE so metal dependence was not required. Then, by combining this new strain with the previous modifications which showed to produce the highest amount of glycoglycerolipids, a new strain would be obtained to study if it is possible to increase the GGL production by decreasing the PE content.

11.2. Regulation of *psd*

Phosphatidylserine decarboxylase gene can be found in an operon where the miniconductance mechanosensitive channel protein (MscM) is also produced. This latter protein, which share the same promoter as *psd,* is not essential for growth cell and its function is related for the response to the membrane tension and avoidance of hyperosmotic shocks (Edwards et al., 2012). Due to the fact that at this part of the project it was not intended to suppress the expression of this operon but decreasing it, this gene codifying for MscM was neglected (Figure 76A).

Whereas it was important to understand the functioning of this operon it was also interesting to know about the regulation of *psd* gene and the effect that touching this gene may have in the metabolism of *E. coli*.

As Figure 76B shows, De Wulf et al reported that there is a transcriptional activation of *psd* when CpxA was produced while Rezuchova et al reported that when CpxR-phosphorilated was present in the cell there was a repression of it (De Wulf et al., 1999; Rezuchova et al., 2003). Cpx is one of the main ESRS (Envelope Stress Response System) in *E. coli* which is responsible for sensing the incorrect folding of proteins in the periplasm. Is a two-component system. When is activated, CpxA is responsible of phosphorylating CpxR, which then represses the expression of several membrane complexes (Raivio, 2014).

Following the transcription and translation, the produced Psd is synthetized as a proenzyme (also called π), which is not active (Figure 76B). To be active, this protein must have an autocatalytic event where Serine 254 plays a key role. This autocatalytic event produces a bond scission between Glycine-253 and Serine-254. This mechanism of cleavage, provides two peptide chains (α and ß) which after the serinolysis entailed by residue Ser-254 results in the formation of an ester bond between aminoacids Ser-254 and Gly-253 and, therefore, both peptide chains which provide then the active form of the Psd protein (Matsumoto et al., 1998; Voelker, 1997). It has been reported that mutations in the aminoacid Serine-254 leads to a huge decrease in the activity or, even, an inactive protein leading to the cell death (Li and Dowhan, 1990).

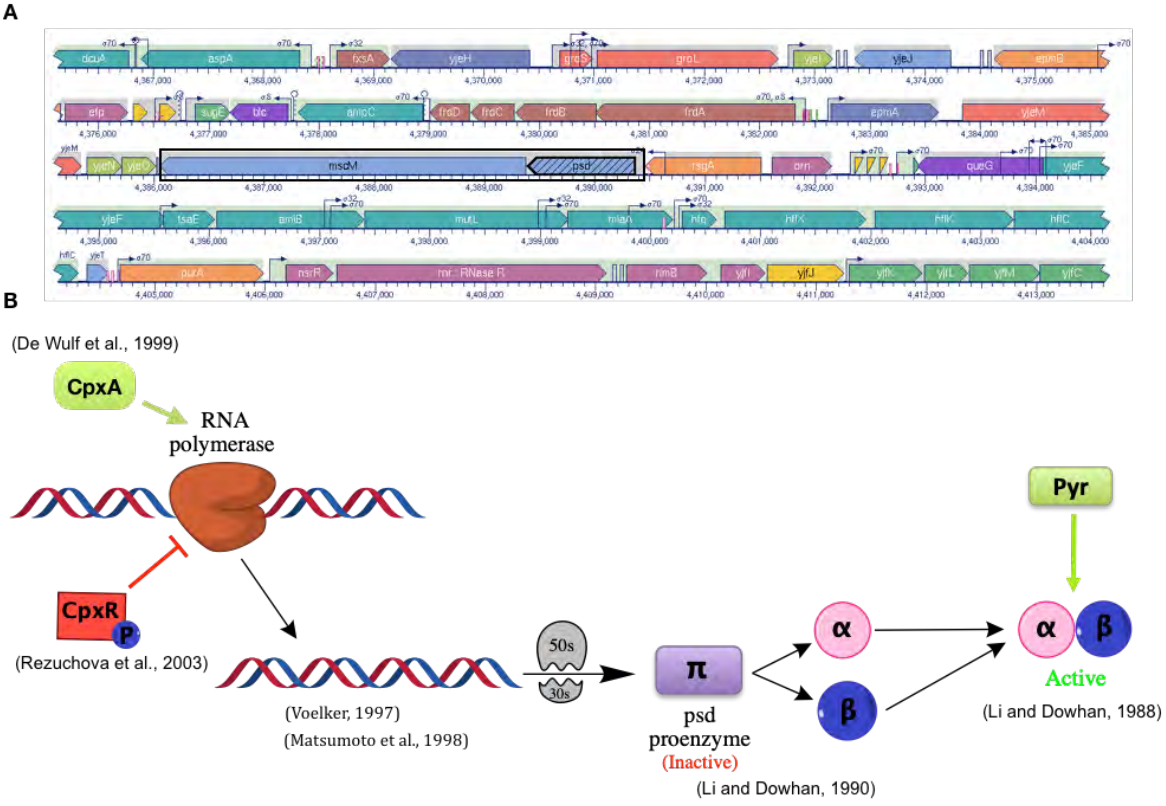

Figure 76. Information about psd (A) Operon of psd (B) Regulation of psd

11.3. Design of the library

The design of this library was based on the work performed by Prof. Marjan de Mey's group at the Ghent University and with her collaboration (Bauwens, 2019). The library designed by this group

was focused on screening the library by fluorescence, which can be followed as a high throughput screening assay (HTS). The idea was based on interchanging the genomic promoter and RBS sites by new designed ones, which would allow a different range of expression. Along with these new sites, and to make sure that all the possible expression coming from the natural promoter was stopped, a strong terminator (T7) was also added to the design previously to the promoters and RBS sites (Figure 77).

To screen the library, they developed a method based on the ability of cI protein, a transcriptional repressor from lambda phage, to repress the transcription of different genes. In this case, this protein would repress the lambda promoter in charge of the expression of a fluorescent protein (mKATE2). This means that it was required for this method to have a plasmid containing this promoter and the fluorescent protein, which from now on is going to be referred to as screening plasmid.

Moreover, to be sure that the expression of this cI protein is the same as the gene of interest making possible to relate the fluorescence levels with the expression of the protein of interest, a coupled transcription of *psd* to *cI* was designed. This means that inside the sequence of *cI*, which is the first gene to be transcribed, there was the promoter for the second gene being this one, *psd*. Finally, to be able to select those cells, which introduced this DNA inside its genome, a chloramphenicol resistance, flanked by FRT sites, was also added. All this was included in the Knock-in (KI) plasmid.

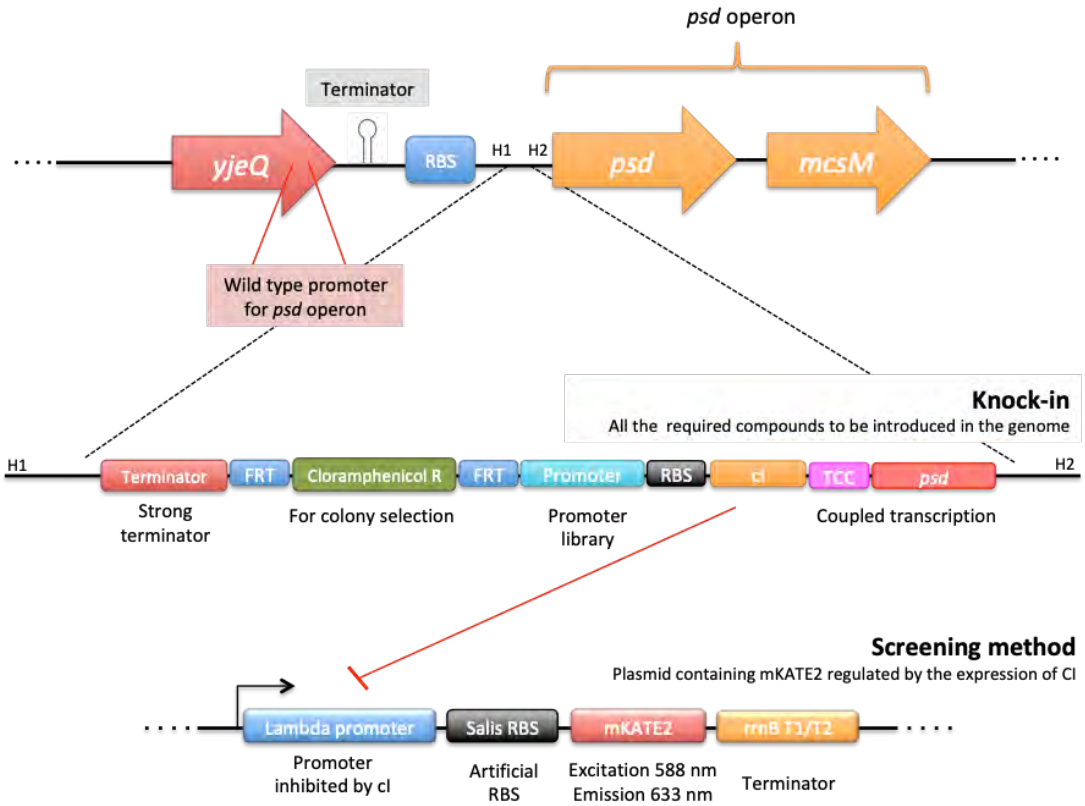

Figure 77. Design of the promoter and RBS library for psd and its screening system

To introduce this DNA into the genome, as it happened with the knockouts, the Datsenko and Wanner approach was used to interchange the designed DNA with the genomic one of *E. coli*

including this way the new promoter and RBS sites along with the screening method. If cells when transformed with this DNA correctly incorporated it in its genome, due to the chloramphenicol resistance cassette, the selection of this colonies would be possible.
Once these cells were plated, different ranges of pink were expected regarding to the expression of *cI* and, therefore, *psd*. Since *cI* represses the expression of mKATE2, if a lot of fluorescence was detected, it meant that *cI* was not highly expressed and, therefore, a low expression of *psd* as well.

11.4. Obtaining the KI plasmid

All these genetic compounds previously mentioned required for the correct function of the library that was placed in a plasmid designed by Prof. Marjan de Mey's group (Bauwens, 2019). Once *psd* gene was subcloned after *cI*, maintaining the coupled transcription, it would be possible to obtain linear DNA to be used for recombining it with the genomic one.
As Figure 77 shows after *cI* sequence it was necessary to subclone *psd*. To do so a Golden Gate Assembly approach was followed using the primers placed in Table 30 (Engler et al., 2008). This type of assembly is based on the ability of type II restriction enzymes to cut after the recognition sequence and the design of specific overlapping sequences of 4 nucleotides.

Table 30. Primers designed to subclone psd into KI plasmid by GGA.

Fragment	Template used	Primers	Length of the fragment (bp)
Backbone	pSC101	fw_vector_psd_gg	2813
		rv_ vector_psd_gg	
psd	Δ*tesA* genome	fw_psd_gg	997
		rv_ psd_gg	

After obtaining the fragments required for the assembly being one the linearized plasmid and the other *psd* gene extracted from Δ*tesA* genome, two different types of transformation were performed. First, electroporation in 50 µL of TOP10 electrocompetent cells was done by using 2 µL of the GGA reaction mixture. After one hour of incubation, 60 and 120 µL of culture were plated onto chloramphenicol plates. In parallel, 2 µL of the same reaction mixture were transformed into DH5α chemical competent cells. The heat shock was performed at 42ºC for 20 seconds followed by 45 minutes of incubation at 37ºC. As it was done with electroporation, 60 and 120 µL of the culture were plated and grown overnight at 37ºC.
The next day, 28 colonies were analyzed by PCR colony using oMEMO1617 and oMEMO8369 as primers where the expected product length was around 1.2 kb if the construct contained *psd* (Figure 78).
As it can be seen, almost all the colonies tested seemed to have a band at the desired length but only colonies 1, 3, 5 and 28 were sent for sequencing (arrow) reporting that only 1 and 28 contained the desired construct. Plasmid from colony 28 was selected to be used for the construction of the library and it was named as KI plasmid.

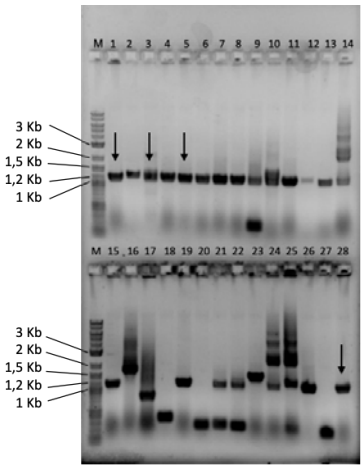

Figure 78. PCR colony from the colonies obtained by GGA. Arrows indicate the colonies sent for sequencing

11.5. Obtaining the library

To obtain this library first of all it was required to create the promoter library and RBS so, after linearizing all the required genetic material, it could be transformed into *ΔtesA* strain (as major producer) and test the range of expression in *psd*.

Based on the work of Coussement et al it was necessary to linearize the KI plasmid deleting at the same time the promoter and RBS sites so these could be added in the primers where they were randomized (Coussement et al., 2017). Figure 79 schematizes the procedure known as Single Strain Annealing (SSA), where it can be seen that when two fragments are meant to be added one primer will contain the library of promoters (in our case randomized N) while the other primer will contain the sequence of the RBS designed by SALIS where a range of expression was expected (Salis et al., 2009). Both primers needed to have homologous sequences to anneal to the linearized plasmid and to the other primer so finally, it would be possible to circulize the plasmid again by CPEC. Once the plasmid was obtained, a library of promoters and RBS was available.

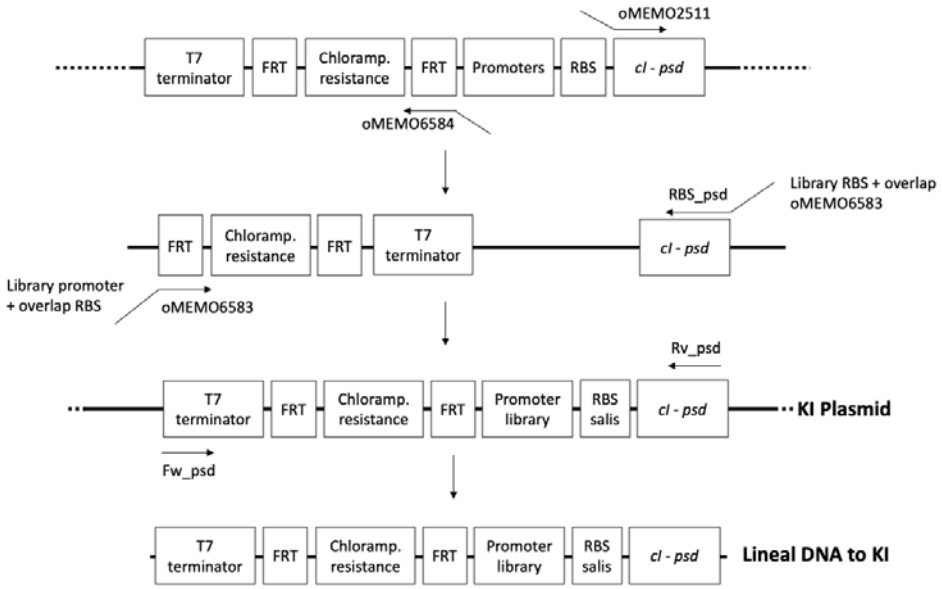

Figure 79. Scheme to obtain the linear DNA to perform the KI

Once this plasmid was obtained, by using a different set of primers, fw_psd and rv_psd, it was possible to linearize the DNA that was required to perform the knock in (Figure 79). The procedure to introduce this linear DNA into the genome of *E. col*i was the same as if the first step of the KO was performed. This means that cells were transformed with pKD46 and, since it was necessary to have the plasmid containing mKATE2 in order to screen the library, it was also required to have the same cells transformed with the screening plasmid, which contained both promoter lambda and mKATE2 (Screening plasmid). Once the linear DNA was transformed into these cells containing both plasmids, the recombinase promoted the recombination between this DNA, which contained homologous sequences at the extremes, and the genome. Thanks to the chloramphenicol resistance cassette it was possible to select those colonies that inserted this foreign DNA containing the promoter and RBS library along with the screening method. Since the content of phosphatidylethanolamine in these cells could be low, it was necessary to supplement these kanamycin and chloramphenicol plates with 50 mM of $MgCl_2$ to ensure the growth of these cells. The whole procedure followed is summarized in figure 80.

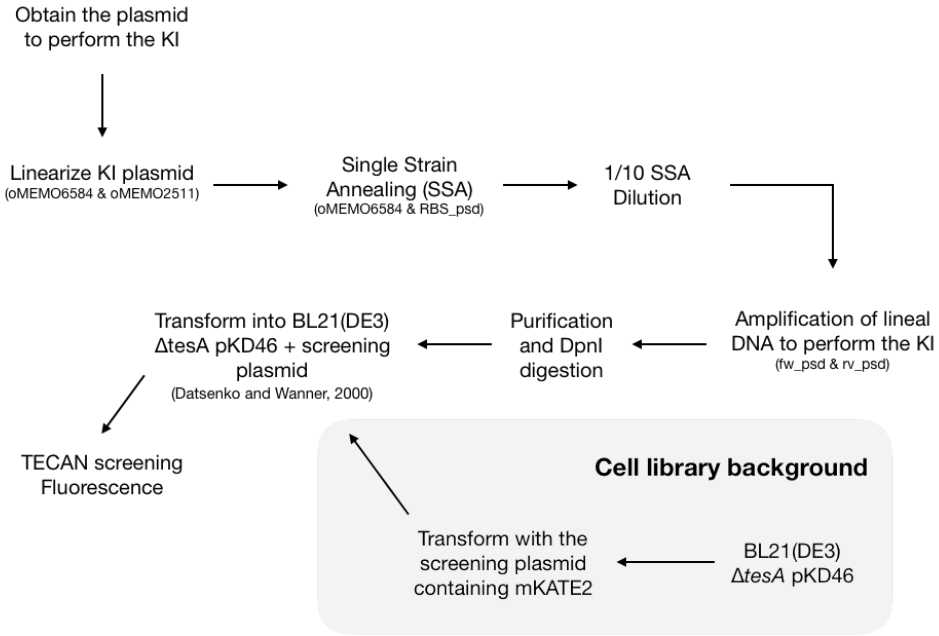

Figure 80. Summary of the procedure to obtain the library

11.6. Library characterization

Once the colonies were obtained showing different ranges of pink an analysis of 264 colonies was performed. The first analyses performed monitored the growth (Figure 81A) showing that almost all the colonies presented the same growth profile with an exponential phase between the first and the fifth hour after initializing the culture. Furthermore, these same colonies were analyzed following the fluorescence emission (633 nm) of mKATE2 protein along time. To have controls and have a reference Δ*tesA* without the screening plasmid (*ΔtesA*) and Δ*tesA* transformed with this plasmid (Δ*tesA*+) were analyzed. Results showed that a wide range of fluorescence was detected indicating the correct randomization of the library (Figure 81B).

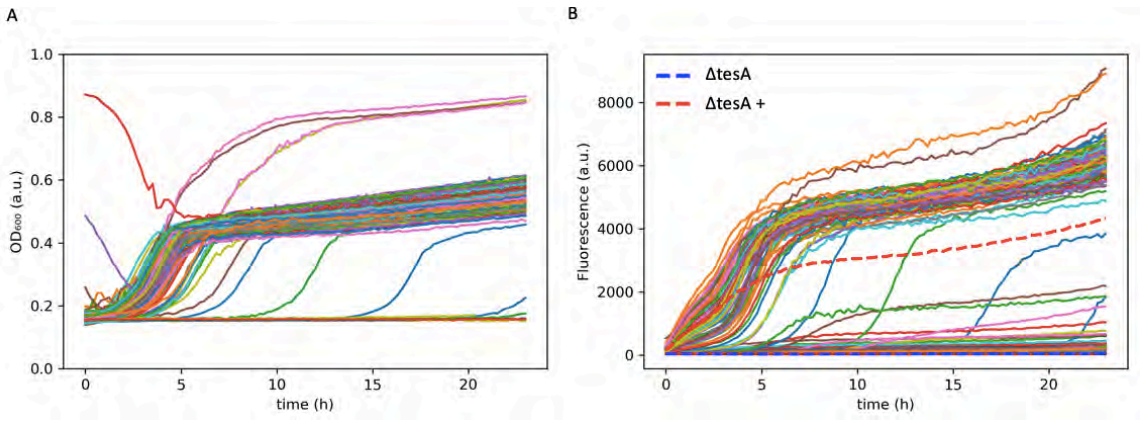

Figure 81. Monitorization of the growth and fluorescence overtime of 264 colonies. (A) Growth (B) Fluorescence

To normalize the results, the fluorescence and growth parameters were graphed and using Δ*tesA* and Δ*tesA*+ strains a metabolic burden of the plasmid indicator was calculated. This metabolic growth allowed us to select those colonies that were interesting to produce glycoglycerolipids. It has to be remembered that those strains which present higher fluorescence are the interesting ones since it means that those strains incorporated weaker promoters than the Δ*tesA* original strain and can produce lower amounts of PE.

As figure 82 shows, all the colonies were surrounding the linear regression. The cluster of cells with high fluorescence/growth make us think that there could be a large number of false positives. Nevertheless, eleven colonies were selected (green) to further analyze the genome and also check how the fluorescence and growth behavior.

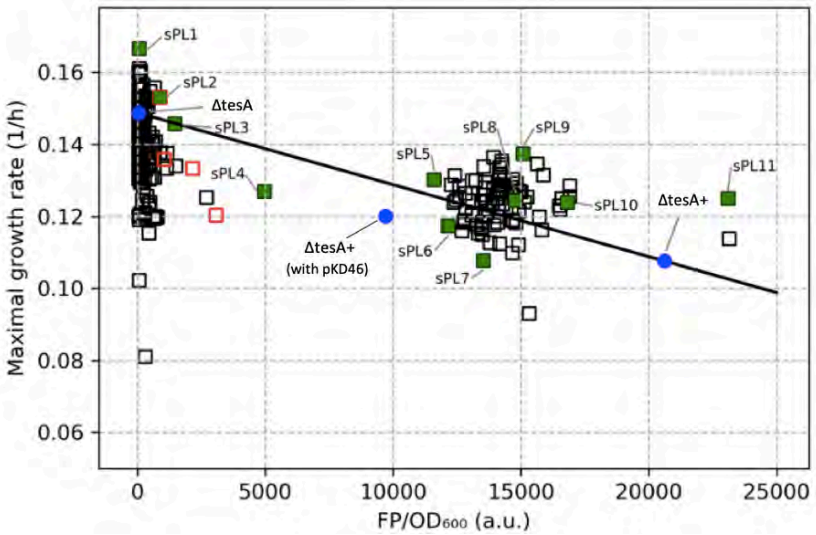

Figure 82. Fluorescence/growth vs maximal growth rate

The colonies in green are the ones selected for further analysis while the blue ones are the controls and the red ones are three colonies that reported higher growths in comparison to the other strains.

As Figure 83 shows, the strain Δ*tesA*, which does not contain the screening plasmid and therefore, does not contain mKATE2, does not show fluorescence. The maximal growth rate is slightly higher

than in the strains of Δ*tesA* that contained the screening plasmid. Between Δ*tesA*+ and Δ*tesA*+_pKD, strain which has the plasmid pKD46, required for the insertion of the linear DNA to remove the native promoter and RBS, it can be seen that there is a decrease in the fluorescence levels. This indicates us that there is an interference in the fluorescence levels when the strain contains pKD46 plasmid.

Regarding the analyzed strains, sPL1, sPL2, sPL3 and sPL4, which were nearby Δ*tesA* strain in figure 83, did not report fluorescence. This result indicates us that these strains probably incorporated a strong promoter which led to a high expression of cI protein that repressed the lambda promoter and, therefore, the expression of mKATE2 protein. This suggest that the strong promoter and RBS will lead to higher expression of Psd and production of PE. On the other hand, strains sPL5 to sPL11 showed high fluorescence levels close to the maximal fluorescence expression of Δ*tesA*+ (BL21 Star (DE3) Δ*tesA* transformed with indicator plasmid). This result means that these strains probably contain weak promoters and RBS that produce less cI-Psd proteins and therefore the expression of mKATE2 is lower. Thus, these strains are very interesting for the project because this high range of fluorescence suggest us that lower amounts of PE are produced. This combined with the previous modifications reported in this project can lead to an increase of the GGL production.

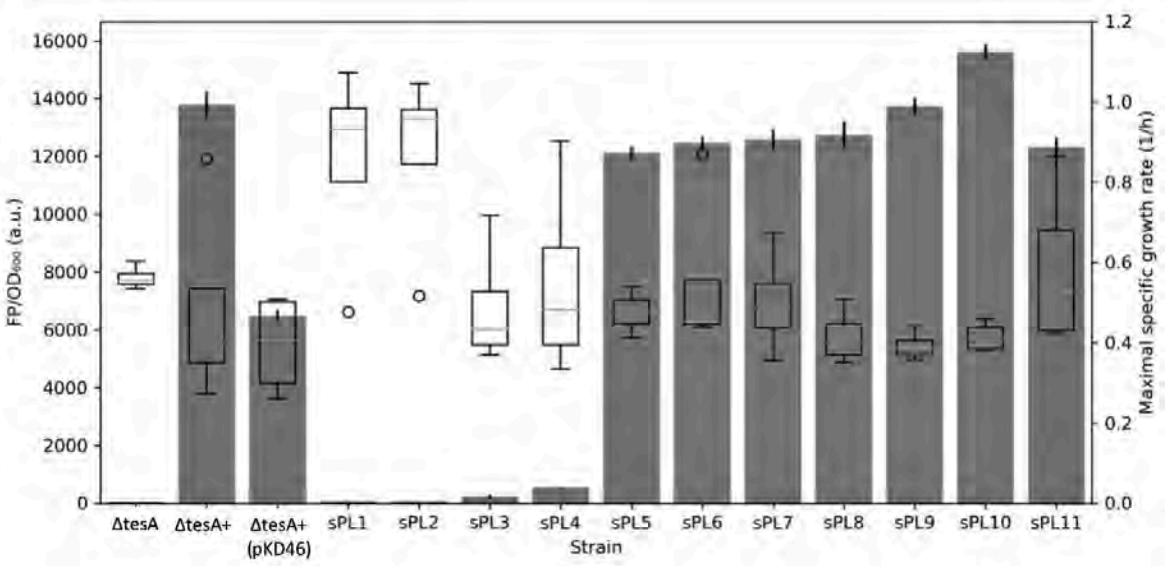

Figure 83. Strain characterization by fluorescence determination and growth rate

These results were very promising but unfortunately a recombination event happened at some point during the analyses and this library was removed from the genome of the cells. In further studies, the process should be repeated again. In the new selected colonies, more studies more studies to characterize the promoter and RBS sequences, quantify the amount of PE and relate it with the glycoglycerolipids production should be performed.

This strain characterization was performed in collaboration with A. Peeters in the group of Prof. Marjan de Mey (University of Ghent).

CONCLUSIONS

12. Conclusions

- Focusing on the availability of the diacylglycerol lipid precursor, the highest GGL production was achieved when overexpressing MG517-PlsC^{H} and removing *tesA* (#2) from the genome of *Escherichia coli*, leading to an almost 2-fold increase in production yield when compared to the WT strain (#1). On the other hand, no increase in the production was observed when *fadE* (#3), a gene involved in the ß-oxidation, was knocked out. Finally, a change in the fatty acid profile was observed when *fabR* (#5 strain) was removed from the genome, resulting in an increase of unsaturated fatty acids titer.
- The highest conversion of acyl donor to phosphatidic acid was obtained by overexpressing PlsC and PlsB acyltransferases in high copy number plasmids (#2, #12 and #13) or coexpressing both acyltransferases in a low copy number polycistronic plasmid (#11)
- A limitation of UDP-Glc, precursor of glycoglycerolipids, was observed in the new engineered strain ΔtesA (#2) that was overcome by the overexpression of GalU (#16) or by knocking out *ushA* gene (#18). In both cases, the production was increased almost 3-fold when compared to WT strain (#1).
- A change in the glycolipids profile was observed when CDH enzyme was expressed in the engineered strains (#14 and #22 strains) even though no increase in the overall production was reported.
- A 3-fold increase in the glycoglycerolipid production was observed using phosphatidylglycerophosphatase B fused to PlsC acyltransferase in ΔtesA strain (#15) although no increase was observed in ΔtesA ΔushA (#23) or ΔtesA overexpressing this new fusion protein along with GalU (#17).
- An increase in the unsaturated fatty acid profile was reported in those strains that either had *fabR* knocked out or were overexpressing *fadR.* Nevertheless, this change was not translated into an increase in the production of GGL (#5, #7 and #21 strains).
- As previously shown in previous studies in our group, heterologous production of non-native glycolipids in *E. coli* results in a decrease of phosphatidylethanolamine levels to maintain the charge balance of the plasma membrane.
- A range of phosphatidylserine decarboxylase expression levels, enzyme responsible for the synthesis of phosphatidylethanolamine, was achieved by designing a library of promoters and RBS. The impact of the expression level in glycoglycerolipid production will be analyzed in further studies.

CONCLUDING HIGHLIGHTS:

1. Metabolically engineered E. coli strains producing glycoglycerolipids (GGL) are reported.

2. Introduction of the glycosyltransferase MG517 from *Mycoplasma genitalium* transfers UDP-Glc onto diacylglycerol (DAG) to generate mono- to tetra-glucosyldiacylglycerols.

3. Knockouts of genes involved in fatty acid degradation and overexpression of genes in the phosphatidic acid biosynthesis pathway boost DAG availability for GGL production.

4. UDP-Glc does not become limiting in the engineered strains.

5. A thioesterase *tesA* KO strain harboring plasmids overexpressing MG517 and a fusion PlsCxPgpB protein (acyltransferase fused to phosphatidic acid phosphatase) is the best GGL producer with a titer of 9.9 nmol GGL per mg of dry cells.

6. GGLs replace phosphatidylethanolamine in the *E. coli* membrane to balance the overall membrane charge.

MATERIAL AND METHODS

13. Material and methods

In this chapter all the information related with the experimental part of the project is described.

13.1. Molecular biology

13.1.1. Genomic DNA extraction and purification

To obtain the genomic DNA to extract genes to be used for further studies it was required to grow a 2 mL preinocculum of the strain of interest, in this project *E. coli* JM109, *E. coli* MG1655 and BL21 Star (DE3) Δ*tesA* at 37ºC.

Once this preinocculum was at stationary phase, it was harvested, rinsed twice with 0.9% NaCl, harvested and then resuspended with 0.5 mL of MilliQ sterile water. Afterwards, it was required to boil it between 90 – 100ºC for 10 minutes, centrifuged again and the supernatant was conserved. In this supernatant is where the genomic DNA was expected to be.

13.1.2. Plasmid DNA extraction and purification

To obtain the plasmid DNA different commercial kits of extraction were used. Before extracting it was required to grow an overnight culture of 6 mL of LB and the required antibiotic with the correspondent strain.

13.1.2.1. High Purity Plasmid Miniprep kit (Clinisciences, Ref. NB-03-002)

The protocol was extracted from the website (Clinisciences, n.d.)

13.1.2.2. InnuPREP Plasmid Mini Kit 2.0 (AnalytikJena, Ref. 845-KS-5041250)

The protocol used was provided by the provider.

13.1.3. DNA Gel extraction

To purify specific bands of DNA from the PCR samples amplified it was required to use GenElute™ Gel Extraction Kit (Sigma Aldrich®) or innuPREP Gel Extraction Kit (Analytik Jena®) following the protocol supplied by the providers.

1% agarose gels were performed to separate the different bands and be able to excise them.

13.1.4. DNA quantification

Two different methods were used to quantify DNA

13.1.4.1. Qubit

This method was based on the capacity of the fluorescent proofs provided by Life Technologies (Thermo Fisher®) to bind to the DNA. Two different kits were available regarding to the expected amount of DNA being High sensitivity the one used to quantify between 10 pg/µL and 100 ng/µL whereas Broad Range (BR) was used to quantify samples from 100 pg/µL to 1000 ng/µL. High sensitivity was used for samples coming from gel extraction or low copy number plasmids while BR was used for samples coming from PCR or high copy number plasmids.

The protocol followed can be found at www.thermofisher.com

13.1.4.2. Nanodrop

To quantify DNA concentration, 2 µL of DNA samples were loaded into the detector and after pressing measure the software gave back the result of the concentration along with the grade of purity of the sample. The blank for the measurement was MilliQ water for those samples that were eluted with this or elution buffer for those samples eluted using this buffer, e.g. plasmids.

13.1.5. Subcloning techniques

13.1.5.1. Gibson assembly

1. PCR amplifying the vector and insert(s), ensuring that at least 20 bp of overlapping sequences exist. Determine DNA concentration (ng/µL) of each assembly piece
2. Thaw assembly master mix (table 29) and keep on ice until sample are ready to be used
3. Mix 15 µL of assembly master mix with the desired amount of each assembly piece (in equimolar amount) without overpassing 22 µL of total volume
4. Incubate at 50°C for 15-60 min
5. Transform cells with 2 µL of assembly mixture and plate in desired antibiotic resistance plates after the required incubation time

Table 31. Mix composition for the Gibson Assembly reaction mixture

Compound	Volume (µL)
5x ISO buffer	320
T5 exonuclease (1000 U/mL)	0.64
Q5 High fidelity polymerase	20
Taq DNA ligase (40000 U/mL)	160
Water (milliQ)	699.36
Total Volume	1200 (aliquot to 15 µL in PCR tubes)

Table 32. 5x ISO buffer preparation

Compound	Volume (µL)
1 M Tris-HCL pH 7.5	3000
2M MgCl2	150
100 mM dNTP mix	240
1M DIT	300
100mM NAD	300
PEG - 8000	1.5 g
Water (milliQ)	2010
Total volume	6000

13.1.5.2. CPEC assembly

1. PCR amplifying the vector and insert(s), ensuring that at least 20 bp of overlapping sequences exist. Determine DNA concentration (ng/µL) of each assembly piece

2. Add 100 ng of the linearized vector backbone and equimolar amounts of the other assembly pieces up to a 25 µL total volume assembly reaction mixture as table 31 shows

Table 33. Gibson Assembly reaction mix, volumes

Compound	Volume (µL)
linearized vector backbone (100 ng)	X
Assembly piece	Y
5X HF Q5 Reaction Buffer	5
dNTP 2mM	5
DMSO	0.75
2U/ml Q5 Polymerase	0.5
Water (milliQ)	up to 25

3. Perform PCR incubation (15 cycles) at desired annealing and elongation temperature, elongation time may change depending on the polymerase used
4. Transform 2 µL of CPEC mixture into Top10 electrocompetent cells and plate in desired antibiotic resistance plates

Note: If Iproof 2x master mix is used instead of Q5 polymerase, the volume required for 25 µL of reaction is 12.5 µL.

13.1.5.3. Golden Gate Assembly

1. Linearize the vector and insert fragments by PCR, adding at the same time the BsaI sites and homologous sequences at 5' of the primer
2. Add 100 ng of the linearized vector backbone and the required equimolar amount of each assembly pieces up to a 20 µL of total volume assembly reaction mix as table 32 shows

Table 34. Golden Gate Assembly reaction mix

Compound	Volume (µL)
Linearized vector backbone (100 ng)	X
Assembly piece	Y
BsaI	1
T4 ligase	1
T4 ligase buffer	2
BSA	2
Water (milliQ)	up to 20

3. Perform PCR incubation with the following cycles and temperatures (table 33):

Table 35. PCR conditions for GGA

Cycles	Time (min)	Temperature (ºC)
50	5	37
	5	16
1	5	37
1	10	50

Cycles	Time (min)	Temperature (ºC)
1	10	80
	∞	16

4. Transform 2 µL of GGA mixture into Top10 electrocompetent cells and plate into the desired antibiotic resistance plates

13.1.5.4. Single Strain Annealing

1. Linearize the plasmid by using oMEMO2511 and oMEMO6584
2. Purify the reaction mixture and digest it with DpnI for 3h at 37ºC
3. Amplify the linearized vector with oMEMO6583 and RBS_psd

13.1.6. Weight marker

The marker used as for the agarose gels were:

- 2 log ladder from New England Biolabs® (this product was renamed as 1 kb plus DNA ladder). To prepare this ladder the instructions provided by NEB were followed and only 4 µL of ladder were loaded into the gel.
- Marker III from Roche®, prepared as the provider indicated. 12 µL of samples were loaded in each gel.

13.1.7. Competent cells

Different methods and protocols were used to prepare competent cells.

13.1.7.1. Electrocompetent cells

The protocols used to prepare electrocompetent cells and the Datsenko and Wanner methods were provided by Metabolic Engineering and Modelling of Microorganisms group from Ghent University.

13.1.7.1.1. Electrocompetent cells to maintain in-80°C

1. Cultivate 4 flasks of target cells 50 mL/flask to OD600 = 0.6, chill on ice for 20 min.
2. Centrifuge 20 min at 1000g, 4°C. Remove the media by careful decanting
3. Resuspend in 50 mL chilled 10% glycerol, centrifuge. Remove the media using a 50 mL sterile pipette
4. Resuspend in 25 mL chilled 10% glycerol, centrifuge. Remove the media using a sterile 25mL pipette
5. Resuspend in 10 mL chilled 10% glycerol, centrifuge.
6. Resuspend in 1 mL chilled 10% glycerol, centrifuge. Remove the media using a 1mL sterile pipette
7. Resuspend in 400 µL 10% glycerol (total 1600 µL, divide into 40 µL/vial). Store at -80°C. Resuspend in 200 µL for linear DNA transformation.

   Note: Always keep the cells in ice

13.1.7.1.2. Fresh electrocompetent cells

1. Prepare a seed culture of the target cells
2. Inoculate 10 mL LB (+antibiotics, inducer) in a 50 mL falcon tube with the seed culture in order to have an OD 660 nm between 0,05 and 0,1 (~0.5 ml)
3. Incubate at the appropriate temperature till OD600 nm 0.6
4. Transfer 10 ml into a chilled falcon tube and place on ice during 30 min. From now on it's important to keep everything cold! Use chilled DI water, chilled recipients and tips!
5. Centrifuge 4 min at 5500 rpm (4 ºC)
6. Resuspend the pellet in 45 ml cold sterile DI water and centrifuge again during 5 min at 5500 rpm (4 ºC)
7. Resuspend the pellet in 1 ml cold sterile DI water and transfer into a cold Eppendorf
8. Centrifuge cold at maximum speed during 20 - 30 seconds
9. Resuspend the pellet in 50 µl cold sterile DI water by moving the pipet point around (do not pipet up and down!)
10. Transfer 50 µl of the suspension into a new cold Eppendorf, what 's left over will be used as a control.

13.1.7.2. Chemical competent cells

The protocol used to make chemical competent cells was extracted from the protocols of Bioengineering master.

1. Cultivate 4 flasks of target cells (50 mL/flask) until it reach and OD600 of 0.6, chill on ice for 20 min.
2. Centrifuge for 5 min at 5000 g at 4°C. Remove the media by decanting carefully.
3. Resuspend in 15 mL chilled sterile 50mM $CaCl_2$.
4. Add 750 µL sterile glycerol to cell suspension until it reaches 20% of glycerol and aliquot 200 µL of this cells into sterile Eppendorf's.
5. Store them at -80°C.

<u>Transformation protocol:</u>

1. Thaw competent cells on ice.
2. Add 50 ng of DNA to 200 µl cells.
3. For negative control, add 5µl of sterile TE buffer or sterile water to 200µl.
4. In case you have a known and well-characterized plasmid, use it as a positive control and add similar DNA concentration to 200µl cells.
5. Incubate 30 min on ice
6. Incubate cells 2 min at 42°C in water bath.
7. Transfer the cells for 5 min on ice.
8. Add 500 µl LB medium to cells and incubate for 45 min on shaker (250 rpm) at 37 °C.

If DNA comes from DNA plasmid extraction:

9. Plate 20 µl on LB plate with ampicillin. Let absorb liquid.

10. Prepare 1/10 dilution of the transformant cells: take 10 µl and add 90 µl of LB. Plate 20 µl on LB plate with ampicillin. Let the liquid absorb.
11. Incubate plates at 37 °C for 12h.
12. As a control plate (in LB plates without ampicillin) non-transformed competent cells, following the same dilutions and incubations steps described above.

If DNA comes from ligation:

9. Plate 200 µl on LB plate with ampicillin. Let absorb the liquid.

13.1.8. DNA digestion

Different digestions of DNA were performed throughout the project

13.1.8.1. DpnI

Using DpnI restriction enzyme was destined to get rid of the methylated DNA used for PCRs.
All the reactions done had 20 µL of total volume. 1 µL of Dpn-HF (NEB), 2 µL of CutSmart and up to 16 µL of sample.

13.1.8.2. Double digestion

To perform this digestion, it was required to use EcoRI-HF and NcoI-HF restriction enzymes. Both enzymes were able to reach a 100% activity by using CutSmart reaction buffer. The conditions used for the digestion can be seen in Table 34.

Table 36. Conditions for the double digestion using NcoI-HF and EcoRI-HF enzymes

DNA sample	16 µL
CutSmart 10x	2 µL
EcoRI-HF	1 µL
NcoI-HF	1 µL
Total volume	20 µL

The reaction was performed at 37ºC for 3 hours and then was stopped as providers indicated.

13.1.9. Ligation

The ligation procedure followed was based on the ability of T4 DNA ligase to form the phosphodiester bond between the joined fragments. New England Biolabs provided this ligase. The ligation procedure followed was based on the indications of the provided and a 1:3 relation of pRSF-1b vector and PlsB gene were used.

13.1.10. Datsenko and Wanner protocol to obtain KO strains

Protocol adapted from Datsenko & Wanner, PNAS 2000 and provided by MEMO group (Ghent University).

13.1.10.1. Procedure: transform with pKD46

1. Plate the strain out on a LB-plate and incubate overnight at 37°C.
2. Suspend 1 colony in 4-5 ml LB and incubate overnight at 37°C on a shaker.
3. Add around 500 µl of the overnight culture to 10 ml LB in order to have an OD 660 nm between 0,05 and 0,1

4. Incubate at 37 °C till OD600nm 0,6
5. Transfer 10 ml into a chilled falcon tube and place on ice during 30 min.
1. From now on it‘s important to keep everything cold! Use chilled tips and recipients!
6. Centrifuge 4 min at 5500 rpm (4 ºC)
7. Resuspend the pellet in 45 ml cold DI water and centrifuge again during 5 min at 5500 rpm (4 ºC)
8. Resuspend the pellet in 1 ml cold DI water and transfer into a cold Eppendorf
9. Centrifuge at maximum speed during 20 - 30 seconds (cold)
10. Resuspend the pellet in 50 µl cold DI water by moving the pipet point around (do not pipet up and down!)
11. Transfer 50 µl of the suspension into a new cold Eppendorf, what ‘s left over will be used as a control
12. Add ~3 µl pKD46 (50ng) and mix gently by turning
13. Transform the strain by electroporation with the Gene Pulser:
    a. Transfer the sample to the gene Pulser Cuvette using a cold tip
    b. Tap gently to make sure that the sample is evenly distributed between the sides of the Cuvette. Also take care that everything is on the bottom of the cuvette.
    c. Insert the cuvette into the chamber slide and push the chamber slide into the chamber until the cuvette is seated between the contacts in the base of the chamber
    d. Pulse once by pressing the 2 red buttons at the same time until the Gene Pulser beeps.
14. Add immediately 1 ml SOC to the cells and transfer into a new tube
15. Also add 1 ml SOC to what is left of the washed cells (= control, see above)
16. Incubate the transformed cells and the control during 1 h at 30 °C on a shaker
17. To obtain single colonies we will plate out our transformed cells in different dilutions on an ampicillin plate (100µg/ml): transfer 100 µl of the culture, dilution 10-1 and 10-2 to the plate and streak out.
18. Incubate overnight at 30 ºC
19. Make 3 precultures in LB+Amp of each time 1 colony and incubate over night at 30°C
20. Transfer of each culture 1 ml into a glycerol tube to store at -80°C and use 1 culture to do transformation 2

13.1.10.2. Procedure: insert linear DNA

1. Add around 800 µl of the overnight culture to 25 ml LB+Amp in order to have an OD 660 nm between 0,05 and 0,1 (without blank). Also add 500 µl 500 mM L-arabinose solution (200 µl / 10 ml culture)
2. Incubate at 30 °C till OD 0,6
3. Transfer 10 ml into a falcon tube and place on ice during 30 min.
1. From now on it‘s important to keep everything cold! Use chilled tips and recipients!
4. Centrifuge 4 min at 5500 rpm (4 ºC)
5. Resuspend the pellet in 45 ml cold water and centrifuge again during 5 min at 5500 rpm (4 ºC)
6. Resuspend the pellet in 1 ml cold water and transfer into a cold Eppendorf
7. Centrifuge at maximum speed during 20 - 30 seconds

8. Resuspend the pellet in 50 µl cold water by moving the pipet point around (do not pipet up and down!)
9. Transfer 50 µl of the suspension into a new cold Eppendorf, what ‘s left over will be used as the control
10. Add 8-10 µl of the DNA (500ng-1µg) and mix gently by turning
11. Transform the strain by electroporation with the Gene Pulser:
    a. Transfer the sample to the gene Pulser Cuvette using a cold tip
    b. Tap gently to make sure that the sample is evenly distributed between the sides of the Cuvette. Also take care that everything is on the bottom of the cuvette.
    c. Insert the cuvette into the chamber slide and push the chamber slide into the chamber until the cuvette is seated between the contacts in the base of the chamber
    d. Pulse once by pressing the 2 red buttons at the same time until the Gene Pulser beeps.
12. Add immediately 1 ml SOC to the cells and transfer into a new tube
13. Also add 1 ml SOC to what is left of the washed cells (= control, see above)
14. Incubate the transformed cells and the control during 2 h at 37 °C on a shaker
15. Transfer 200 µl of the culture on a Chloramphenicol (250 µg/ml) plate or a Kanamycin (50 µg/ml) plate.
16. Centrifuge the rest of the culture for 1 min at 13.000 rpm, resuspend the pellet in 200 µl LB and spread the entire sample on a second plate
17. Incubate overnight at 37 ºC
18. Perform a colony PCR using out-primers on as much colonies as possible, strike them out on Kanamycin or Chloramphenicol plates.
19. Check the colony PCR on an analytical gel
20. Take 3 positives and plate them out on the corresponding selective plate, incubate at 42 °C
21. Check 3 colonies of each plate again with colony PCR using the out-primers
22. Take 3 positives and make precultures at 37°C
23. Transfer of each culture 1 ml into a glycerol tube to store at -80°C and use 1 culture to do transformation 3

13.1.10.3. Procedure: removal antibiotic resistance cassette

1. Add around 500 µl of the overnight culture to 10 ml LB + Chl/Kan in order to have an OD 660 nm between 0,05 and 0,1
2. Incubate at 37 °C till OD600nm 0,6
3. Transfer 10 ml into a falcon tube and place on ice during 30 min.
   From now on it‘s important to keep everything cold! Use chilled tips and recipients!
4. Centrifuge 4 min at 5500 rpm (4 ºC)
5. Resuspend the pellet in 45 ml cold water and centrifuge again during 5 min at 5500 rpm (4 ºC)
6. Resuspend the pellet in 1 ml cold water and transfer into a cold Eppendorf
7. Centrifuge at maximum speed during 20 - 30 seconds
8. Resuspend the pellet in 50 µl cold water by moving the pipet point around (do not pipet up and down!)

9. Transfer 50 µl of the suspension into a new cold Eppendorf, what 's left over will be used as the control
10. Add ~4 µl pCP20 (50ng) and mix gently by turning
11. Transform the strain by electroporation with the Gene Pulser:
    a. Transfer the sample to the gene Pulser Cuvette using a cold tip
    b. Tap gently to make sure that the sample is evenly distributed between the sides of the Cuvette. Also take care that everything is on the bottom of the cuvette.
    c. Insert the cuvette into the chamber slide and push the chamber slide into the chamber until the cuvette is seated between the contacts in the base of the chamber
    d. Pulse once by pressing the 2 red buttons at the same time until the Gene Pulser beeps.
12. Add immediately 1 ml SOC to the cells and transfer into a new tube
13. Also add 1 ml SOC to what is left of the washed cells (= control, see above)
14. Incubate the transformed cells and the control during 1 h at 30 °C on a shaker
15. Make following dilutions 100, 10-1,10-2 in LB
16. Transfer 100 µl of each dilution on a LB plate with Ampicillin
17. Incubate overnight at 30 ºC
18. Perform a PCR on a few colonies (ca 10), strike them out on LB plates
19. Plate 3 positives out on LB plates and put them at 42 °C
20. Check 3 colonies of each plate again with colony PCR using out primers,. Prepare for 1 colony 100 µl HiFi PCR mix, this can be sent to sequence after PCR purification.
21. Take 3 positives and make precultures at 37°C
22. Transfer of each culture 1 ml into a glycerol tube to store at -80°C and use 1 culture to do transformation 1 if necessary

13.2. Cultures

13.2.1. Preinocculum

In sterile tubes of 15 mL, 3 or 6 mL of LB broth with antibiotics (if required) are added. One unique colony is picked from plates and grown overnight at the desired temperature.

13.2.2. Minimal medium

Dilution stocks for the minimal medium were prepared using distilled water.

Table 37. Stock for the minimal medium preparation

Medium		Concentration	Sterilization procedure
Glucose		20% (w/v)	Autoclave (120ºC for 20 min)
$MgSO_4$		0.2 M	Filtration 0.22 µm
$FeCl_3$		30 mM	Filtration 0.22 µm
Thiamine		1 mg/mL	Filtration 0.22 µm
MES 5x	**$Na_2HPO_4 \cdot 2H_2O$**	0.42 M	Autoclave (120ºC for 20 min)
	NaH_2PO_4	0.23 M	

Medium		Concentration	Sterilization procedure
	NaCl	0.21 M	
	$(NH_4)_2SO_4$	0.18 M	
TES	**$AlCl_3 \cdot 6H_2O$**	0.17 mM	Filtration 0.22 µm (pH 2)
	$CoCl_3 \cdot 6H_2O$	0.59 mM	
	H_3BO_3	0.16 mM	
	$NiCl_2 \cdot 6H_2O$	0.04 mM	
	$ZnSO_4 \cdot 7H_2O$	3.03 mM	
	$CuSO_4 \cdot 7H_2O$	5.43 mM	
	$MnCl_2 \cdot 4H_2O$	7.17 mM	
	$NaMoO_4$	0.10 mM	
	$CaCl_2$	9.73 mM	

Volumes of each stock used to perform the different cultures can be seen in next table:

Table 38. Volumes used for the different cultures

	50 mL	125 mL	300 mL	Units
Water	32.94	83.45	197.44	mL
MES 5x	8.48	21.20	51	mL
Glucose	6.24	15.63	37.50	mL
$MgSO_4$	0.50	1.25	3	mL
$FeCl_3$	0.20	0.50	1.20	mL
Thiamine	0.50	0.12	3	mL
TES	0.14	0.35	0.86	mL
Antibiotics	0.05	0.12	0.30	mL
Preinocculum	1	2.5	6	mL

13.2.3. Antibiotics

The antibiotics used during this project were 1000x concentrated as Table 37 shows.

Table 39. Antibiotic stocks

	Concentration
Ampicillin	100 mg/mL
Kanamycin	50 mg/mL
Chloramphenicol	34 mg/mL*
Streptomycin	100 mg/mL

All the antibiotics, when diluted were filtered through a filter of 0.22 µM and stored at -21ºC. Except for chloramphenicol, which was diluted using ethanol for analysis (*), all the antibiotics were diluted in MilliQ water.

13.2.4. IPTG

The stock prepared for the inducer was 1 M (1000x) meaning that 2.4 g were added to 10 mL of water and then filtered through a filter of 0.22 µM and stored at -21ºC. If 300 mL of culture was used the required amount of IPTG was 300 µL.

13.2.5. Preparation of cultures for further analysis

Once the cultures are grown the procedure to rinse it and store it was the following:

1. Centrifuge the cultures at 9000 rpm for 15 minutes
2. Discard the supernatant
3. Rinse the pellet with 0.9% NaCl (1/3 of the culture volume)
4. Centrifuge at 9000 rpm for 15 minutes
5. Discard supernatant
6. Resuspend the pellet with 0.9% NaCl (1/10 of total volume) and place it into a new tube (50 mL)
7. Centrifuge at 4500 rpm for 20 minutes
8. Discard supernatant and conserve the pellet

13.2.6. Lyophilization

After the culture had been induced to obtain cells a lyophilization procedure needs to be performed.

1. Centrifuge cultures at 9000 rpm for 15 minutes
2. Discard supernatant
3. Wash the pellet with 100 mL of 0.9% NaCl (if 300 mL)
4. Centrifuge 9000 rpm for 15 minutes
5. Discard supernatant
6. Wash the pellet with 25 mL of 0.9 NaCl and put it into a lyophilization tube
7. Centrifuge at 4500 rpm for 15 minutes
8. Discard supernatant
9. Store it at -80ºC for at least couple of hours
10. Put the samples in the lyophilization machine and let them lyophilize for, at least, overnight.
11. Once the cultures are lyophilized store them at -21ºC.

13.3. GGL analysis

13.3.1. Folch extraction

1. Weigh 40 mg of lyophilized biomass and put it in a lyophilization tube. Add 800 µL of chloroform and 640 µL of methanol. Put the sample into ultrasound bath for 30 minutes
2. Filter the sample with a 0.45 µm filter organic solvent resistant
3. Add 480 µL of chloroform and 480 µL of water into the sample filtered and put it in ultrasound bath for 15 minutes

4. Centrifuge the samples from the bath for 10 minutes at 13.000 rpm
5. With a Hamilton syringe take the lower band of the mixture that contains the organic solvent and place it into a glass assay tube (800 µL)
6. Dry the sample with N_2 flux

The volumes of each solvent were depending on the lyophilized mass used. For these assays, the amounts can be seen in tables 38 and 39.

Table 40. Volumes used to extract the lipid fraction in the first extraction

Sample (mg)	Chloroform (µL)	Methanol (µL)	Water (µL)	Total volume (µL)
10	200	160	-	360
15	300	240	-	540
20	400	320	-	720
40	800	640	-	1440

Table 41. Volumes used for the second extraction

Sample (mg)	Chloroform (µL)	Methanol (µL)	Water (µL)	Total volume (µL)
10	120	-	120	240
15	180	-	180	360
20	240	-	240	480
40	480	-	480	960

13.3.2. Anthrone reaction

13.3.2.1. Sample preparation

To prepare the samples for this assay, glass tubes obtained from Folch extraction (7.3.1. Folch extraction) were used as follows.

1. 30 µL of chloroform and 70 µL of ethanol were added to each tube in order to resuspend the organic fraction
2. 900 µL of MQ water was added to each tube so a 1 mL of final volume was obtained

13.3.2.2. Standard curve

Concentrations between 0 and 120 µg of glucose (duplicated) were prepared with a final volume of 1 mL from glucose Standard solution of 0.15 mg/mL. The standard curve of calibration was prepared regarding the following proportions to 1 mL as table 40 shows.

Table 42. Glucose standard curve

Glucose (µg)	Glucose standard solution (µL)	Water (µL)
0	0	1000
10	67	933
20	133	866
30	200	800
60	400	600
90	600	400
120	800	200
150	1000	0

13.3.2.3. Anthrone reagent preparation

This reagent was required to be prepared every time the assay was performed. The reagent was at 0.2% w/v anthrone in sulfuric acid (96%).

13.3.2.4. Anthrone assay

1. 2 mL of 85% phosphoric acid was added to all the samples to be analyzed (samples and calibration curve).
2. Tubes were covered and heated at 100ºC for 10 minutes
3. After cooling down the samples, 4 mL of anthrone reagent was added, well mixed and then heated again at 100ºC or 14 minutes
4. After cooling down, the samples are read at 625 nm

13.3.3. TLC analysis

13.3.3.1. One-dimension TLC

This type of TLC was used to confirm the presence of glycolipids in the assays were anthrone was performed.

13.3.3.1.1. Mobile phase preparation

The composition of this mobile phase consisted in 4 mL of deionized water, 25 mL of methanol and 65 mL of chloroform; added, preferably, by increasing value of volatility as mentioned. A piece of absorbent paper was added to the container before covering it in order to help to saturate the atmosphere. The container needed to remain unused during at least 15 minutes in order to let it condition before the TLC assay.

13.3.3.1.2. Sample preparation

A volume of 75 µL of the organic fraction obtained during Folch extraction procedure was poured into a HPLC glass vial and evaporated by N_2. The organic fraction was then resuspended with 35 µL of chloroform in order to concentrate the organic phase and 14 µL of the samples are added on the silica thin layer.

13.3.3.1.3. Standards

To check the correct performance of the TLC, standards of MGDAG, DGDAG, CL, PG and PE were used. MGDAG and DGDAG were obtained from Matreya LLC. (ref. 1058 and 1059) while cardiolipin, phosphatidylethanolamine and phosphatidylglycerol were obtained from Sigma Aldrich (ref. 21979-25MGF, P6386-5MG and 841138-25MG).

13.3.3.1.4. Thin Layer chromatography procedure

TLC was then introduced in the container, covered and left until the mobile phase reached the top part. After that time the TLC was removed and well dried before immersing it in a 0.5 % α-naphthol solution. The TLC was again dried with the help of a dryer and submerged into a sulfuric acid

solution (450 mL H_2O:450 mL MetOH:100 mL H_2SO_4). After that, TLC was warmed by introducing it inside the stove until the appearance of the purple spots.

13.3.3.2. Two-dimension TLC

To perform the assays related with radiolabeled ^{14}C-acetate, two-dimension TLC were performed. The dimensions of the silica used were 7.5x7.5 cm and two different mobile phases were used:
A mixture of ethyl acetate, isopropyl alcohol, chloroform, methanol and 0.25% KCl in water (25:25:25:11:9) was used as a first dimension.
The samples were added at the left side of the silica so it could be turned to the left side in order to perform the second elution.
Once the first dimension was run it was required to dry well the TLC before using the second mobile phase so the compounds could be correctly separated.
For the second mobile phase a mixture of chloroform, methanol and water (65:25:4) was used.
After running the second mobile phase, TLCs were dried and then, to confirm that they were OK, they were submerged into a solution of 0.5% of α-naphthol (methanol and water; 1:1) and after it were dried, they were put into to stove until the characteristic dots appeared.
On the other hand, if the TLC was used to quantify phospholipids, after running the second mobile phase, TLC were dried and stored until the radioactivity was measured.
The exposure time to read the TLC was of 16 hours.

13.4. Enzymatic activities

13.4.1. Glycosyltransferase activity

To determine the glycosyltransferase activity, it was required to use C6-NBD-ceramide, which is a fluorescent ceramide, so the formation of MGalCer and DGalCer could be detected by Agilent® HPLC 1200.

13.4.1.1. Preparation of cell extract for glycosyltransferase activity

Cultures of 50 mL defined medium were set up. When the $OD_{600\ nm}$ was between 0.15-0.2 the induction was performed. The cultures were stopped at an end point (24 hours) after the induction. They were centrifuged at 5000 g for 10 minutes then rinsed twice with 0.9% NaCl and finally resuspended with 1.5 mL of extraction buffer at pH 8 (20 mM $MgCl_2$, 20 mM HEPES, 10 mM CHAPS, 20% Glycerol and 0.5% NaCl). PMSF was added (1 mM) and then an overnight incubation at 4ºC and agitation was followed.
After the incubation time, 4 cycles of 2 min (20 s ON and 10 s OFF) of sonication were followed in order to lyse the cells. Cell debris was removed by centrifugation (13000 rpm, 5 minutes). Protein quantification was determined by BCA assay.

13.4.1.2. Glycosyltransferase determination

The reaction was performed in 10 mM HEPES, 5 mM CHAPS, 10% Glycerol, 0.25 mM $MgCl_2$, 25 µM C6-NBD-ceramide, 25 µM BSA and 1 mM UDP – Galactose and water. Cell lysate volume was

variable depending on the amount of protein obtained during the sonication (Andrés et al., 2011). Reaction was stopped by 80% of methanol at 2, 4, 6, 8, 10 and 12 minutes.

To analyze the reaction, 1200 Agilent® HPLC that detects fluorescence was used so it was possible to follow the products obtained during the reaction. The donor of the reaction was UDP – Galactose (Sigma Aldrich®) and the acceptor was the ceramide. HPLC conditions were 75:25 acetonitrile, water using a C18 Nova-pak® Waters. The samples were excited at 458 nm and read at 530 nm. Ceramide-C6-NBD eluted at ≈ 6.2 min, MonoGalCerNBD at ≈ 2.35 minutes and finally DiGalCerNBD at ≈ 1.74 minutes.

13.4.1.3. Quantification of the formed products

To determine the amount of product formed all the product was calculated by the following equation:

$$Product\ (\mu M) = \frac{(Area_{MGalCer} + Area_{DGalCer})}{(Area_{MGalCer} + Area_{DGalCer} + Area_{Cer})} x\ [Ceramide_0]$$

13.4.2. Acyltransferase activity

To calculate the enzymatic activity of the acyltransferases it was required to use a coupled assay where the formation of TNB by the interaction of the released CoA with DTNB. Once TNB was formed it was possible to detect it by microplate reader.

13.4.2.1. Cell extract preparation for the acyltransferase enzymatic activity assay

50 mL of cell culture were harvested by centrifugation (5000 g, 10 min, 4ºC), washed twice with 0.9% NaCl and suspended in 20 mM phosphate buffer (pH 6.5), 50 mM NaCl, 10 mM MgCl2 and 1 mM dithiothreitol (Mora-Buyé et al., 2012). Bacteria was disrupted by sonication at 0ºC for 4 cycles of 2 minutes (20 s ON, 10 s OFF), cell debris was removed by centrifugation (13000 rpm, 5 minutes). The total protein quantification was determined by BCA assay.

13.4.2.2. Acyltransferase reaction

PlsC and PlsB activities assays were set-up to analyze HsCoA formation with the Ellman's reagent (5,5'-dithio-bis-2-nitrobenzoic acid, DTNB). The reaction mixture contained 0.1 M Tris-HCl (pH 9), 0.5 mM $MgCl_2$, 1 mg mL^{-1} BSA, 0.1 mM DTNB, 50 mg L^{-1} oleoylCoA and freshly prepared cell extract. The reaction was started by adding 50 mg mL^{-1} of oleoyl lysophosphatidic acid for plsC acyltransferase, or 50 mg $^{L-1}$ glycerol – 3 – phosphate for acyltransferase plsB. The formation of 2-nitro-5-thiobenzoate (TNB) was monitored by absorbance (405 nm) every 30 seconds for 10 minutes at 30ºC. All the reagents were obtained at Sigma Aldrich®.

13.4.3. PgpB activity

Phosphatase activity of PgpB was determined by Malachite green assay kit. To perform this assay cultures of 50 mL of minimal medium were set up. Cells were induced with IPTG (1 mM) and grown overnight when the OD was around 0.15. Cultures were rinsed twice with 0.9% NaCl and they were resuspended using a buffer that contained maleic acid – Tris (25 mM) at pH 7, 20% of glycerol and 50 mM of $MgCl_2$. Samples were sonicated and centrifuged so the supernatant could be collected.

Reaction conditions contained 0.1 mM of LPA, 1 mM TRITON X-100, 2 mM $MgCl_2$ and 10 mM ß-mercaptoethanol. The amount of protein used for this assay was between 1.5 and 6 µg. Reaction was performed at 30ºC and samples were taken every 2 minutes. After following Sigma Aldrich® indications for the assay the microplates were read at 620 nm.
Standard curve was prepared as provider indicated.

13.4.4. CDH activity

To quantify the activity of CDH an assay based on Bulawa *et al* was performed (Bulawa and Raetzs, 1984). Cultures of 50 mL were grown in minimal medium for 24h (induction with IPTG when OD was around 0.2). After rinsing them, they were resuspended with Tris (25 mM) pH7 and TRITON X-100 (1 mM). After this, these samples were sonicated until the protein was freed.
Reaction buffer used contained a mixture of Tris (25 mM), TRITON X-100 (1 mM), $MgCl_2$ (1 mM) and 18:1 CDP-DG provided by Sigma Aldrich® (0.1 mM). Since CDH uses CDP-DAG to obtain PA releasing CMP, phosphatase alkaline (12 DEA units) were used to transform this CMP into cytidine so the released phosphate could be quantified by malachite green. Standard curve used was prepared as provider Sigma Aldrich indicated.
The reaction was performed at 37ºC and samples were taken every 2 minutes for 12 minutes.

13.5. Lipid analysis

13.5.1. Lipid profile by gas chromatography

To analyze the abundance of the lipids for each strain it was required to obtain the methyl ester that could, afterwards, be analyzed by gas chromatography. To do so, the following protocol was followed:

1. Grow and induce the cultures
2. Lyophilize the desired cultures
3. Weigh 20 mg of the lyophilized biomass
4. Extract the lipid fraction by Folch extraction, take 100 µL of the organic fraction, place it into a glass container and evaporate the organic phase by N_2 flow
5. Resuspend the lipid fraction with 2.5 mL of n-hexane
6. Add 25 µL of 2M KOH
7. Vortex the samples for 30 seconds
8. Take 100 µL of the upper phase and analyze it by GC

Table 43. Conditions for the Gas Chromatography analysis

Equipment	Agilent 7890A GC
Column	TRB-WaxOmega (Teknocroma®) (TR-840232) (30 m, 0.25 mm, 0.25 µm)
Temperature	200ºC
Pressure	20 psi
Injection volume	2 µL
Injector temperature	250ºC
Gas	Helium

Run time	20 min
Split ratio	25:1
Detector	FID
Detector temperature	250ºC
Standard used	SUPELCO 37 Component FAME mix

13.5.2. Lipid profile identification by GC-MS

To analyze the mass spectrum of each lipid found and confirm which lipid was, a GC-MS was used. The main differences in the protocols were the equipment used for the analysis. Conditions used for this assay can be found in the following table.

Table 44. Conditions for the mass analysis

Equipment	HRGC 6890N		
Column	HP-5MS UI (Agilent J&W GC Columns®) (Ref. 19091S-433UI) (30 m, 0.25 mm, 0.25 µm)		
Pressure	10.5 psi		
Temperature	T(ºC)	Time (min)	Ramp. (ºC/min)
	100	4	15
	300	20	-
Injection volume	2 µL		
Split ratio	10:1		
Detector	Agilent 5973 Inert Gas		
Detector temperature	200ºC MS Source: 230ªC MS quadrupole 150ºC		
Mass range	EM Scan 45 – 700 Da		
Threshold	30		
Solvent delay	3 min		
Run time	20 min		
Gas	Helium		
m/z library	NIST14		

13.5.3. Phospholipid analysis by radiolabeled ^{14}C-acetate

To analyze by radioactivity the abundance of the phospholipids found in the new strains it was required to use ^{14}C-acetate sodium salt, which was supplemented to the minimal medium so it could be used as a carbon source for the production of fatty acids. For each culture 10 µCi were added and the final concentration in the cultures was 3.8 µM.

Table 45. Stocks used for the minimal medium supplemented with ^{14}C-acetate

Medium	Concentration	Sterilization procedure
Glucose	20% (w/v)	Autoclave (120ºC for 20 min)
$MgSO_4$	0.2 M	Filtration 0.22 µm
$FeCl_3$	30 mM	Filtration 0.22 µm
Thiamine	1 mg/mL	Filtration 0.22 µm

Medium		Concentration	Sterilization procedure
MES 5x	$Na_2HPO_4 \cdot 2H_2O$	0.42 M	Autoclave (120ºC for 20 min)
	NaH_2PO_4	0.23 M	
	NaCl	0.21 M	
	$(NH_4)_2SO_4$	0.18 M	
TES	$AlCl_3 \cdot 6H_2O$	0.17 mM	Filtration 0.22 µm (pH 2)
	$CoCl_3 \cdot 6H_2O$	0.59 mM	
	H_3BO_3	0.16 mM	
	$NiCl_2 \cdot 6H_2O$	0.04 mM	
	$ZnSO_4 \cdot 7H_2O$	3.03 mM	
	$CuSO_4 \cdot 7H_2O$	5.43 mM	
	$MnCl_2 \cdot 4H_2O$	7.17 mM	
	$NaMoO_4$	0.10 mM	
	$CaCl_2$	9.73 mM	
^{14}C-acetate sodium salt		3.8 µM (final)	-

The protocol used for this assay was the following one:

1. 50 mL of cultures were set up supplemented each one with 10 µCi of ^{14}C-acetate
2. All the cultures were induced with IPTG 1 M when the OD was between 0.15 and 0.2
3. Cultures were grown overnight at 37ºC and 200 rpm
4. Centrifuge at 9000 rpm for 15 minutes and the supernatant was discarded
5. Rinse with 25 mL of 0.9% NaCl
6. Centrifuge at 9000 rpm for 15 min and the supernatant was discarded
7. Rinse with 5 mL of 0.9% NaCl
8. Centrifuge at 9000 rpm for 15 min. Supernatant was discarded
9. Cultures were resuspended with 1 mL of 0.9% NaCl and placed into two Eppendorfs (500 µL each fraction)
10. Centrifuged at maximum speed for 1 minute and the supernatant was discarded
11. Samples were resuspended with 0.67 mL of chloroform and 0.33 mL of methanol and were placed into an ultrasound bath for 30 min
12. Samples were centrifuged at maximal speed for 10 min and different fractions of 75 µL were taken
13. The obtained fractions were dried and then concentrated with the desired amount of chloroform for the 1D and 2D TLC.

13.6. Protein quantification

13.6.1. MG517 quantification by fluorescence

To quantify the amount of MG517 glycosyltransferase protein present in the different engineered strains a reports gene was used. In this case, the reported gene was mCherry fluorescent protein. The protocol to quantify this protein was the following one:

1. Grow 50 mL of minimal medium
2. Induce with IPTG when OD is around 0.15-0.2

3. Let the cultures grow for 24h at 37ºC and 200 rpm
4. Rinse the cultures twice with 0.9% NaCl
5. Resuspend this 50 mL of cultures with 1.5 mL of extraction buffer (20 mM $MgCl_2$, 20 mM HEPES, 10 mM CHAPS, 20% Glycerol and 0.5% NaCl) at pH 8
6. Take 1 mL of this sample and place it into fluorescent cuvettes
7. Read the fluorescence using a fluorimeter (excitation 550 nm; emission 607 nm)

13.6.2. Total protein quantification

13.6.2.1. BCA

This assay was used for all those samples that contained detergent such as CHAPS, glycerol, etc. This kit was provided by Thermo Scientific and the protocol followed was the one indicated by the provider for microplate reader.

The standard curve used for these assays can be found in the following table. The initial stock was of 2mg/mL of BSA provided by Thermo Scientific.

Table 46. BCA standard curve

Name	Concentration (mg/mL)	Volume from stock (µL)	Volume of solvent (µL)
A	2	300 (stock 2mg/mL)	0
B	1.5	375 (stock 2mg/mL)	125
C	1	325 (stock 2mg/mL)	325
D	0.75	175 (from B)	175
E	0.50	325 (from C)	325
F	0.25	325 (from E)	325
G	0.125	325 (from F)	325
H	0.025	100 (from G)	400
I	0	0	300

From this standard curve and, for all the tested samples, 25 µL of each one were mixed with 200 µL of the working reagent. Microplates were read at 595 nm after a 30 min incubation at 37ºC.

13.6.2.2. Bradford

For Bradford reaction, the same standard curve was used but only adding 5 µL of each standard and sample and mixed with 250 µL of Bradford reagent. Once it was mixed, 5 min incubation at room temperature was performed followed by a reading at 595 nm using a microplate reader.

REFERENCES

14. References

Abdel-Mawgoud, A.M., Stephanopoulos, G., 2018. Simple glycolipids of microbes: Chemistry, biological activity and metabolic engineering. Synth. Syst. Biotechnol. 3, 3–19. https://doi.org/10.1016/j.synbio.2017.12.001

Akasaka, H., Mizushina, Y., Yoshida, K., Ejima, Y., Mukumoto, N., Wang, T., Inubushi, S., Nakayama, M., Wakahara, Y., Sasaki, R., 2016. MGDG extracted from spinach enhances the cytotoxicity of radiation in pancreatic cancer cells. Radiat. Oncol. 11, 153. https://doi.org/10.1186/s13014-016-0729-0

Akasaka, H., Sasaki, R., Yoshida, K., Takayama, I., Yamaguchi, T., Yoshida, H., Mizushina, Y., 2013. Monogalactosyl diacylglycerol, a replicative DNA polymerase inhibitor, from spinach enhances the anti-cell proliferation effect of gemcitabine in human pancreatic cancer cells. Biochim. Biophys. Acta 1830, 2517–2525. https://doi.org/10.1016/j.bbagen.2012.11.004

Aliye, N., Fabbretti, A., Lupidi, G., Tsekoa, T., Spurio, R., 2015. Engineering color variants of green fluorescent protein (GFP) for thermostability, pH-sensitivity, and improved folding kinetics. Appl. Microbiol. Biotechnol. 99, 1205–1216. https://doi.org/10.1007/s00253-014-5975-1

Alsina, C., Faijes, M., Planas, A., 2019. Glycosynthase-type GH18 mutant chitinases at the assisting catalytic residue for polymerization of chitooligosaccharides. Carbohydr. Res. 478, 1–9. https://doi.org/10.1016/j.carres.2019.04.001

Anastas, P.T., Kirchhoff, M.M., 2002. Origins, current status, and future challenges of green chemistry. Acc. Chem. Res. 35, 686–94.

Andrés, E., Biarnés, X., Faijes, M., Planas, A., 2012. Bacterial glycoglycerolipid synthases: processive and non-processive glycosyltransferases in mycoplasma. Biocatal. Biotransformation 30, 274–287. https://doi.org/10.3109/10242422.2012.674733

Andrés, E., Martínez, N., Planas, A., 2011. Expression and characterization of a *Mycoplasma genitalium* glycosyltransferase in membrane glycolipid biosynthesis: Potential target against Mycoplasma infections. J. Biol. Chem. 286, 35367–35379. https://doi.org/10.1074/jbc.M110.214148

Aragunde, H., Castilla, E., Biarnés, X., Faijes, M., Planas, A., 2014. A transitional hydrolase to glycosynthase mutant by Glu to Asp substitution at the catalytic nucleophile in a retaining glycosidase. Carbohydr. Res. 389, 85–92. https://doi.org/10.1016/j.carres.2014.02.003

Bahia, F.M., De Almeida, G.C., De Andrade, L.P., Campos, C.G., Queiroz, L.R., Da Silva, R.L.V., Abdelnur, P.V., Corrêa, J.R., Bettiga, M., Parachin, N.S., 2018. Rhamnolipids production from sucrose by engineered *Saccharomyces cerevisiae*. Sci. Rep. 8, 1–10. https://doi.org/10.1038/s41598-018-21230-2

Bailey, R.W., 1958. The reaction of pentoses with anthrone. Biochem. J. 68, 669–672.

Banat, I.M., Makkar, R.S., Cameotra, S.S., 2000. Potential commercial applications of microbial surfactants. Appl. Microbiol. Biotechnol. 53, 495–508.

Barbier, M., Damron, F.H., 2016. Rainbow Vectors for Broad-Range Bacterial Fluorescence Labeling. PLoS One 11, e0146827. https://doi.org/10.1371/journal.pone.0146827

Barnes, E.M., Wakil, S.J., 1968. Studies on the mechanism of fatty acid synthesis. J. Biol. Chem. 243, 2955 . 2962.

Bauwens, D., 2019. Metabolic engineering strategies for improving microbial synthesis of chitooligosaccharides. Ghent University.

Behren, S., Westerlind, U., 2019. Glycopeptides and -Mimetics to Detect, Monitor and Inhibit Bacterial and Viral Infections: Recent Advances and Perspectives. Molecules 24, 1004. https://doi.org/10.3390/molecules24061004

Beld, J., Lee, D.J., Burkart, M.D., 2015. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. Mol. Biosyst. 11, 38–59. https://doi.org/10.1039/C4MB00443D

Bentley, G.J., Jiang, W., Guamán, L.P., Xiao, Y., Zhang, F., 2016. Engineering *Escherichia coli* to produce branched-chain fatty acids in high percentages. Metab. Eng. 38, 148–158. https://doi.org/10.1016/j.ymben.2016.07.003

Bernstein, R.L., Robbins, P.W., 1965. Control aspects of Uridine 5'-diphosphate glucose and Thymidine 5'-diphosphate glucose synthesis by microbial enzymes. J. Biol. Chem. 240, 391–397.

Blanchard, J.E., Withers, S.G., 2001. Rapid screening of the aglycone specificity of glycosidases: Applications to enzymatic synthesis of oligosaccharides. Chem. Biol. 8, 627–633. https://doi.org/10.1016/S1074-5521(01)00038-2

Blunt, J.W., Carroll, A.R., Copp, B.R., Davis, R.A., Keyzers, R.A., Prinsep, M.R., 2018. Marine natural products. Nat. Prod. Rep. 35, 8–53. https://doi.org/10.1039/c7np00052a

Bonner, W.M., Bloch, K., 1972. Purification and Properties of Fatty Acyl Thioesterase I from *E. coli*. J. Biol. Chem. 247, 3124–3133.

Brandenburg, K., Wagner, F., Müller, M., Heine, H., Andrä, J., Koch, M.H.J., Zähringer, U., Seydel, U., 2003. Physicochemical characterization and biological activity of a glycoglycerolipid from *Mycoplasma fermentans*. Eur. J. Biochem. 270, 3271–3279. https://doi.org/10.1046/j.1432-1033.2003.03719.x

Bulawa, C.E., Raetzs, C.R.H., 1984. Isolation and Characterization of *Escherichia coli* Strains Defective in CDP-diglyceride Hydrolase. J. Biol. Chem. 259.

Campbell, J.W., Cronan, J.E., 2002. The enigmatic *Escherichia coli fadE* gene is *yafH*. J. Bacteriol. 184, 3759–3764. https://doi.org/10.1128/JB.184.13.3759-3764.2002

Cao, Y., Cheng, T., Zhao, G., Niu, W., Guo, J., Xian, M., Liu, H., 2016. Metabolic engineering of *Escherichia coli* for the production of hydroxy fatty acids from glucose. BMC Biotechnol. 16, 1–9. https://doi.org/10.1186/s12896-016-0257-x

Carreño, L.J., Kharkwal, S.S., Porcelli, S. a, 2014. Optimizing NKT cell ligands as vaccine adjuvants. Immunotherapy 6, 309–20. https://doi.org/10.2217/imt.13.175

Celińska, E., Borkowska, M., Białas, W., Kubiak, M., Korpys, P., Archacka, M., Ledesma-Amaro, R., Nicaud, J.-M., 2019. Genetic engineering of Ehrlich pathway modulates production of higher alcohols in engineered *Yarrowia lipolytica*. FEMS Yeast Res. 19. https://doi.org/10.1093/femsyr/foy122

Chalfie, M., 1995. Green fluorescent protein. Photochem. Photobiol. 62, 651–6.

Chen, G.S., Segel, I.H., 1968. Purification and properties of glycogen phosphorylase from *Escherichia coli*. Arch. Biochem. Biophys. 127, 175–186.

Chennamadhavuni, D., Saavedra-Avila, N.A., Carreno, L.J., Guberman-Pfeffer, M.J., Arora, P., Yongqing, T., Pryce, R., Koay, H.-F., Godfrey, D.I., Keshipeddy, S., Richardson, S.K., Sundararaj, S., Lo, J.H., Wen, X., Gascon, J.A., Yuan, W., Rossjohn, J., Le Nours, J., Porcelli, S.A., Howell, A.R., 2018. Dual Modifications of alpha-Galactosylceramide Synergize to Promote Activation of Human Invariant Natural Killer T Cells and Stimulate Anti-tumor Immunity. Cell Chem. Biol. 25, 925. https://doi.org/10.1016/j.chembiol.2018.06.008

Chester, M.A., 1998. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of glycolipids--recommendations 1997. Eur. J. Biochem. 257, 293–8.

Cho, H., Cronan, J.E., 1995. Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. J. Biol. Chem. https://doi.org/10.1074/jbc.270.9.4216

Chong, H., Li, Q., 2017. Microbial production of rhamnolipids: Opportunities, challenges and strategies. Microb. Cell Fact. 16, 1–12. https://doi.org/10.1186/s12934-017-0753-2

Clinisciences, n.d. High purity Plasmid Miniprep Kit [WWW Document]. URL https://www.clinisciences.com/es/file_fournisseur/dc83_nb-03-0001.pdf (accessed 5.9.19).

Codera, V., Gilbert, H.J., Faijes, M., Planas, A., 2015. Carbohydrate-binding module assisting glycosynthase-catalysed polymerizations. Biochem. J. 470, 15–22. https://doi.org/10.1042/BJ20150420

Coelho, M.A.Z., Amaral, P.F.F., Belo, I., 2010. *Yarrowia lipolytica* : an industrial workhorse, in: Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology. pp. 930–944.

Colombo, D., Tringali, C., Franchini, L., Cirillo, F., Venerando, B., 2011. Glycoglycerolipid analogues

inhibit PKC translocation to the plasma membrane and downstream signaling pathways in PMA-treated fibroblasts and human glioblastoma cells, U87MG. Eur. J. Med. Chem. 46, 1827–1834. https://doi.org/10.1016/j.ejmech.2011.02.043

Conrad, S.E., Wold, M., Campbell, J.L., 1979. Origin and direction of DNA replication of plasmid RSF1030. Proc. Natl. Acad. Sci. 76, 736–740. https://doi.org/10.1073/pnas.76.2.736

Corti, M., Cantù, L., Brocca, P., Del Favero, E., 2007. Self-assembly in glycolipids. Curr. Opin. Colloid Interface Sci. 12, 148–154. https://doi.org/10.1016/j.cocis.2007.05.002

Costantino, V., D'Esposito, M., Fattorusso, E., Mangoni, A., Basilico, N., Parapini, S., Taramelli, D., 2005. Damicoside from *Axinella damicornis*: the influence of a glycosylated galactose 4-OH group on the immunostimulatory activity of alpha-galactoglycosphingolipids. J. Med. Chem. 48, 7411–7417. https://doi.org/10.1021/jm050506y

Coussement, P., Bauwens, D., Maertens, J., De Mey, M., 2017. Direct Combinatorial Pathway Optimization. ACS Synth. Biol. 6, 224–232. https://doi.org/10.1021/acssynbio.6b00122

Couvreur, P., Vauthier, C., 2006. Nanotechnology: Intelligent design to treat complex disease. Pharm. Res. https://doi.org/10.1007/s11095-006-0284-8

Czerwiec, Q., Idrissitaghki, A., Imatoukene, N., Nonus, M., Thomasset, B., Nicaud, J.-M., Rossignol, T., 2019. Optimization of cyclopropane fatty acids production in *Yarrowia lipolytica*. Yeast 36, 143–151. https://doi.org/10.1002/yea.3379

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. 97, 6640–6645. https://doi.org/10.1073/pnas.120163297

Davis, M.S., Solbiati, J., Cronan, J.E., 2000. Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli* . J. Biol. Chem. 275, 28593–28598. https://doi.org/10.1074/jbc.M004756200

De Bruyn, F., De Paepe, B., Maertens, J., Beauprez, J., De Cocker, P., Mincke, S., Stevens, C., De Mey, M., 2015a. Development of an in vivo glucosylation platform by coupling production to growth: Production of phenolic glucosides by a glycosyltransferase of *Vitis vinifera*. Biotechnol. Bioeng. 112, 1594–1603. https://doi.org/10.1002/bit.25570

De Bruyn, F., Van Brempt, M., Maertens, J., Van Bellegem, W., Duchi, D., De Mey, M., 2015b. Metabolic engineering of *Escherichia coli* into a versatile glycosylation platform: Production of bio-active quercetin glycosides. Microb. Cell Fact. 14, 1–12. https://doi.org/10.1186/s12934-015-0326-1

De Wulf, P., Kwon, O., Lin, E.C.C., 1999. The CpxRA signal transduction system of *Escherichia coli*: Growth-related autoactivation and control of unanticipated target operons. J. Bacteriol. 181, 6772–6778.

DeChavigny, A., Heacock, P.N., Dowhan, W., 1991. Sequence and inactivation of the *pss* gene of *Escherichia coli*. Phosphatidylethanolamine may not be essential for cell viability. J. Biol. Chem. 266, 5323–5332.

Dellomonaco, C., Clomburg, J.M., Miller, E.N., Gonzalez, R., 2011. Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. Nature 476, 355–359. https://doi.org/10.1038/nature10333

Dewald, J.H., Cavdarli, S., Steenackers, A., Delannoy, C.P., Mortuaire, M., Spriet, C., Noel, M., Groux-Degroote, S., Delannoy, P., 2018. TNF differentially regulates ganglioside biosynthesis and expression in breast cancer cell lines. PLoS One 13, e0196369. https://doi.org/10.1371/journal.pone.0196369

Dewald, J.H., Colomb, F., Bobowski-Gerard, M., Groux-Degroote, S., Delannoy, P., 2016. Role of Cytokine-Induced Glycosylation Changes in Regulating Cell Interactions and Cell Signaling in Inflammatory Diseases and Cancer. Cells 5. https://doi.org/10.3390/cells5040043

Dillon, D.A., Wu, W.I., Riedel, B., Wissing, J.B., Dowhan, W., Carman, G.M., 1996. The *Escherichia coli pgpB* gene encodes for a diacylglycerol pyrophosphate phosphatase activity. J. Biol. Chem. 271, 30548–53. https://doi.org/10.1074/JBC.271.48.30548

Dowhan, W., 1997. Molecular basis for membrane phospholipid diversity: why are there so many lipids? Annu. Rev. Biochem. 66, 199–232. https://doi.org/10.1146/annurev.biochem.66.1.199

Du, W., Gervay-Hague, J., 2005. Efficient Synthesis of α-Galactosyl Ceramide Analogues Using Glycosyl Iodide Donors. Org. Lett. 7, 2063–2065. https://doi.org/10.1021/ol050659f

Du, W., Kulkarni, S.S., Gervay-Hague, J., 2007. Efficient, one-pot syntheses of biologically active α-linked glycolipids. Chem. Commun. 2336. https://doi.org/10.1039/b702551c

Durand, G., Seta, N., 2000. Protein glycosylation and diseases: Blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring. Clin. Chem. 46, 795–805. https://doi.org/10.1093/clinchem/46.6.795

Edwards, M.D., Black, S., Rasmussen, T., Rasmussen, A., Stokes, N.R., Stephen, T.-L., Miller, S., Booth, I.R., 2012. Characterization of three novel mechanosensitive channel activities in *Escherichia coli*. Channels (Austin). 6, 272–281. https://doi.org/10.4161/chan.20998

Engler, C., Kandzia, R., Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3. https://doi.org/10.1371/journal.pone.0003647

Faijes, M., Planas, A., 2007. In vitro synthesis of artificial polysaccharides by glycosidases and glycosynthases. Carbohydr. Res. 342, 1581–1594. https://doi.org/10.1016/j.carres.2007.06.015

Faijes, M., Saura-Valls, M., Pérez, X., Conti, M., Planas, A., 2006. Acceptor-dependent regioselectivity of glycosynthase reactions by *Streptomyces E383A* beta-glucosidase. Carbohydr. Res. 341, 2055–2065. https://doi.org/10.1016/j.carres.2006.04.049

Faivre, V., Rosilio, V., 2010. Interest of glycolipids in drug delivery: from physicochemical properties to drug targeting. Expert Opin. Drug Deliv. 7, 1031–1048. https://doi.org/10.1517/17425247.2010.511172

Fan, J., Jiang, D., Zhao, Y., Liu, J., Zhang, X.C., 2014. Crystal structure of lipid phosphatase *Escherichia coli* phosphatidylglycerophosphate phosphatase B. Proc. Natl. Acad. Sci. U. S. A. 111, 7636–7640. https://doi.org/10.1073/pnas.1403097111

Folch, J., Lees, M., Sloane Stanley, G.H., 1957. A simple method for the isolation and purification of total lipides from animal tissues. J. Biol. Chem. 226, 497–509. https://doi.org/10.1371/journal.pone.0020510

Freeze, H.H., Schachter, H., Kinoshita, T., 2015. Genetic Disorders of Glycosylation, in: Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., Schnaar, R.L., Seeberger, P.H. (Eds.), . Cold Spring Harbor (NY), pp. 569–582. https://doi.org/10.1101/glycobiology.3e.045

Fujita, Y., Matsuoka, H., Hirooka, K., 2007. Regulation of fatty acid metabolism in bacteria. Mol. Microbiol. 66, 829–839. https://doi.org/10.1111/j.1365-2958.2007.05947.x

Funk, C.R., Zimniak, L., Dowhan, W., 1992. The *pgpA* and *pgpB* genes of *Escherichia coli* are not essential: evidence for a third phosphatidylglycerophosphate phosphatase. J. Bacteriol. 174, 205–213.

Furukawa, Koichi, Hamamura, K., Aixinjueluo, W., Furukawa, Keiko, 2006. Biosignals modulated by tumor-associated carbohydrate antigens: novel targets for cancer therapy. Ann. N. Y. Acad. Sci. 1086, 185–198. https://doi.org/10.1196/annals.1377.017

Gan, B.S., Kim, J., Reid, G., Cadieux, P., Howard, J.C., 2002. Lactobacillus fermentum RC-14 inhibits Staphylococcus aureus infection of surgical implants in rats. J. Infect. Dis. 185, 1369–1372. https://doi.org/10.1086/340126

Ghachi, M. El, Derbise, A., Bouhss, A., Mengin-Lecreulx, D., 2005. Identification of Multiple Genes Encoding Membrane Proteins with Undecaprenyl Pyrophosphate Phosphatase (UppP) Activity in *Escherichia coli*. J. Biol. Chem. 280, 18689–18695. https://doi.org/10.1074/jbc.M412277200

Ghosh, B., Lai, Y.-H., Shih, Y.-Y., Pradhan, T.K., Lin, C.-H., Mong, K.-K.T., 2013. Total synthesis of a glycoglycerolipid from *Meiothermus taiwanensis* through a one-pot glycosylation reaction and exploration of its immunological properties. Chem. Asian J. 8, 3191–3199.

https://doi.org/10.1002/asia.201300933
Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C. a, Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345. https://doi.org/10.1038/nmeth.1318
Glaser, L., Melo, A., Paul, R., 1967. Uridine diphosphate sugar hydrolase. Purification of enzyme and protein inhibitor. J. Biol. Chem. 242, 1944–54.
Green, P.R., Merrill, A.H., Bell, R.M., 1981. Membrane phospholipid synthesis in <i> Escherichia coli </>. Purification, reconstitution, and characterization of sn-glycerol-3-phosphate acyltransferase. J. Biol. Chem. 256, 11151–9.
Gregoriadis, G., Wills, E.J., Swain, C.P., Tavill, A.S., 1974. Drug-carrier potential of liposomes in cancer chemotherapy. Lancet (London, England) 1, 1313–6.
Grisewood, M.J., Hernández-Lozada, N.J., Thoden, J.B., Gifford, N.P., Mendez-Perez, D., Schoenberger, H.A., Allan, M.F., Floy, M.E., Lai, R.Y., Holden, H.M., Pfleger, B.F., Maranas, C.D., 2017. Computational Redesign of Acyl-ACP Thioesterase with Improved Selectivity toward Medium-Chain-Length Fatty Acids. ACS Catal. 7, 3837–3849. https://doi.org/10.1021/acscatal.7b00408
Groux-Degroote, S., Guerardel, Y., Delannoy, P., 2017. Gangliosides: Structures, Biosynthesis, Analysis, and Roles in Cancer. Chembiochem 18, 1146–1154. https://doi.org/10.1002/cbic.201600705
Groux-Degroote, S., Rodriguez-Walker, M., Dewald, J.H., Daniotti, J.L., Delannoy, P., 2018. Gangliosides in Cancer Cell Signaling. Prog. Mol. Biol. Transl. Sci. 156, 197–227. https://doi.org/10.1016/bs.pmbts.2017.10.003
Gumperz, J.E., Roy, C., Makowska, A., Lum, D., Sugita, M., Podrebarac, T., Koezuka, Y., Porcelli, S. a, Cardell, S., Brenner, M.B., Behar, S.M., 2000. Murine CD1d-restricted T cell recognition of cellular lipids. Immunity 12, 211–221. https://doi.org/10.1016/S1074-7613(00)80174-0
Hakomori, S., 1996. Tumor Malignancy Defined by Aberrant Glycosylation and Sphingo(glyco)lipid Metabolism ’. Cancer Res. 5309–5318.
Hakomori, S., Handa, K., 2002. Glycosphingolipid-dependent cross-talk between glycosynapses interfacing tumor cells with their host cells: Essential basis to define tumor malignancy. FEBS Lett. 531, 88–92. https://doi.org/10.1016/S0014-5793(02)03479-8
He, H., Lu, Y., Qi, J., Zhu, Q., Chen, Z., Wu, W., 2019. Adapting liposomes for oral drug delivery. Acta Pharm. Sin. B 9, 36–48. https://doi.org/10.1016/j.apsb.2018.06.005
He, L., Xiao, Y., Gebreselassie, N., Zhang, F., Antoniewicz, M.R., Tang, Y.J., Peng, L., 2014. Central metabolic responses to the overproduction of fatty acids in *Escherichia coli* based on 13C-metabolic flux analysis. Biotechnol. Bioeng. 111, 575–585. https://doi.org/10.1002/bit.25124
Heinz, E., Christie, W.W., 1996. Plant glycolipids: structure, isolation and analysis. In: Advances in Lipid Methodology, The Oily Press lipid library. Oily Press.
Hernández Lozada, N.J., Lai, R.Y., Simmons, T.R., Thomas, K.A., Chowdhury, R., Maranas, C.D., Pfleger, B.F., 2018. Highly Active C8-Acyl-ACP Thioesterase Variant Isolated by a Synthetic Selection Strategy. ACS Synth. Biol. 7, 2205–2215. https://doi.org/10.1021/acssynbio.8b00215
Holst, O., 2008. Glycolipids: Occurrence, Significance, and Properties, in: Fraser-Reid, B.O., Tatsuta, K., Thiem, J. (Eds.), Glycoscience: Chemistry and Chemical Biology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1603–1627. https://doi.org/10.1007/978-3-540-30429-6_39
Hölzl, G., Dörmann, P., 2007. Structure and function of glycoglycerolipids in plants and bacteria. Prog. Lipid Res. 46, 225–243. https://doi.org/10.1016/j.plipres.2007.05.001
Hossain, H., Wellensiek, H.J., Geyer, R., Lochnit, G., 2001. Structural analysis of glycolipids from *Borrelia burgdorferi*. Biochimie 83, 683–692. https://doi.org/10.1016/S0300-9084(01)01296-2
Hossain, S.A., Tanizawa, K., Kazuta, Y., Fukui, T., 1994. Overproduction and characterization of recombinant UDP-glucose pyrophosphorylase from *Escherichia coli* K-12. J. Biochem. 115,

965–72.

Huang, Y.-L., Hung, J.-T., Cheung, S.K.C., Lee, H.-Y., Chu, K.-C., Li, S.-T., Lin, Y.-C., Ren, C.-T., Cheng, T.-J.R., Hsu, T.-L., Yu, A.L., Wu, C.-Y., Wong, C.-H., 2013. Carbohydrate-based vaccines with a glycolipid adjuvant for breast cancer. Proc. Natl. Acad. Sci. U. S. A. 110, 2517–2522. https://doi.org/10.1073/pnas.1222649110

Janßen, H.J., Steinbüchel, A., 2014. Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. Biotechnol Biofuels 7, 7. https://doi.org/10.1186/1754-6834-7-7

Jia, Y., Akache, B., Deschatelets, L., Qian, H., Dudani, R., Harrison, B.A., Stark, F.C., Chandan, V., Jamshidi, M.P., Krishnan, L., McCluskie, M.J., 2019. A comparison of the immune responses induced by antigens in three different archaeosome-based vaccine formulations. Int. J. Pharm. 561, 187–196. https://doi.org/10.1016/j.ijpharm.2019.02.041

Joyce, S., Woods, a S., Yewdell, J.W., Bennink, J.R., De Silva, a D., Boesteanu, A., Balk, S.P., Cotter, R.J., Brutkiewicz, R.R., 1998. Natural ligand of mouse CD1d1: cellular glycosylphosphatidylinositol. Science (80-.). 279, 1541–1544. https://doi.org/10.1126/science.279.5356.1541

Kalisch, B., Dormann, P., Holzl, G., 2016. DGDG and Glycolipids in Plants and Algae., Subcellular biochemistry. United States. https://doi.org/10.1007/978-3-319-25979-6_3

Karlsson, K., 1989. Animal glycosphingolipids as membrane attachment sites for bacteria. Annu. Rev. Biochem. 58, 309–350.

Karlsson, K., 1986. Animal glycolipids as attachment sites for microbes. Chem. Phys. Lipids 42, 153–172.

Karp, P.D., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., Krummenacker, M., Subhraveti, P., Weerasinghe, D., Gama-Castro, S., Huerta, A.M., Muñiz-Rascado, L., Bonavides-Martinez, C., Weiss, V., Peralta-Gil, M., Santos-Zavaleta, A., Schröder, I., Mackie, A., Gunsalus, R., Collado-Vides, J., Keseler, I.M., Paulsen, I., 2014. The EcoCyc Database. EcoSal Plus 6. https://doi.org/10.1128/ecosalplus.ESP-0009-2013

Kavoosi, M., Creagh, A.L., Kilburn, D.G., Haynes, C.A., 2007. Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in *Escherichia coli*. Biotechnol. Bioeng. 98, 599–610. https://doi.org/10.1002/bit.21396

Kharkwal, S.S., Arora, P., Porcelli, S.A., 2016. Glycolipid activators of invariant NKT cells as vaccine adjuvants. Immunogenetics 68, 597–610. https://doi.org/10.1007/s00251-016-0925-y

Kiem, P. Van, Minh, C. Van, Nhiem, N.X., Cuong, N.X., Tai, B.H., Quang, T.H., Anh, H.L.T., Yen, P.H., Ban, N.K., Kim, S.H., Xin, M., Cha, J.-Y., Lee, Y.-M., Kim, Y.H., 2012. Inhibitory effect on TNF-alpha-induced IL-8 secretion in HT-29 cell line by glyceroglycolipids from the leaves of *Ficus microcarpa*. Arch. Pharm. Res. 35, 2135–2142. https://doi.org/10.1007/s12272-012-1210-8

Killian, J.A., Koorengevel, M.C., Bouwstra, J.A., Gooris, G., Dowhan, W., de Kruijff, B., 1994. Effect of divalent cations on lipid organization of cardiolipin isolated from *Escherichia coli* strain AH930. Biochim. Biophys. Acta 1189, 225–232.

Kim, J., Yoo, H.-W., Kim, M., Kim, E.-J., Sung, C., Lee, P.-G., Park, B.G., Kim, B.-G., 2018. Rewiring FadR regulon for the selective production of omega-hydroxy palmitic acid from glucose in *Escherichia coli*. Metab. Eng. 47, 414–422. https://doi.org/10.1016/j.ymben.2018.04.021

Klement, M.L.R., Öjemyr, L., Tagscherer, K.E., Widmalm, G., Wieslander, Å., 2007. A processive lipid glycosyltransferase in the small human pathogen *Mycoplasma pneumoniae*: Involvement in host immune response. Mol. Microbiol. 65, 1444–1457. https://doi.org/10.1111/j.1365-2958.2007.05865.x

Krengel, U., Bousquet, P.A., 2014. Molecular recognition of gangliosides and their potential for cancer immunotherapies. Front. Immunol. 5, 325. https://doi.org/10.3389/fimmu.2014.00325

Lee, Y.-L., Chen, J.C., Shaw, J.-F., 1997. The Thioesterase I of *Escherichia coli* Has Arylesterase Activity and Shows Stereospecificity for Protease Substrates. Biochem. Biophys. Res.

Commun. 231, 452–456. https://doi.org/10.1006/bbrc.1997.5797
Lemaire, G., Tenu, J.-P., Petit, J.-F., Lederer, E., 1986. Natural and synthetic trehalose diesters as immunomodulators. Med. Res. Rev. 6, 243–274. https://doi.org/10.1002/med.2610060302
Lewin, T.M., Wang, P., Coleman, R.A., 1999. Analysis of amino acid motifs diagnostic for the sn-glycerol-3-phosphate acyltransferase reaction. Biochemistry 38, 5764–5771. https://doi.org/10.1021/bi982805d
Leyva, A., Quintana, A., Sánchez, M., Rodríguez, E.N., Cremata, J., Sánchez, J.C., 2008. Rapid and sensitive anthrone-sulfuric acid assay in microplate format to quantify carbohydrate in biopharmaceutical products: Method development and validation. Biologicals 36, 134–141. https://doi.org/10.1016/j.biologicals.2007.09.001
Li, I.A., Popov, A.M., Sanina, N.M., Kostetskii, E.Y., Novikova, O.D., Reunov, A. V., Nagorskaya, V.P., Portnyagina, O.Y., Khomenko, V.A., Shnyrov, V.L., 2004. Physicochemical and Immune Properties of Glycoglycerolipids from *Laminaria japonica* in Immunostimulating Complexes (ISCOMs). Biol. Bull. 31, 244–249. https://doi.org/10.1023/B:BIBU.0000030144.98256.ee
Li, Q.X., Dowhan, W., 1990. Studies on the mechanism of formation of the pyruvate prosthetic group of phosphatidylserine decarboxylase from *Escherichia coli*. J. Biol. Chem. 265, 4111–4115.
Lightners, V.A., Larsons, T.J., Tailleuri, P., Kantorg, G.D., Raetzg, C.R.H., Bell, R.M., Modrich, P., 1980. Membrane Phospholipid Synthesis in *Escherichia coli* 255, 9413–9420.
Linington, R.G., Robertson, M., Gauthier, A., Finlay, B.B., MacMillan, J.B., Molinski, T.F., van Soest, R., Andersen, R.J., 2006. Caminosides B−D, Antimicrobial Glycolipids Isolated from the Marine Sponge *Caminus s phaeroconia*. J. Nat. Prod. 69, 173–177. https://doi.org/10.1021/np050192h
Liu, T., Khosla, C., 2010. Genetic Engineering of *Escherichia coli* for Biofuel Production. Annu. Rev. Genet. 44, 53–69. https://doi.org/10.1146/annurev-genet-102209-163440
Liu, Z., Guo, J., 2017. NKT-cell glycolipid agonist as adjuvant in synthetic vaccine. Carbohydr. Res. 452, 78–90. https://doi.org/10.1016/j.carres.2017.10.006
Lourith, N., Kanlayavattanakul, M., 2009. Natural surfactants used in cosmetics: glycolipids. Int. J. Cosmet. Sci. 31, 255–261. https://doi.org/10.1111/j.1468-2494.2009.00493.x
Lu, M., Kleckner, N., 1994. Molecular cloning and characterization of the pgm gene encoding phosphoglucomutase of *Escherichia coli*. J. Bacteriol. 176, 5847–51.
Lu, X., Vora, H., Khosla, C., 2008. Overproduction of free fatty acids in *E. coli* : Implications for biodiesel production. Metab. Eng. 10, 333–339. https://doi.org/10.1016/j.ymben.2008.08.006
Maeda, N., Kokai, Y., Hada, T., Yoshida, H., Mizushina, Y., 2013. Oral administration of monogalactosyl diacylglycerol from spinach inhibits colon tumor growth in mice. Exp. Ther. Med. 5, 17–22. https://doi.org/10.3892/etm.2012.792
Maeda, N., Matsubara, K., Yoshida, H., Mizushina, Y., 2011. Anti-cancer effect of spinach glycoglycerolipids as angiogenesis inhibitors based on the selective inhibition of DNA polymerase activity. Mini Rev. Med. Chem. 11, 32–38.
Magnuson, K., Jackowski, S., Rock, C.O., Cronan, Jr., J.E., 1993. Regulation of Fatty Acid Synthesis in *Escherichia coli* . Microbiol. Rev. 57, 522–542.
Mao, Z., Shin, H.-D., Chen, R.R., 2006. Engineering the E. coli UDP-glucose synthesis pathway for oligosaccharide synthesis. Biotechnol. Prog. 22, 369–74. https://doi.org/10.1021/bp0503181
Marcella, A.M., Barb, A.W., 2017. The R117A variant of the *Escherichia coli* transacylase FabD synthesizes novel acyl-(acyl carrier proteins). Appl. Microbiol. Biotechnol. 101, 8431–8441. https://doi.org/10.1007/s00253-017-8586-9
Marchant, R., Banat, I.M., 2012. Biosurfactants: a sustainable replacement for chemical surfactants? Biotechnol. Lett. 34, 1597–1605. https://doi.org/10.1007/s10529-012-0956-x
Marella, E.R., Holkenbrink, C., Siewers, V., Borodina, I., 2018. Engineering microbial fatty acid metabolism for biofuels and biochemicals. Curr. Opin. Biotechnol. 50, 39–46.

https://doi.org/10.1016/j.copbio.2017.10.002
Marrakchi, H., Zhang, Y.-M., Rock, C.O., 2002. Mechanistic diversity and regulation of Type II fatty acid synthesis. Biochem. Soc. Trans. 30, 30–33. https://doi.org/10.1042/BST0301050
Matsumoto, K., Okada, M., Horikoshi, Y., 1998. Cloning , Sequencing , and Disruption of the *Bacillus subtilis psd* Gene Coding for Phosphatidylserine Decarboxylase 180, 100–106.
Meng, X., Shang, H., Zheng, Y., Zhang, Z., 2013. Free fatty acid secretion by an engineered strain of *Escherichia coli*. Biotechnol. Lett. 35, 2099–2103. https://doi.org/10.1007/s10529-013-1305-4
Meyer, H.-P., Schmidhalter, R.D., 2012. Microbial Expression Systems and Manufacturing from a Market and Economic Perspective, in: Innovations in Biotechnology. pp. 211–250. https://doi.org/10.5772/29417
Mnif, I., Ghribi, D., 2016. Glycolipid biosurfactants: main properties and potential applications in agriculture and food industry. J. Sci. Food Agric. 96, 4310–4320. https://doi.org/10.1002/jsfa.7759
Mora-Buyé, N., Faijes, M., Planas, A., 2012. An engineered *E.coli* strain for the production of glycoglycerolipids. Metab. Eng. 14, 551–559. https://doi.org/10.1016/j.ymben.2012.06.001
Moran-Zorzano, M.T., Viale, A.M., Munoz, F.J., Alonso-Casajus, N., Eydallin, G.G., Zugasti, B., Baroja-Fernandez, E., Pozueta-Romero, J., 2007. Escherichia coli AspP activity is enhanced by macromolecular crowding and by both glucose-1,6-bisphosphate and nucleotide-sugars. FEBS Lett. 581, 1035–1040. https://doi.org/10.1016/j.febslet.2007.02.004
Murakami, S., Shimamoto, T., Nagano, H., Tsuruno, M., Okuhara, H., Hatanaka, H., Tojo, H., Kodama, Y., Funato, K., 2015. Producing human ceramide-NS by metabolic engineering using yeast *Saccharomyces cerevisiae*. Sci. Rep. 5, 16319. https://doi.org/10.1038/srep16319
Muthusamy, K., Gopalakrishnan, S., Ravi, T.K., Sivachidambaram, P., 2008. Biosurfactants: Properties, commercial production and application. Curr. Sci. 94, 736–747.
My, L., Ghandour Achkar, N., Viala, J.P., Bouveret, E., 2015. Reassessment of the Genetic Regulation of Fatty Acid Synthesis in *Escherichia coli* : Global Positive Control by the Dual Functional Regulator FadR. J. Bacteriol. 197, 1862–72. https://doi.org/10.1128/JB.00064-15
My, L., Rekoske, B., Lemke, J.J., Viala, J.P., Gourse, R.L., Bouveret, E., 2013. Transcription of the *Escherichia coli* fatty acid synthesis operon FabHDG is directly activated by FadR and inhibited by ppGpp. J. Bacteriol. 195, 3784–3795. https://doi.org/10.1128/JB.00384-13
Nair, S., Dhodapkar, M. V, 2017. Natural Killer T Cells in Cancer Immunotherapy. Front. Immunol. 8, 1178. https://doi.org/10.3389/fimmu.2017.01178
Neu, H.C., 1967. The 5'-nucleotidase of *Escherichia coli*. I. Purification and properties. J. Biol. Chem. 242, 3896–904.
Nitschke, M., Costa, S.G.V.A.O., 2007. Biosurfactants in food industry. Trends Food Sci. Technol. 18, 252–259. https://doi.org/10.1016/J.TIFS.2007.01.002
O'Brien, W.J., Frerman, F.E., 1977. Evidence for a complex of three beta-oxidation enzymes in *Escherichia coli*: induction and localization. J. Bacteriol. 132, 532–40.
Pan, H., Zhang, L., Li, X., Guo, D., 2017. Biosynthesis of the fatty acid isopropyl esters by engineered *Escherichia coli*. Enzyme Microb. Technol. 102, 49–52. https://doi.org/10.1016/j.enzmictec.2017.03.012
Pandey, R.P., Malla, S., Simkhada, D., Kim, B.-G., Sohng, J.K., 2013. Production of 3-O-xylosyl quercetin in *Escherichia coli*. Appl. Microbiol. Biotechnol. 97, 1889–1901. https://doi.org/10.1007/s00253-012-4438-9
Pandey, R.P., Parajuli, P., Koirala, N., Lee, J.H., Park, Y. Il, Sohng, J.K., 2014. Glucosylation of isoflavonoids in engineered *Escherichia coli*. Mol. Cells 37, 172–177. https://doi.org/10.14348/molcells.2014.2348
Papanikolaou, S., Aggelis, G., 2010. Yarrowia lipolytica: A model microorganism used for the production of tailor-made lipids. Eur. J. Lipid Sci. Technol. 112, 639–654. https://doi.org/10.1002/ejlt.200900197

Park, Y.-K., Dulermo, T., Ledesma-Amaro, R., Nicaud, J.-M., 2018. Optimization of odd chain fatty acid production by *Yarrowia lipolytica*. Biotechnol. Biofuels 11, 158. https://doi.org/10.1186/s13068-018-1154-4

Parsons, J.B., Rock, C.O., 2013. Bacterial lipids: metabolism and membrane homeostasis. Prog. Lipid Res. 52, 249–76. https://doi.org/10.1016/j.plipres.2013.02.002

Percival Zhang, Y.H., Himmel, M.E., Mielenz, J.R., 2006. Outlook for cellulase improvement: Screening and selection strategies. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2006.03.003

Pérez, X., Faijes, M., Planas, A., 2011. Artificial mixed-linked β-glucans produced by glycosynthase-catalyzed polymerization: tuning morphology and degree of polymerization. Biomacromolecules 12, 494–501. https://doi.org/10.1021/bm1013537

Pozzo, T., Romero-García, J., Faijes, M., Planas, A., Nordberg Karlsson, E., 2017. Rational design of a thermostable glycoside hydrolase from family 3 introduces β-glycosynthase activity. Glycobiology 27, 165–175. https://doi.org/10.1093/glycob/cww081

Qiao, K., Imam Abidi, S.H., Liu, H., Zhang, H., Chakraborty, S., Watson, N., Kumaran Ajikumar, P., Stephanopoulos, G., 2015. Engineering lipid overproduction in the oleaginous yeast *Yarrowia lipolytica*. Metab. Eng. 29, 56–65. https://doi.org/10.1016/j.ymben.2015.02.005

Quan, J., Tian, J., 2014. Circular Polymerase Extension Cloning, in: Methods in Molecular Biology (Clifton, N.J.). pp. 103–117. https://doi.org/10.1007/978-1-62703-764-8_8

Rabu, C., McIntosh, R., Jurasova, Z., Durrant, L., 2012. Glycans as targets for therapeutic antitumor antibodies. Future Oncol. 8, 943–960. https://doi.org/10.2217/fon.12.88

Raivio, T.L., 2014. Everything old is new again: An update on current research on the Cpx envelope stress response. Biochim. Biophys. Acta - Mol. Cell Res. 1843, 1529–1541. https://doi.org/10.1016/j.bbamcr.2013.10.018

Rakhuba, D., Novik, G., Dey, E.S., 2009. Application of supercritical carbon dioxide (scCO2) for the extraction of glycolipids from *Lactobacillus plantarum* B-01. J. Supercrit. Fluids 49, 49–51.

Rand, R.P., Luzzati, V., 1968. X-ray diffraction study in water of lipids extracted from human erythrocytes: the position of cholesterol in the lipid lamellae. Biophys. J. 8, 125–37. https://doi.org/10.1016/S0006-3495(68)86479-3

Ray, T.K., Cronan, J.E., 1987. Acylation of glycerol 3-phosphate is the sole pathway of de novo phospholipid synthesis in *Escherichia coli*. J. Bacteriol. 169, 2896–8.

Razin, S., Hayflick, L., 2010. Highlights of mycoplasma research—An historical perspective. Biologicals 38, 183–190. https://doi.org/10.1016/j.biologicals.2009.11.008

Razin, S., Yogev, D., Naot, Y., 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. Mol. Biol. Rev. 62, 1094–156.

Reid, G., 2001. Probiotic agents to protect the urogenital tract against infection. Am. J. Clin. Nutr. 73.

Rezuchova, B., Miticka, H., Homerova, D., Roberts, M., Kormanec, J., 2003. New members of the *Escherichia coli* σE regulon identified by a two-plasmid system. FEMS Microbiol. Lett. 225, 1–7. https://doi.org/10.1016/S0378-1097(03)00480-4

Rigouin, C., Lajus, S., Ocando, C., Borsenberger, V., Nicaud, J.M., Marty, A., Averous, L., Bordes, F., 2019. Production and characterization of two medium-chain-length polydroxyalkanoates by engineered strains of *Yarrowia lipolytica*. Microb. Cell Fact. 18, 99. https://doi.org/10.1186/s12934-019-1140-y

Rocha-Martin, J., Harrington, C., Dobson, A.D.W., O'Gara, F., 2014. Emerging strategies and integrated systems microbiology technologies for biodiscovery of marine bioactive compounds. Mar. Drugs 12, 3516–3559. https://doi.org/10.3390/md12063516

Rock, C.O., Jackowski, S., 2002. Forty Years of Bacterial Fatty Acid Synthesis. Biochem. Biophys. Res. Commun. 292, 1155–1166. https://doi.org/10.1006/bbrc.2001.2022

Rodrigues, L., Banat, I.M., Teixeira, J., Oliveira, R., 2007. Strategies for the prevention of microbial biofilm formation on silicone rubber voice prostheses. J. Biomed. Mater. Res. Part B Appl.

Biomater. 81B, 358–370. https://doi.org/10.1002/jbm.b.30673

Sajna, K.V., Sukumaran, R.K., Jayamurthy, H., Reddy, K.K., Kanjilal, S., Prasad, R.B.N., Pandey, A., 2013. Studies on biosurfactants from Pseudozyma sp. NII 08165 and their potential application as laundry detergent additives. Biochem. Eng. J. 78, 85–92. https://doi.org/10.1016/J.BEJ.2012.12.014

Sakakura, C., Sweeney, E. a, Shirahama, T., Hagiwara, a, Yamaguchi, T., Takahashi, T., Hakomori, S., Igarashi, Y., 1998. Selectivity of sphingosine-induced apoptosis. Lack of activity of DL-erythyro-dihydrosphingosine. Biochem. Biophys. Res. Commun. 246, 827–830. https://doi.org/S0006291X98987198 [pii]

Salis, H.M., Mirsky, E.A., Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. 27, 946–950. https://doi.org/10.1038/nbt.1568

San, K.-Y., Li, M., 2017. Bacteria and method for synthetizing fatty acids. US9598696B2. https://doi.org/10.1145/634067.634234.

Santiveri, C., 2018. Metabolic engineering on *E. coli* fatty acid and glucose pathways to enhance the glycoglycerolipids production.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods. https://doi.org/10.1038/nmeth.2089

Schramm, L., 2010. Surfactants: Fundamentals and Applications in the Petroleum Industry, PALAIOS. https://doi.org/10.2307/3515635

Scott, A., Noga, M.J., de Graaf, P., Westerlaken, I., Yildirim, E., Danelon, C., 2016. Cell-Free Phospholipid Biosynthesis by Gene-Encoded Enzymes Reconstituted in Liposomes. PLoS One 11, e0163058. https://doi.org/10.1371/journal.pone.0163058

Shcherbo, D., Merzlyak, E.M., Chepurnykh, T. V., Fradkov, A.F., Ermakova, G. V., Solovieva, E.A., Lukyanov, K.A., Bogdanova, E.A., Zaraisky, A.G., Lukyanov, S., Chudakov, D.M., 2007. Bright far-red fluorescent protein for whole-body imaging. Nat. Methods 4, 741–746. https://doi.org/10.1038/nmeth1083

Shetlar, M.R., 1952. Use of Anthrone Reaction for Determination of Carbohydrates in the Presence of Serum Protein. Anal. Chem. 24, 1844–1846.

Shin, K.S., Kim, S., Lee, S.K., 2016. Improvement of free fatty acid production using a mutant acyl-CoA thioesterase I with high specific activity in *Escherichia coli*. Biotechnol. Biofuels 9. https://doi.org/10.1186/s13068-016-0622-y

Simons, R.W., Egan, P.A., Chute, H.T., Nunn, W.D., 1980. Regulation of fatty acid degradation in *Escherichia coli* : Isolation and Characterization of Strains Bearing Insertion and Temperature-Sensitive Mutations in gene *fadR*. J. Bacteriol. 142, 621–632.

Springer, T. a, 1990. Adhesion receptors of the immune system. Nature 346, 425–434. https://doi.org/10.1038/346425a0

Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S.B., Keasling, J.D., 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature 463, 559–562. https://doi.org/10.1038/nature08721

Stefanetti, G., Hu, Q.Y., Usera, A., Robinson, Z., Allan, M., Singh, A., Imase, H., Cobb, J., Zhai, H., Quinn, D., Lei, M., Saul, A., Adamo, R., MacLennan, C.A., Micoli, F., 2015. Sugar-Protein Connectivity Impacts on the Immunogenicity of Site-Selective *Salmonella* O-Antigen Glycoconjugate Vaccines. Angew. Chemie - Int. Ed. 54, 13198–13203. https://doi.org/10.1002/anie.201506112

Sueoka, E., Nishiwaki, S., Okabe, S., Iida, N., Suganuma, M., Yano, I., Aoki, K., Fujiki, H., 1995. Activation of Protein Kinase C by Mycobacterial Cord Factor, Trehalose 6-Monomycolate, Resulting in Tumor Necrosis Factor-α Release in Mouse Lung Tissues. Japanese J. Cancer Res. 86, 749–755. https://doi.org/10.1111/j.1349-7006.1995.tb02464.x

Tan, Z., Black, W., Yoon, J.M., Shanks, J. V., Jarboe, L.R., 2017. Improving *Escherichia coli* membrane integrity and fatty acid production by expression tuning of FadL and OmpF. Microb. Cell Fact. 16. https://doi.org/10.1186/s12934-017-0650-8

Turnquist, R.L., Hansen, R.G., 1973. 2 Uridine Diphosphoryl Glucose Pyrophosphorylase, in: The Enzymes. Academic Press, pp. 51–71. https://doi.org/10.1016/S1874-6047(08)60062-1

Uniprot, n.d. mCherry [WWW Document]. URL https://www.uniprot.org/uniprot/X5DSL3 (accessed 4.12.19).

Val-Cid, C., Biarnés, X., Faijes, M., Planas, A., 2015. Structural-Functional Analysis Reveals a Specific Domain Organization in Family GH20 Hexosaminidases. PLoS One 10, e0128075. https://doi.org/10.1371/journal.pone.0128075

Vill, V., Hashim, R., 2002. Carbohydrate liquid crystals: Structure - Property relationship of thermotropic and lyotropic glycolipids. Curr. Opin. Colloid Interface Sci. https://doi.org/10.1016/S1359-0294(02)00091-2

Voelker, D.R., 1997. Phosphatidylserine decarboxylase. Biochim. Biophys. Acta - Lipids Lipid Metab. 1348, 236–244. https://doi.org/10.1016/S0005-2760(97)00101-X

Wang, X., Zou, X., Yin, J., Qiu, L., Lu, M., Hu, J., 2015. Carbohydrate-based vaccine adjuvants – discovery and development. Expert Opin. Drug Discov. 10, 1133–1144. https://doi.org/10.1517/17460441.2015.1067198

Wennekes, T., Van Den Berg, R.J.B.H.N., Boot, R.G., Van Der Marel, G. a., Overkleeft, H.S., Aerts, J.M.F.G., 2009. Glycosphingolipids - Nature, function, and pharmacological modulation. Angew. Chemie - Int. Ed. 48, 8848–8869. https://doi.org/10.1002/anie.200902620

Weyler, C., Heinzle, E., 2015. Multistep synthesis of UDP-glucose using tailored, permeabilized cells of *E. coli*. Appl. Biochem. Biotechnol. 175, 3729–3736. https://doi.org/10.1007/s12010-015-1540-3

White, S.W., Zheng, J., Zhang, Y.-M., Rock, C.O., 2005. The structural biology of type II Fatty Acid Biosynthesis. Annu. Rev. Biochem. 74, 791–831. https://doi.org/10.1146/annurev.biochem.74.082803.133524

Wicke, C., Huners, M., Wray, V., Nimtz, M., Bilitewski, U., Lang, S., 2000. Production and structure elucidation of glycoglycerolipids from a marine sponge-associated microbacterium species. J. Nat. Prod. 63, 621–626. https://doi.org/10.1021/np990313b

Wikstrom, M., Kelly, A.A., Georgiev, A., Eriksson, H.M., Klement, M.R., Bogdanov, M., Dowhan, W., Wieslander, A., 2009. Lipid-engineered *Escherichia coli* membranes reveal critical lipid headgroup size for protein function. J. Biol. Chem. 284, 954–965. https://doi.org/10.1074/jbc.M804482200

Wikstrom, M., Xie, J., Bogdanov, M., Mileykovskaya, E., Heacock, P., Wieslander, A., Dowhan, W., 2004. Monoglucosyldiacylglycerol, a foreign lipid, can substitute for phosphatidylethanolamine in essential membrane-associated functions in *Escherichia coli*. J. Biol. Chem. 279, 10484–10493. https://doi.org/10.1074/jbc.M310183200

Woo, J.E., Seong, H.J., Lee, S.Y., Jang, Y.-S., 2019. Metabolic Engineering of *Escherichia coli* for the Production of Hyaluronic Acid From Glucose and Galactose. Front. Bioeng. Biotechnol. 7, 351. https://doi.org/10.3389/fbioe.2019.00351

Xie, J., Bogdanov, M., Heacock, P., Dowhan, W., 2006. Phosphatidylethanolamine and monoglucosyldiacylglycerol are interchangeable in supporting topogenesis and function of the polytopic membrane protein lactose permease. J. Biol. Chem. 281, 19172–19178. https://doi.org/10.1074/jbc.M602565200

Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A.G.W., Collins, C.H., Koffas, M. a G., 2013. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. Nat. Commun. 4, 1409. https://doi.org/10.1038/ncomms2425

Yao, J., Rock, C.O., 2013. Phosphatidic Acid Synthesis in Bacteria. Biochim. Biophys. Acta 1831, 495–502. https://doi.org/10.1016/j.bbalip.2012.08.018

Yemm, E.W., Willis, A.J., 1954. The estimation of carbohydrates in plant extracts by anthrone. Biochem. J. 57, 508–514.

Yu, H., Santra, A., Li, Y., McArthur, J.B., Ghosh, T., Yang, X., Wang, P.G., Chen, X., 2018. Streamlined chemoenzymatic total synthesis of prioritized ganglioside cancer antigens. Org. Biomol. Chem.

16, 4076–4080. https://doi.org/10.1039/c8ob01087k

Yunoki, K., Sato, M., Seki, K., Ohkubo, T., Tanaka, Y., Ohnishi, M., 2009. Simultaneous quantification of plant glyceroglycolipids including sulfoquinovosyldiacylglycerol by HPLC-ELSD with binary gradient elution. Lipids 44, 77–83. https://doi.org/10.1007/s11745-008-3248-4

Zha, W., Rubin-Pitel, S.B., Shao, Z., Zhao, H., 2009. Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. Metab. Eng. 11, 192–198. https://doi.org/10.1016/j.ymben.2009.01.005

Zhang, F., Carothers, J.M., Keasling, J.D., 2012a. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat. Biotechnol. 30, 354–359. https://doi.org/10.1038/nbt.2149

Zhang, F., Ouellet, M., Batth, T.S., Adams, P.D., Petzold, C.J., Keasling, J.D., Mukhopadhyay, A., 2012b. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. Metab. Eng. 14, 653–660. https://doi.org/10.1016/j.ymben.2012.08.009

Zhang, J., Li, C., Yu, G., Guan, H., 2014. Total synthesis and structure-activity relationship of glycoglycerolipids from marine organisms. Mar. Drugs 12, 3634–3659. https://doi.org/10.3390/md12063634

Zhang, T., de Waard, A.A., Wuhrer, M., Spaapen, R.M., 2019. The Role of Glycosphingolipids in Immune Cell Functions. Front. Immunol. 10, 1–22. https://doi.org/10.3389/fimmu.2019.00090

Zhang, X., Agrawal, A., San, K.-Y., 2012. Improving fatty acid production in *Escherichia coli* through the overexpression of malonyl coA-acyl carrier protein transacylase. Biotechnol. Prog. 28, 60–65. https://doi.org/10.1002/btpr.716

Zhang, Y.-M., Marrakchi, H., Rock, C.O., 2002. The FabR (YijC) Transcription Factor Regulates Unsaturated Fatty Acid Biosynthesis in *Escherichia coli*. J. Biol. Chem. 277, 15558–15565. https://doi.org/10.1074/jbc.M201399200

Zhang, Y.-M., Rock, C.O., 2008. Thematic review series: Glycerolipids. Acyltransferases in bacterial glycerophospholipid synthesis. J. Lipid Res. 49, 1867–74. https://doi.org/10.1194/jlr.R800005-JLR200

Zhou, D., Mattner, J., Cantu, C., Schrantz, N., Yin, N., Gao, Y., Sagiv, Y., Hudspeth, K., Wu, Y.-P., Yamashita, T., Teneberg, S., Wang, D., Proia, R.L., Levery, S.B., Savage, P.B., Teyton, L., Bendelac, A., 2004. Lysosomal glycosphingolipid recognition by NKT cells. Science (80-.). 306, 1786–1789. https://doi.org/10.1126/science.1103440

ANNEXES

15. Annexes

15.1. Strains used in this project

15.1.1. BL21 Star (DE3) (WT)

This strain used did not contain any further genetic modification on its genome than the one that already contained BL21 Star (DE3) strain.

Genotype:

F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3)

This strain was used as a control since it was the strain used in previous studies performed in our group.

15.1.2. BL21 Star (DE3) Δ*fadE*

This strain apart from the modifications of BL21 Star (DE3) strain it also had a deletion on the *fadE* gene which is related with the ß-oxidation pathways. The whole gene was removed conserving in its place a scratch of 35 NT coming from the FRT recombined sites.

Genotype:

F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $fadE^-$

When the sequencing of this removed region was compared to the WT strain it was seen that there was an insertion of adenine in the middle of the sequence. This region was not codifying for anything so it was not important (Figure 84).

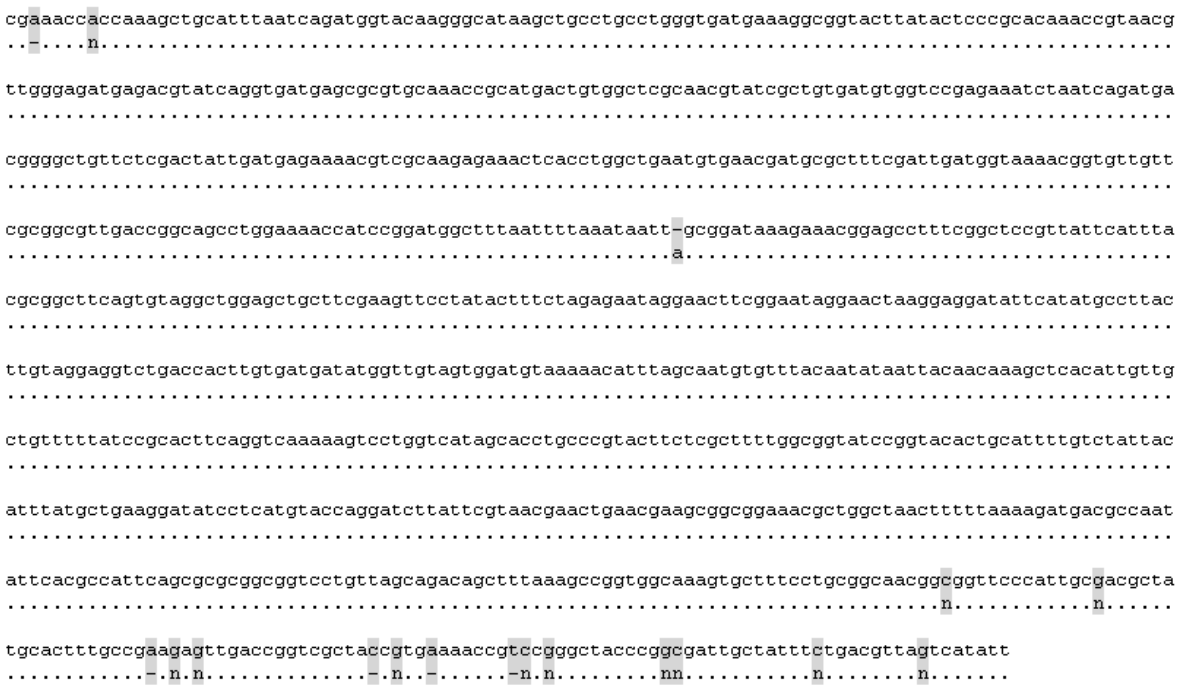

Figure 84. Sequencing of BL21 Star (DE3) ΔfadE

15.1.3. BL21 Star (DE3) Δ*tesA*

This strain apart from the modifications of BL21 Star (DE3) strain it also had a deletion on the *tesA* gene which is responsible for the hydrolysis of acyl groups to produce free fatty acids. It was not possible to remove the entire gene due there were CDS from other genes so only the middle part was removed making the protein inactive.

Genotype:

F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $tesA^-$

It seems that there was an insertion at 5' but the electropherogram confirmed that the sequence was OK.

```
 779 ggatctccgttgc-tttatgagtcatgatttactaaaggctgcaactgcttcgccatccagtcggcaataaacggctgggcgtcgcggtt
   1 ..............t.............................................................................

 868 gggatgaataccgtcatcctgcatccattgtggcttgaggtagacctcttccataaaaaagggcagcagcggaacatcaaactctttggc
  91 .............................................................................................

 958 gagtttggggtaaatggcgctaaaggcttcattataacggcgagtgtaggctggagctgcttcgaagttcctatactttctagagaatag
 181 .............................................................................................

1048 gaacttcggaataggaactaaggaggatattcatatgtaccgacgttttactctgccacttatcattcaacaaggcaggccaggccgcgc
 271 .............................................................................................

1138 tggcagacattcgatacccggcgctcaggctatcacccagaatcaataacgtgtccgctgcggcggcacggaaggttaacaggaccagga
 361 .............................................................................................

1228 acaggaagggcaaatgccagcggaaaacattgttgaagttcatcatcttaagaagtccgtcggtcaggggagcatgaact
 451 ...................................................................................
```

Figure 85. BL21 Star (DE3) ΔtesA sequencing

15.1.4. BL21 Star (DE3) Δ*tesA* Δ*fadE*

To obtain this KO, first it was used BL21 Star (DE3) Δ*tesA* as initial strain and then *fadE* was removed entirely.

Genotype:

F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $fadE^-$ $tesA^-$

```
 570 taaccattagcagccttgcaaagggcgaaaccacc-aaagctgcatttaatcagatggtacaagggcataagctgcctgcctgggtgatgaaaggcggtacttatactcc
 959 ....................................a..............................................................

 679 cgcacaaaccgtaacgttgggagatgagacgtatcaggtgatgagcgcgtgcaaaccgcatgactgtggctcgcaacgtatcgctgtgatgtggtccgagaaatctaatc
 849 .............................................................................................

 789 agatgacggggctgttctcgactattgatgagaaaacgtcgcaagagaaactcacctggctgaatgtgaacgatgcgctttcgattgatggtaaaacggtgttgttcgcg
 739 .............................................................................................

 899 gcgttgaccggcagcctggaaaaccatccggatggctttaattttaaataatt-gcggataaagaaacggagcctttcggctccgttattcatttacgcggcttcagtgt
 629 ..........................................................a..................................

1008 aggctggagctgcttcgaagttcctatactttctagagaataggaacttcggaataggaactaaggaggatattcatatgccttacttgtaggaggtctgaccacttgtg
 519 .............................................................................................

1118 atgatatggttgtagtggatgtaaaaacatttagcaatgtgtttacaatataattacaacaaagctcacattgttgctgtttttatccgcacttcaggtcaaaaagtcct
 409 ....................................a........................................................

1228 ggtcatagcacctgcccgtacttctcgcttttggcggtatccggtacactgcattttgtctattacatttatgctgaaggatatcctcatgtaccaggatcttattcgta
 299 .............................................................................................

1338 acgaactgaacgaagcggcggaaacgctggctaacttttttaaaagatgacgccaatattcacgccattcagcgcgcggcggtcctgttagcagacagctttaaagccggt
 189 .............................................................................................

1448 ggcaaagtgctttcctgcggcaacggcggttcccattgcgacgctatgcactttgccgaagagttgaccggtcgctacc
  79 ...............................................................................
```

Figure 86. FadE region sequencing of the double knockout BL21 Star (DE3) ΔtesA ΔfadE

There were some insertions but in a non-codifying region.

15.1.5. BL21 Star (DE3) Δ*tesA* Δ*fabR*

As it happened in the previous strain, Δ*tesA* was used as an initial strain and then *fabR*, genetic regulator that inhibits the biosynthesis of FA, was removed from the genome.

Genotype:

F⁻ dcm ompT hsdS(rB⁻ mB⁻) gal λ(DE3) tesA⁻ fabR⁻

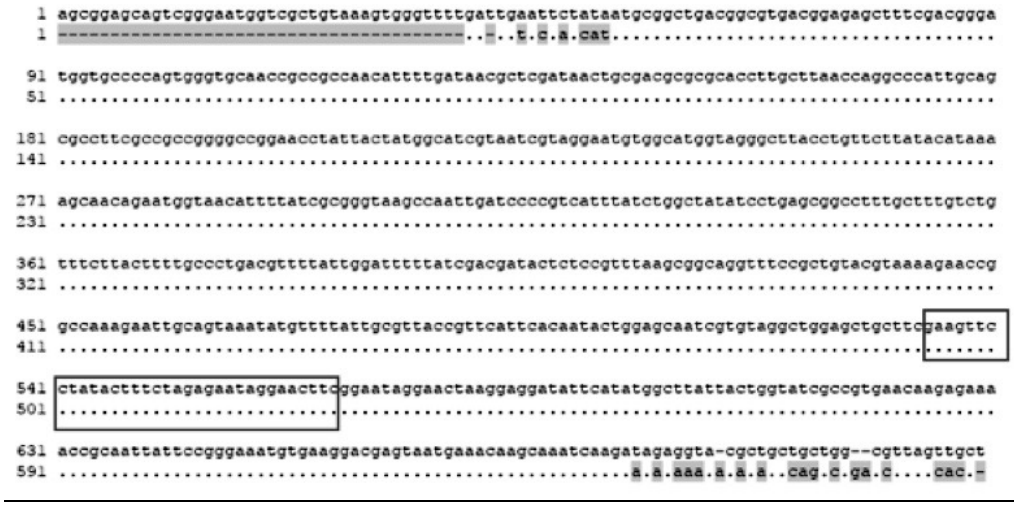

Figure 87. FabR sequence of BL21 Star (DE3) ΔtesA ΔfabR

There were no mismatches found in this knockout. The extreme regions of the sequencing did not have good quality and the electropherogram was not concluding.

15.1.6. BL21 Star (DE3) Δ*tesA* Δ*ushA*

Using BL21 Δ*tesA* strain as an initial strain, *ushA*, the enzyme responsible for the conversion of UDP-glucose to glucose-1-phosphate, was removed entirely from the genome of this strain.

Genotype:

F⁻ dcm ompT hsdS(rB⁻ mB⁻) gal λ(DE3) tesA⁻ ushA⁻

Figure 88. UshA removal sequence of BL21 Star (DE3) ΔtesA ΔushA

15.1.7. Summary

Table 47. Strain genotypes

Strain	Referred as	Gene removal	Genotype
BL21 Star (DE3)	WT	None	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3)
BL21 Star (DE3) Δ*fadE*	Δ*fadE*	*fadE*	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $fadE^-$
BL21 Star (DE3) Δ*tesA*	Δ*tesA*	*tesA*	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $tesA^-$
BL21 Star (DE3) Δ*tesA* Δ*fadE*	Δ*tesA* Δ*fadE*	*fadE* and *tesA*	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $fadE^-$ $tesA^-$
BL21 Star (DE3) Δ*tesA* Δ*fabR*	Δ*tesA* Δ*fabR*	*fabR* and *tesA*	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $tesA^-$ $fabR^-$
BL21 Star (DE3) Δ*tesA* Δ*fabR*	Δ*tesA* Δ*ushA*	*ushA* and *tesA*	F^- dcm ompT hsdS(rB^- mB^-) gal λ(DE3) $tesA^-$ $ushA^-$

15.2 Production summary of the strains

Table 48. Overall glycoglycerolipid production of the engineered strains

	Strain (genotype)	OD	$[Glc]_T$ (µg/mg)	%GGL composition				$[GGL]_T$ (nmol/mg)	Ratio GGL strain/wt(#1)	p-value strain vs wt (#1)
				%M	%D	%Tri	%Tetra			
#0	WT/ mg517	3.6 ± 0.1	0.84	36	31	15	18	2.17 ± 0.50	0.7	0.002*
#1	WT/ mg517-plsC^H	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0	-
#2	ΔtesA/ mg517-plsC^H	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7	0.023*
#3	ΔfadE/ mg517-plsC^H	3.1 ± 0.3	1.51	20	47	20	13	3.72 ± 0.70	1.1	0.118
#4	ΔtesA ΔfadE/ mg517-plsC^H	3.1 ± 0.7	0.92	31	39	8	22	2.31 ± 0.70	0.7	0.054
#5	ΔtesA ΔfabR/ mg517-plsC^H	2.1 ± 0.3	2.04	18	50	12	20	4.84 ± 0.70	1.5	0.024*
#6	ΔtesA/fadR-mg517-plsC^H	1.7 ± 0.2	2.40	12	60	12	16	5.75 ± 0.50	1.8	0.015*
#7	ΔtesA ΔfabR/ fadR-mg517-plsC^H	1.9 ± 0.1	2.29	14	64	17	6	5.86 ± 0.50	1.8	0.012*
#8	ΔtesA/ mg517	3.2 ± 0.8	0.92	32	42	17	6	2.65 ± 0.50	0.8	0.158
#9	ΔtesA/ mg517-plsC^L	2.7 ± 0.3	0.92	16	40	31	14	2.11 ± 0.50	0.6	0.029*
#10	ΔtesA/ mg517-plsB^L	2.0 ± 0.6	1.14	16	47	24	13	2.72 ± 0.70	0.8	0.206
#11	ΔtesA/ mg517-plsC^L·plsB^L	1.5 ± 0.3	2.30	11	50	29	11	5.35 ± 0.50	1.6	0.019*
#12	ΔtesA/ mg517-plsB^H	1.7 ± 0.2	2.06	17	54	17	12	5.11 ± 1.20	1.6	0.038*
#13	ΔtesA/ mg517-plsC^H·plsB^H	1.9 ± 0.3	2.72	4	45	40	10	5.95 ± 0.60	1.8	0.011*
#14	ΔtesA/ mg517·cdh-plsC^H·plsB^H	1.7 ± 0.1	2.20	4	86	6	5	5.71 ± 1.40	1.8	0.051
#15	ΔtesA/ mg517-plsCxpgpB^H	1.6 ± 0.2	4.83	4	45	29	23	9.93 ± 1.20	3.0	0.001*
#16	ΔtesA/ mg517-plsC^H-galU	1.7 ± 0.2	4.30	25	21	20	34	9.08 ± 0.30	2.8	0.002*
#17	ΔtesA/mg517-plsCxpgpB^H-galU	1.9 ± 0.3	2.45	10	50	18	21	5.44 ± 1.50	1.7	0.128
#18	ΔtesA ΔushA/ mg517-plsC^H	2.2 ± 0.6	3.41	6	71	13	9	8.63 ± 1.50	2.6	0.006*
#19	ΔtesA ΔushA/ mg517-plsC^H·plsB^H	1.7 ± 0.1	4.02	2	63	15	18	9.13 ± 0.70	2.8	0.002*
#20	ΔtesA ΔushA/ mg517-plsC^H-galU	1.8 ± 0.2	2.85	33	32	19	16	7.28 ± 0.60	2.2	0.004*
#21	ΔtesA ΔushA/ fadR-mg517-plsC^H·plsB^H	2.9 ± 0.1	2.28	26	34	16	25	5.29 ± 0.50	1.6	0.013*
#22	ΔtesA ΔushA/ mg517·cdh - plsC^H·plsB^H	1.8 ± 0.2	1.43	6	82	9	3	3.81 ± 0.30	1.2	0.195
#23	ΔtesA ΔushA/ mg517-plsCxpgpB^H	1.7 ± 0.1	3.09	6	63	17	13	7.22 ± 0.70	2.2	0.017*

P-value includes the statistical analysis using a T-student analysis with a 95% of confidence.

** Significant increase/decrease of the GGL production compared to #1 strain (reference strain).*

15.3. Plasmids used in this project

Name	Gene(s)	Promoter	Inducer	Origin of replication	Resistance	Use	Reference
mg517	*mg517*	T7	IPTG	ColE1	Ampicillin	Obtain GL	(Mora-Buyé et al., 2012)
mg517xmCherry	*mg517* and *mCherry*	T7	IPTG	ColE1	Ampicillin	Quantify MG517	This work
Mg517·cdh	*mg517* and *cdh*	T7	IPTG	ColE1	Ampicillin	Obtain GL by increasing PA	This work
pET44b-mg517 Y169F	*mg517* mutated Y169F	T7	IPTG	ColE1	Ampicillin	Obtain GGL	This work
plsC^{L}	*plsC*	T7	IPTG	p15a ori	Chloramphenicol	Increase production of PA	This work
plsC^{H}	*plsC*	T7	IPTG	RSF1030	Kanamycin	Increase PA	(Mora-Buyé et al., 2012)
plsB^{L}	*plsB*	T7	IPTG	p15a ori	Chloramphenicol	Increase LPA	This work
plsB^{H}	*plsB*	T7	IPTG	RSF1030	Kanamycin	Increase LPA	This work
plsC^{L}·plsB^{L}	*plsC* and *plsB*	T7	IPTG	p15a ori	Chloramphenicol	Increase PA and LPA	This work
plsC^{H}·plsB^{H}	*plsC* and *plsB*	T7	IPTG	RSF1030	Kanamycin	Increase PA and LPA	This work
plsCxpgpB^{H}	*plsC* and *pgpB*	T7	IPTG	RSF1030	Kanamycin	Increase PA and DAG	This work
fadR	*fadR*	T7	IPTG	pSC101	Streptomycin	Increase FA biosynthesis	This work
galU	*galU*	T7	IPTG	CloDF13	Streptomycin	Increase UDP-Glc	(Mora-Buyé et al., 2012)
galU^{L}	*galU*	T7	IPTG	pSC101	Streptomycin	Increase UDP-Glc	This work

Name	Gene(s)	Promoter	Inducer	Origin of replication	Resistance	Use	Reference
pKD46	*λ recombinase*	AraC	Arabinose	OriR101	Ampicillin	Recombination	(Datsenko and Wanner, 2000)
pKD3	*Chloramphenicol resistance*	R6K	-	R6K γ ori	Ampicillin and chloramphenicol	Resistance cassette flanked by FRT sites	(Datsenko and Wanner, 2000)
pKD4	*Kanamycin resistance*	R6K	-	R6K γ ori	Ampicillin and kanamycin	Resistance cassette flanked by FRT sites	(Datsenko and Wanner, 2000)
pCP20	*Flipase gene*	-	-	repA101	Ampicillin and chloramphenicol	Flipase used to remove FRT cassette	(Datsenko and Wanner, 2000)
KI plasmid	*psd*	p22	-	ColE1	Chloramphenicol	To perform the library	This work
Screening plasmid	*mKATE2*	Lambda promoter	-	pSC101	Kanamycin	To screen the library	This work

15.4. Primers used in this project

Primer name	Sequence 5' → 3'	Use
oMEMO4468	AGCGGATAAAGAAACGGAGCCTTTCGGCTCCGTTATTCATTTACGCGGCTTCAGTGTAGGCTGGAGCTGCTTC	FadE KO
oMEMO4480	TACATCCACTACAACCATATCATCACAAGTGGTCAGACCTCCTACAAGTAAGGCATATGAATATCCTCCTTAG	
oMEMO4508	ACCGACGCAGGAAATATGAC	Check FadE
oMEMO4509	CAGTGCAATGGCGCAGGATG	
oMEMO4464	CTTTGGCGAGTTTGGGGTAAATGGCGCTAAAGGCTTCATTATAACGGCGAGTGTAGGCTGGAGCTGCTTC	TesA KO
oMEMO4465	CGGCCTGGCCTGCCTTGTTGAATGATAAGTGGCAGAGTAAAACGTCGGTACATATGAATATCCTCCTTAG	
oMEMO4466	GCCAATTCCACTGGAACATC	Check TesA
oMEMO4467	CCGGTGAGGATGGAGAGTTC	
fabR_dats_fwC	AGAACCGGCCAAAGAATTGCAGTAAATATGTTTTATTGCGTTACCGTTCATTCACAATACTGGAGCAATCGTGTAGGCTGGAGCTGCTTC	FabR KO
comp C	ATACTCTCCGTTTAAGCGGCAGGTTTCCGCTGTACGTAAAAGAACCGGCCAAAGAATTGCAGTAAATATG	
fabR_dats_rvD	ATTACTCGTCCTTCACATTTCCCGGAATAATTGCGGTTTTCTCTTGTTCACGGCGATACCAGTAATAAGCCATATGAATATCCTCCTTAG	
comp_D	AGCAACTAACGCCAGCAGCAGCGTACCTCTATCTTGATTTGCTTGTTTCATTACTCGTCCTTCACATTTC	
oMEMO2506_SthA_ seq_rv	GGAATGGTCGCTGTAAAGTG	Check FabR
reverse_fabR_out	GCGTACCTCTATCTTGATTTG	
ushA_dats_fwA	ACAGAATTTCTAATCTGGATGCAGATTTATCTTCACCGGACGCAGACTTGTCTATGATGTCGCGTCATACGTGTAGGCTGGAGCTGCTTC	UshA KO
ushA_dats_rvB	AAATTTGCTGATATCGCCCGCCGCGATTAAGCATTGTGCCGGATGCAAACATCCGGCACTTTCGGATTACCATATGAATATCCTCCTTAG	
comp_A	ATTACCAGACTAACATACCTGTATGCGTCGTCTGAAGGAAGTCTCAACGCCGAATACAGAATTTCTAATC	
comp_B	TGGCGGCAGGCGATCTGGCAAAGATCCTCGATGCCAAATTTGCTGATATC	
oMEMO2890	TGTTCGGGCCTTGGTCGTTG	Check UshA
oMEMO2891	TTGAGATGGCCGATCCGATG	
oMEMO3168	CCGGTGCCCTGAATGAACTG	
ushA_dins_fw	GCTTTCCGCGAATATCTACC	Check Kanamycin
oMEMO4560	CAATTCCCCTGTAGAAATAATTTATGCTATATATCTTTCGTCTTATTATTACC	$plsC^{L}$
oMEMO4561	CTCAGCGGTGGCAGCAGCCTTTAAACTTTTCCGGCGGCTTC	
oMEMO4562	GAAGCCGCCGGAAAAGTTTAAAGGCTGCTGCCACCGCTGAG	
oMEMO4563	GGTAATAATAAGACGAAAGATATATAGCATAAATTATTTCTACAGGGGAATTG	

Primer name	Sequence 5' → 3'	Use
oMEMO3479	CCTCAAGACCCGTTTAGAGG	Check plsC^{L}
oMEMO1568	CATTCGATGGTGTCCGGGATCTC	
oMEMO4552	CAATTCCCCTGTAGAAATAATTTATGACTTTCTGCTATCCTTG	plsB^{L}
oMEMO4553	GTTATTGCTCAGCGGTGGCATTACCCTTCGCCCTGCGTCGCACTC	
oMEMO4554	GAGTGCGACGCAGGGCGAAGGGTAATGCCACCGCTGAGCAATAAC	
oMEMO4555	CAAGGATAGCAGAAAGTCATAAATTATTTCTACAGGGGAATTG	
oMEMO1560	CATTCGATGGTGTCCGGGATCTC	Check plsB^{L}
oMEMO4383	TGCCGTTTGTGATGGCTTCC	
oMEMO4554	GAGTGCGACGCAGGGCGAAGGGTAATGCCACCGCTGAGCAATAAC	plsC·plsB^{L}
oMEMO4684	CATATGTATATCTCCTTCTTGAATTCTTAAACTTTTCCGGCGGCTT	
oMEMO4553	GTTATTGCTCAGCGGTGGCATTACCCTTCGCCCTGCGTCGCACTC	
oMEMO4685	GAATTCAAGAAGGAGATATACATATGATGACTTTCTGCTATCCTTG	
fw_plsB_pRSFB	TATACCATGGCTTTCTGCTATCCTTG	plsB^{H}
rv_plsB_pRSFB	CTTAGAATTCTTACCCTTCGCCCTGC	
fw_pRSF_pRSFB	ATGCGAATTCCCTAGGCTGCTGCCACCGCTGAG	
rv_pRSF_pRSFB	CATACCATGGTATATCTCCTTATTAAAGTTAAAC	
fw_vector_Gibson CB	GACGAAAGATATATAGCATAGTTAAACAAAATTATTTCTACA	plsC·plsB^{H}
rv_vector_Gibson CB	GAGAGTGCGACGCAGGGCGAAGGGTAACTGCCACCGCTGAGCAATAAC	
fw_CB_Gibson	CCGCTGAGCAATAACATGCTATATATCTTTCGTCTTATTATT	
rv_CB_gibson	AAAATTATTTCTACAGGTTACCCTTCGCCCTGCGTCGCACTC	
fw_vector_pRSFCpgpB	GCGCAACGAGAACAAGAAAGTTAACTGCCACCGCTGAGCAATAAC	plsCxpgpB^{H}
rv_vector_pRSFCpgpB	GGGTCGGGGTCGGCGTCGGCGTCGGGGTCGGAACTTTTCCGGCGGCTTCGCGTTC	
fw_pgpB_pRSFCpgpB	GACGCCGACCCCGACCCCGACGCCGACCCCGATGCGTTCGATTGCCAGACG	
rv_pgpB_pRSFCpgpB	GTTATTGCTCAGCGGTGGCAGTTAACTTTCTTGTTCTCGTTGCGCTATTTCG	
fw_mg517_mcherry	ATGCGGATCCCCGGCCTAGGCTG	mCherry-mg517
rv_mg517_cherry	ATGCGAATTCGTTATCTGATTTAGATTCCAAAACATG	
fw_mcherry_mcherry	AAGGGAATTCACGCTAGCGCAACCGGAC	
rv_mcherry_mcherry	TTTTGGATCCCTACTTGTACAGCTCGTCCATGCC	

Primer name	Sequence 5' → 3'	Use
fw_pET44b_policistronic	GTGTGAGATTTTGCGTTAAACGAAAGGCTCAGTCGAAAG	mg517·cdh
rv_pET44b_policistronic	GAAGACCCGCTTTTTTCATTAGTCACCTCCGCTACTGCCGCCAGGCAAATTC	
fw_cdh_policistronic	GAATTTGCCTGGCGGCAGTAGCGGAGGTGACTAATGAAAAAAGCGGGTCTTC	
rv_cdh_policistronic	CTTTCGACTGAGCCTTTCGTTTAACGCAAAATCTCACAC	
fw_fadR_p5T7fadR	ACTTTAATAAGGAGATATACCATGGTCATTAAGGCGCAAAG	$fadR^{L}$
rv_fadR_p5T7fadR	GGTGGCAGCAGCCTAGGTGTCTTATCGCCCCTGAATGGCTAAATCACCC	
fw_p5T7_p5T7fadR	GATTTAGCCATTCAGGGGCGATAAGACACCTAGGCTGCTGCCACC	
rv_p5T7_p5T7fadR	CTTTGCGCCTTAATGACCATGGTATATCTCCTTATTAAAGT	
fw_vector_psd_gg	GTTACCGGTCTCCAACATTATCATCACCACCATCCT	KI plasmid
rv_vector_psd_gg	GTTACCGGTCTCCATTCCGCTTAAAGAACATGTG	
fw_psd_gg	GTTACCGGTCTCGGAATTAGACCTGGTCTTTTTTGTCG	
rv_psd_gg	GTTACCGGTCTCGTGTTAAATTCATTTAAACTTTCGCTACA	
oMEMO8369	GTGGCCTGAAGAGACGTTTGG	Check *psd*
oMEMO1617	TCGCCACCTCTGACTTGAGC	
oMEMO2511_RV_p22_tail	TGTCAAGAAATTTATAAATGAAGC	To linearize KI plasmid to SSA
oMEMO6584	TATAATGTGTACATAAACACAAGCTC	
oMEMO6583	GTGTTTATGTACACATTATANNNNNNNNNNNNNNNNNTGTCAAGAAATTTATAAAT	Promoter library
RBS_psd	TGTACATAAACACCCACMANTCCTACTANSCWCTCCGTACAGGCATGAGCACAAAAAAG	RBS library
fw_psd	TCGTCAACCAATGGGCTGGCGTCGTGTTCTG	linearize plasmid to KI
rv_psd	CCCTTTTTTTCAGGATCCGGCATGTAGGCCGGATCAGGAACGACAAAACAATGGCCTGGAGGCTACCTGCTGCCACCGCTGAGCAATAAC	
fw_check promoter	CGACAGATTCCTGGGATAAG	check promoter and RBS
rv_check promoter	GCATTCATCAGGCGGGCAAG	

15.5. Characterization of mg517xmCherry protein

Fluorescent protein mCherry was fused to MG517 glycosyltransferase and subcloned into pET44b(+) plasmid (see 3.3.3). Using fluorescent proteins as reporter genes is being widely used to locate proteins *in vivo* and study gene expression. The most popular studied and used one to do so is Green Fluorescent Protein (GFP), which was firstly isolated from the jellyfish *Aequorea victoria*. This protein is characterized for producing green light due to an energy transfer reaction. Since its identification, several studies lead to the discovery of new proteins that vary on the emission range being these yellow (YFW), blue (BFP) and red (mPlum, mCherry, mStrawberry, mOrange and mCitrine).

The selected protein for this study was mCherry due that the emission range was different from the C6-NBD-ceramide and this allowed an enzymatic characterization. Its properties are summarized in table 45. (Aliye et al., 2015; Barbier and Damron, 2016; Chalfie, 1995; Shcherbo et al., 2007; Uniprot, n.d.)

Table 49. Properties of mCherry protein

Molecular weight	28.8 KDa
Length	256 aminoacids
Isoelectric point	6.23
λ excitation	540 – 590 nm
λ emission	550 – 650 nm

Prior to use it to compare the GT present in the cells, this new protein was tested and characterized in BL21 Star (DE3). To quantify the presence of MG517, cultures were incubated in the same buffer extraction used for the enzymatic characterization of MG517 for 24 hours before lysing the cells and exciting the samples. The wavelength of excitation used was 550 nm and the maximum emission took place at 607 nm. The equipment required to read and analyze the data was F-2500 Fluorimeter Spectrophotometer (Hitachi®). A comparison between BL21 Star (DE3) cells with and without the fusion protein was performed and the results showed that it was possible to use this construct to compare the expression level protein in the different engineered strains (figure 89A).

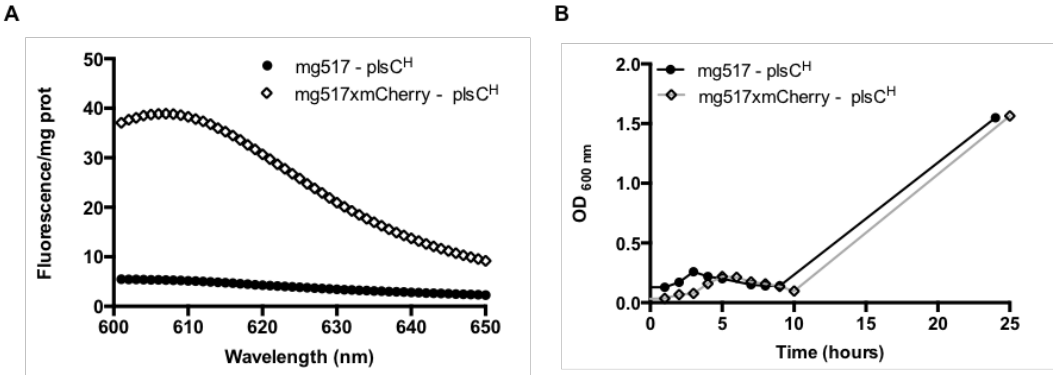

Figure 89. Characterization of MG517-mCherry fusion protein (mg517xmCherry) compared to the wild type glycosyltransferase protein

Moreover, the growth of these same cells was also followed reporting that no differences in the growth profile were found. These results meant that there were no differences in the growth related to using or not the reporter gene *mCherry* (figure 89B).

15.6. Gas chromatography chromatograms and confirmation

Different samples were analyzed by gas chromatography in order to determine the acid methyl esters found in those samples. The samples analyzed were the different engineered strains where since there had genomic modifications it was interesting to determine if those had an effect on the lipid profile.

To confirm the peaks observed in the chromatograms, a CG-MS was also used and compared to a database where the mass spectrometry profile was confirmed. Figure 90 shows a chromatogram of normal gas chromatography using a non-polar column.

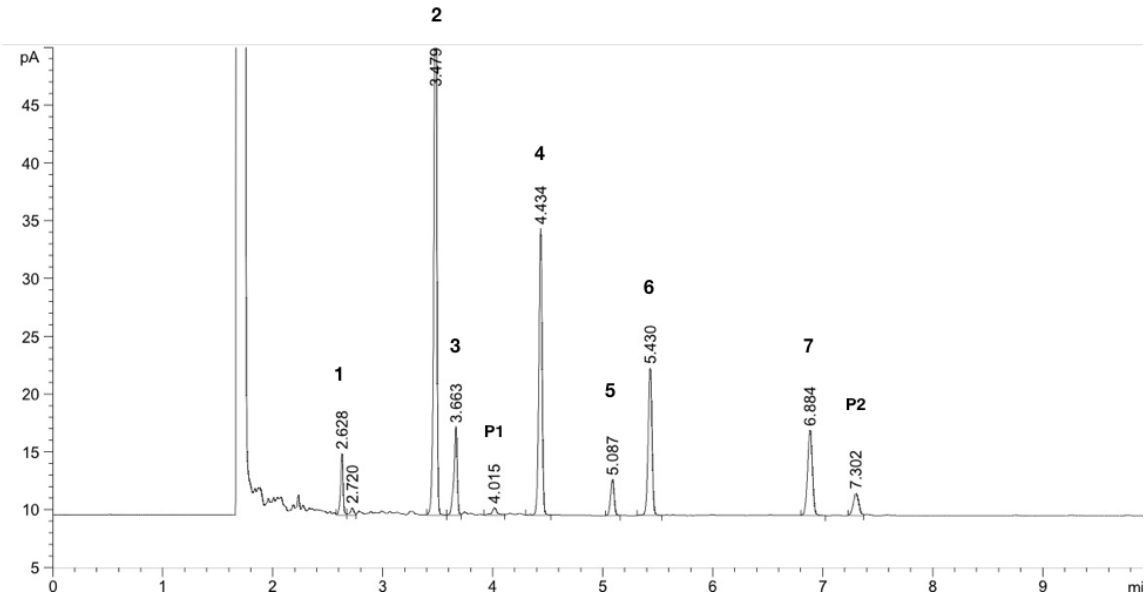

Figure 90. Gas chromatogram of the methyl ester fatty acids

All the methyl esters found were numbered from1 to 7 so it would be easily to follow the confirmation by masses.

Table 50. Identified fatty acids

Fatty acid	Name	Retention time	CAS number	Chromatogram number assigned
C14:0	Myristic acid	2.6 min	124-10-7	1
C16:0	Palmitic acid	3.5 min	2091-29-4	2
C16:1	Palmitoleic acid	3.6 min	57-10-3	3
C17:0Δ	2-hexyl-Cyclopropaneoctanoic acid	4.4 min	5618-00-8	4
C18:0	Stearic acid	5.1 min	112-61-8	5
C18:1	Oleic acid	5.4 min	112-80-1	6
C19:0Δ	2-octyl-Cyclopropaneoctanoic acid	6.8 min	3971-54-8	7

15.6.1. Myristic acid

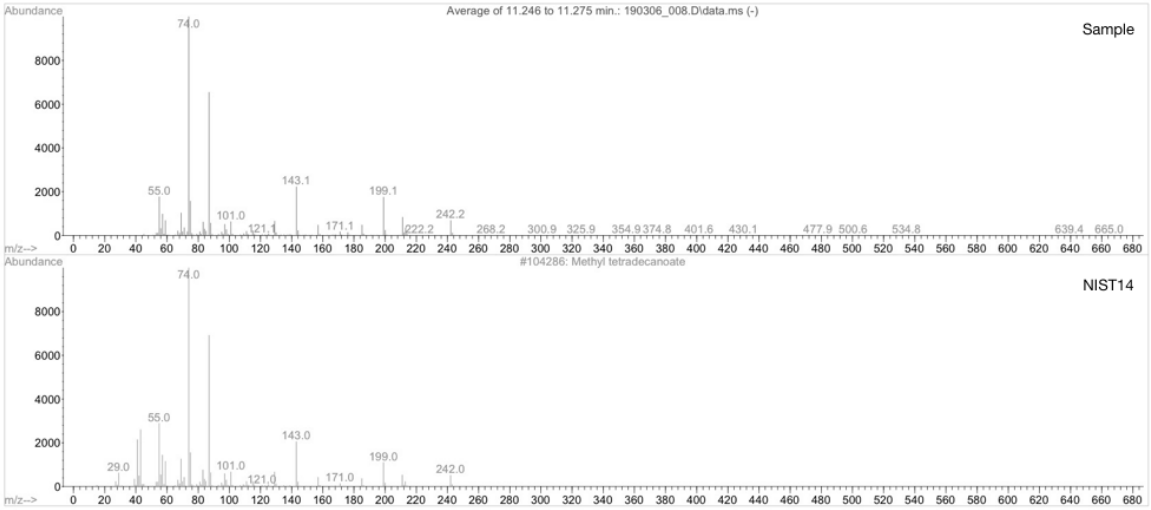

Figure 91. Myristic acid mass spectrum (C14:0)

15.6.2. Palmitic acid

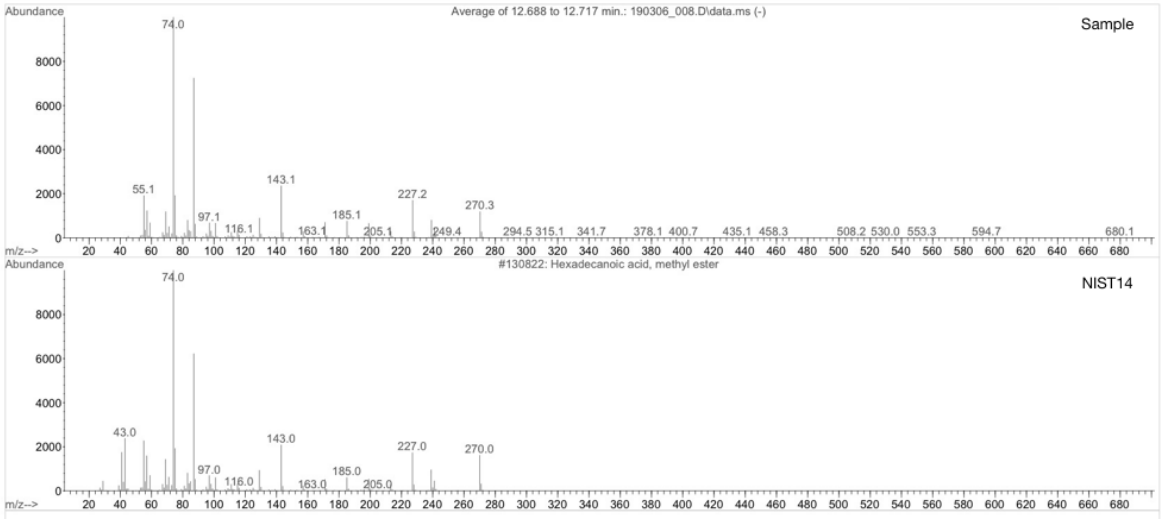

Figure 92. Palmitic acid mass spectrum (C16:0)

15.6.3. Palmitoleic acid

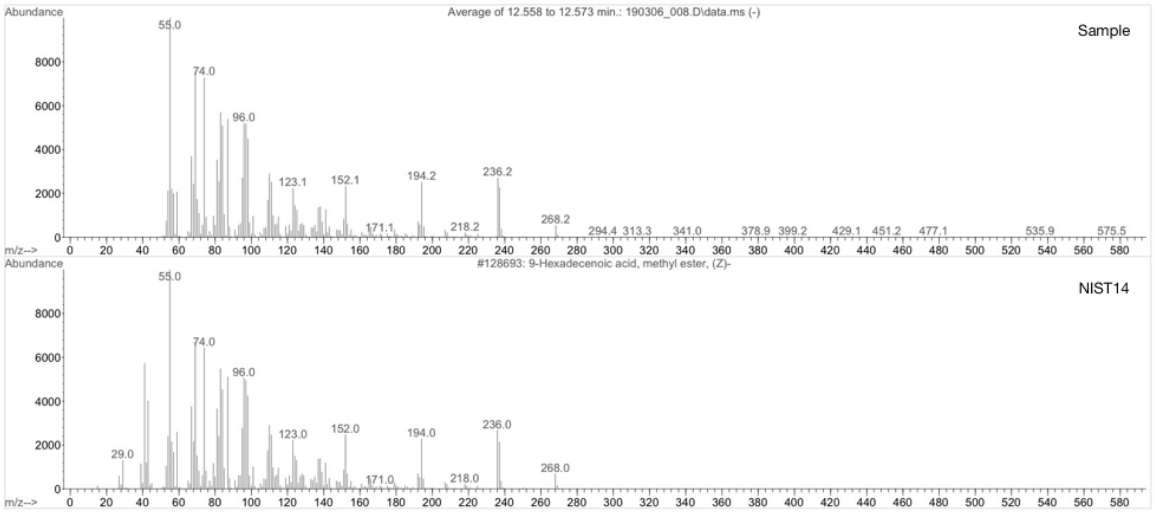

Figure 93. Palmitoleic acid mass spectrum (C16:1)

15.6.4. 2 - hexyl - cyclopropaneoctanoic acid

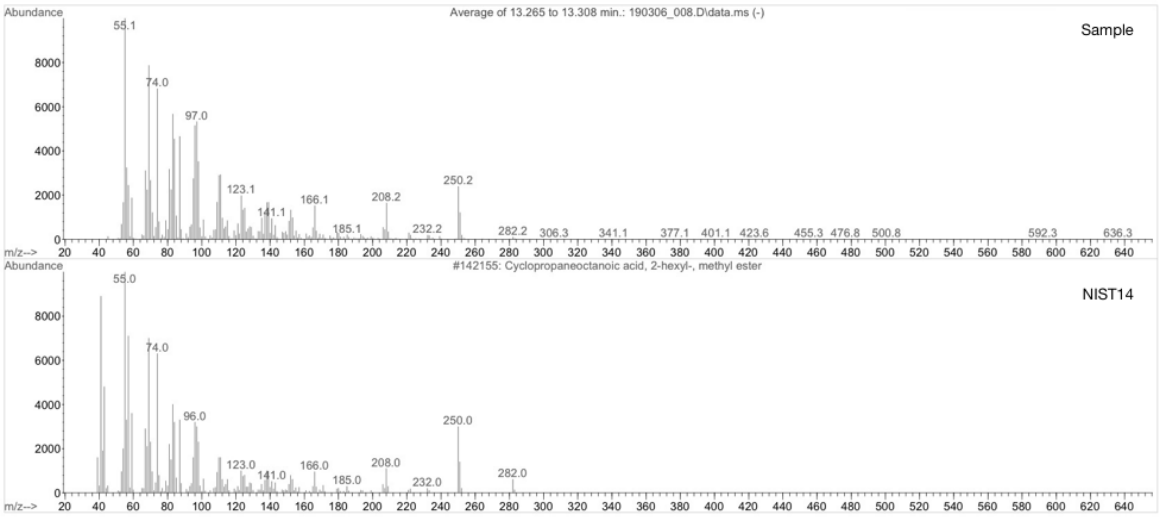

Figure 94. 2 – hexyl – cyclopropaneoctanoic acid mass spectrum (C17:0Δ)

15.6.5. Stearic acid

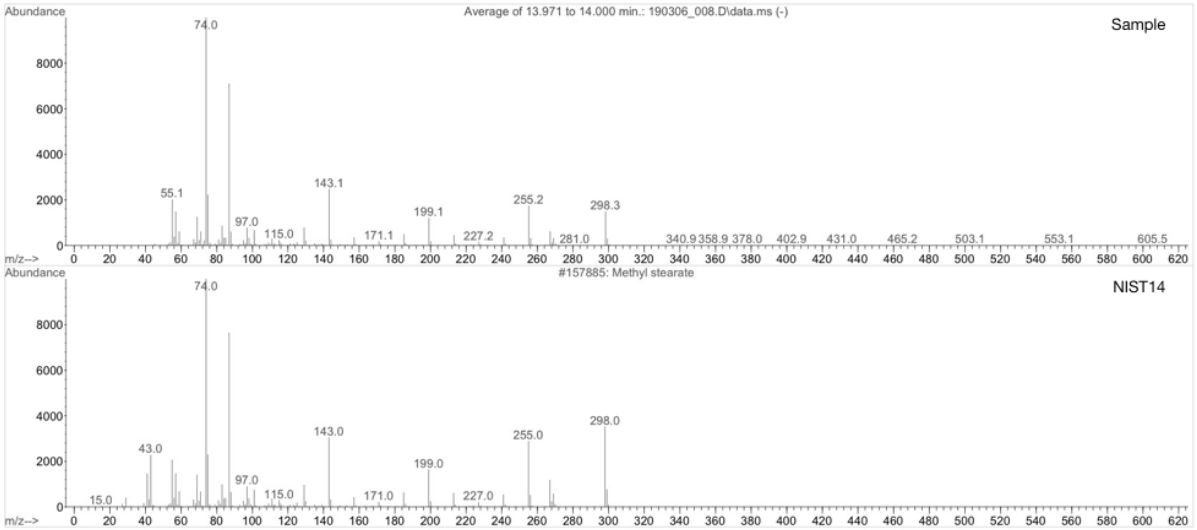

Figure 95. Stearic acid mass spectrum (C18:0)

15.6.6. Oleic acid

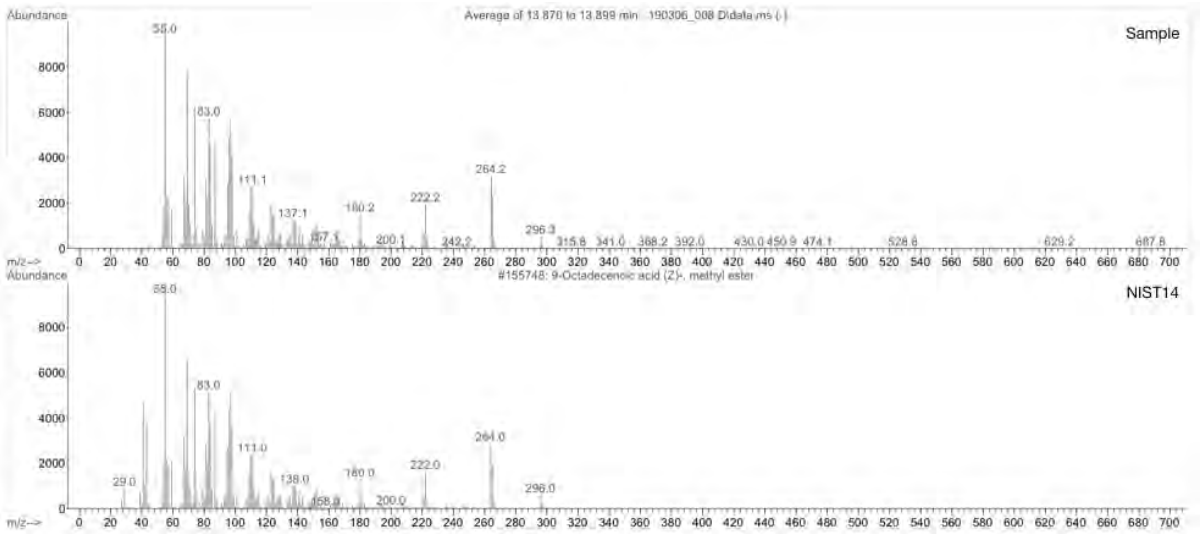

Figure 96. Oleic acid mass spectrum (C18:1)

15.6.7. 2 - octyl - cyclopropaneoctanoic acid

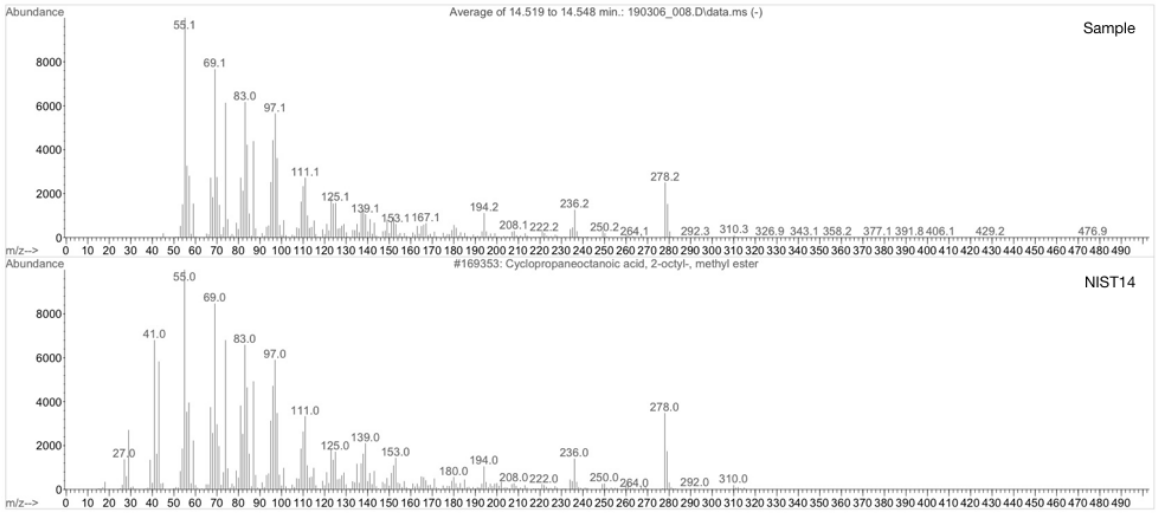

Figure 97. 2 – octyl – cycopropaneoctanoic acid (C19:0Δ)

15.6.8. Subproducts

There were some subproducts observed in these chromatograms that were also analyzed by MS. In figure 97, these subproducts are indicated as P1 and P2. P1 was identified as benzene derivate that could come from the Folch extraction and the use of plastics. On the other hand, the other product observed was an 9-octadecenamide, which is an amide that derivates from oleic fatty acid that can be also be known as oleamide.

15.6.8.1. Benzenepropanoic acid

In the chromatogram (Figure 90) can be found as P1

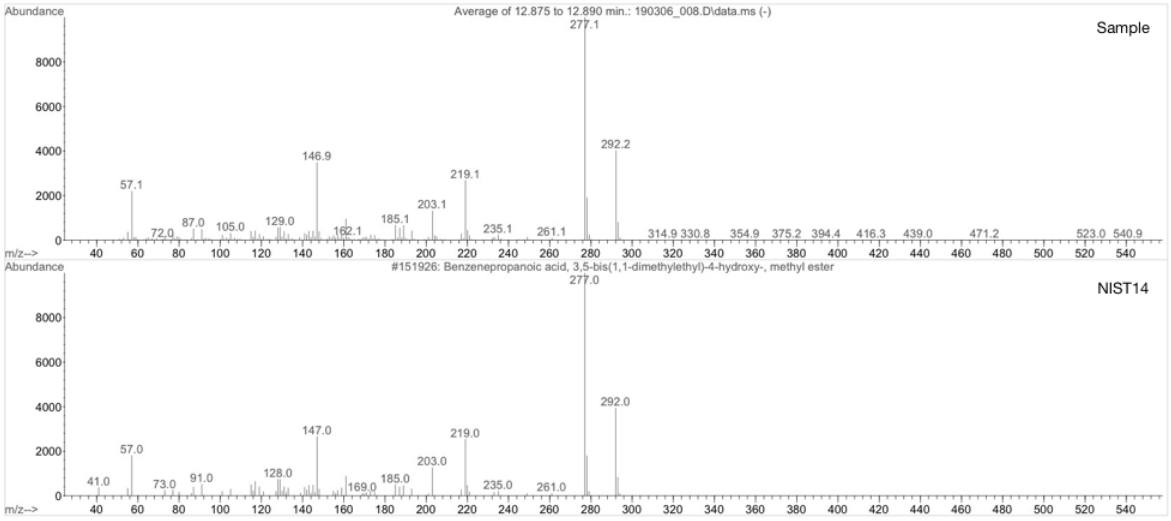

Figure 98. Subproduct 1 mass spectrum identified as benzenepropanoic acid

15.6.8.2. 9-octadecenamide (Oleamide)

In the chromatogram (Figure 90) can be found as P2

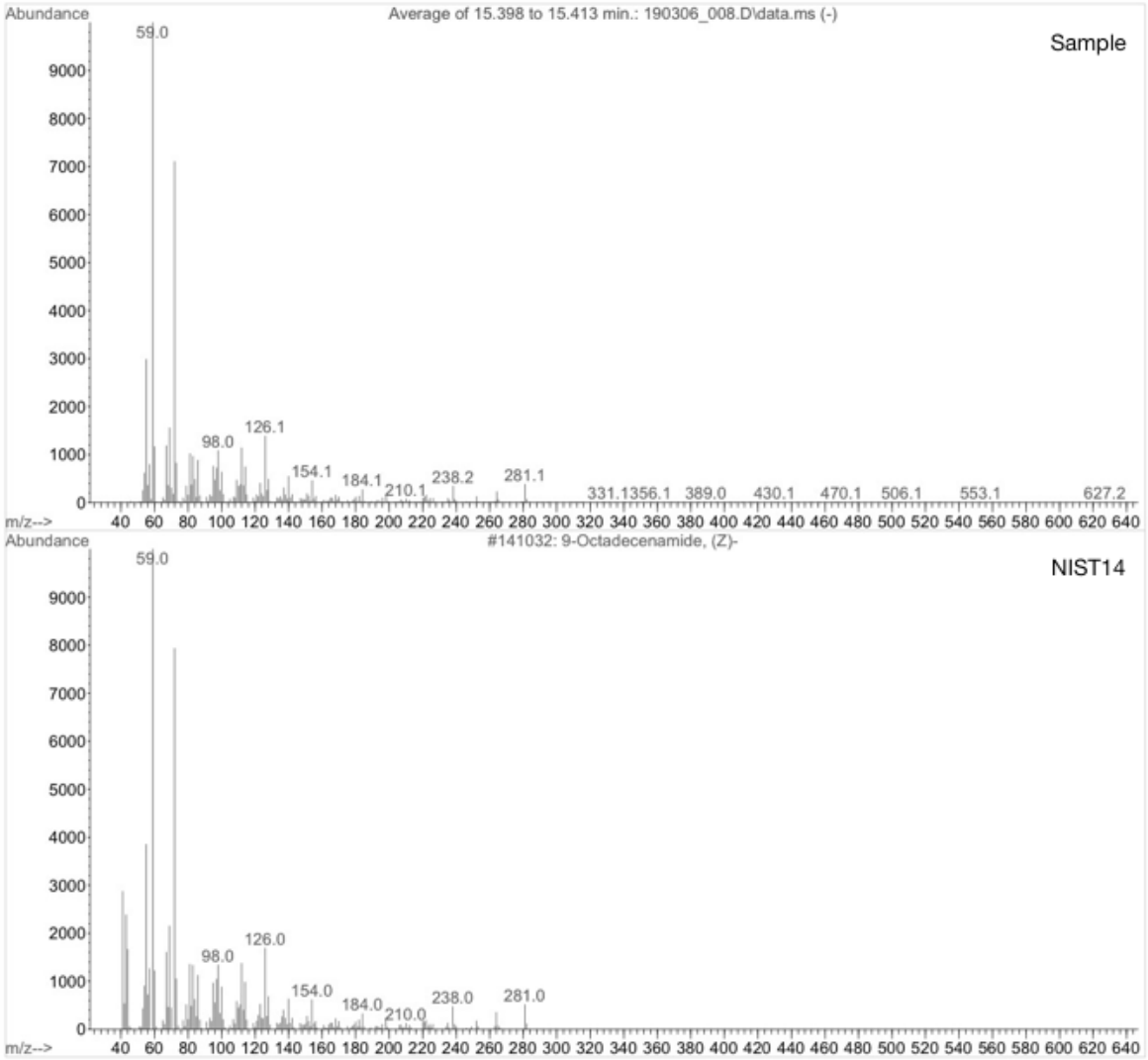

Figure 99. Subproduct 2 mass spectrum identified as oleamide

15.7. Areas and percentage of the radiometric assay

Control	Area	%
Area PE	9	62.7
Area PG + CL	6	37.3
Area GGL	-	-
Total	15	100

(#1) WT/mg517-plsC^{H}	Area	%
Area PE	16	35.2
Area PG + CL	15	34.1
Area GGL	14	30.7
Total	45	100

(#2) ΔtesA/mg517-plsC^{H}	Area	%
Area PE	16	35.2
Area PG + CL	15	34.1
Area GGL	14	30.7
Total	45	100

(#2) ΔtesA/mg517-plsCxpgpB^{H}	Area	%
Area PE	3	12.5
Area PG + CL	10	35.8
Area GGL	14	51.7
Total	27	100

(#2) ΔtesA ΔushA/mg517-plsC^{H}	Area	%
Area PE	14	36.6
Area PG + CL	10	26.4
Area GGL	14	36.9
Total	38	100

16. Publication

Metabolic engineering for glycoglycerolipids production in *E. coli*: Tuning phosphatidic acid and UDP-glucose pathways

Orive-Milla N, Delmulle T, de Mey M, Faijes M, Planas A. *Metab Eng*. 2020;61:106-119. doi:10.1016/j.ymben.2020.05.010

Metabolic Engineering 61 (2020) 106–119

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Metabolic engineering for glycoglycerolipids production in *E. coli*: Tuning phosphatidic acid and UDP-glucose pathways

Nuria Orive-Milla[a], Tom Delmulle[b], Marjan de Mey[b], Magda Faijes[a,*], Antoni Planas[a,**]

[a] Laboratory of Biochemistry, Institut Químic de Sarrià, University Ramon Llull, Via Augusta 350, E-08017, Barcelona, Spain
[b] Centre for Synthetic Biology (CSB), Ghent University, Coupure Links 653, B-9000, Ghent, Belgium

ARTICLE INFO

Keywords:
Glycoglycerolipids
Biosynthesis
Metabolic engineering
E. coli
Diacylglycerol
Glycosyltransferase
Sugar nucleotide
Phosphatidic acid

ABSTRACT

Glycolipids are target molecules in biotechnology and biomedicine as biosurfactants, biomaterials and bioactive molecules. An engineered *E. coli* strain for the production of glycoglycerolipids (GGL) used the MG517 glycolipid synthase from *M. genitalium* for glucosyl transfer from UDPGlc to diacylglycerol acceptor (Mora-Buyé et al., 2012). The intracellular diacylglycerol pool proved to be the limiting factor for GGL production. Here we designed different metabolic engineering strategies to enhance the availability of precursor substrates for the glycolipid synthase by modulating fatty acids, acyl donor and phosphatidic acid biosynthesis. Knockouts of *tes*A, *fad*E and *fab*R genes involved in fatty acids degradation, overexpression of the transcriptional regulator FadR, the acyltransferases PlsB and C, and the pyrophosphatase Cdh for phosphatidic acid biosynthesis, as well as the phosphatase PgpB for conversion to diacylglycerol were explored with the aim of improving GGL titers. Among the different engineered strains, the Δ*tes*A strain co-expressing MG517 and a fusion PlsCxPgpB protein was the best producer, with a 350% increase of GGL titer compared to the parental strain expressing MG517 alone. Attempts to boost UDPGlc availability by overexpressing the uridyltransferase GalU or knocking out the UDP-sugar diphosphatase encoding gene *ush*A did not further improve GGL titers. Most of the strains produced GGL containing a variable number of glucosyl units from mono-to tetra-saccharides. Interestingly, the strains co-expressing Cdh showed a shift in the GGL profile towards the diglucosylated lipid (up to 80% of total GGLs) whereas the strains with a *fad*R knockout presented a higher amount of unsaturated acyl chains. In all cases, GGL production altered the lipidic composition of the *E. coli* membrane, observing that GGL replace phosphatidylethanolamine to maintain the overall membrane charge balance.

1. Introduction

Glycolipids (GL) are widespread products present in the cellular membranes of all organisms as structural components and modulators of cell permeability, stability and rigidity (Holst, 2008). GL are composed of a lipidic moiety linked by glycosidic bond to a sugar that confers them amphipathic properties with many applications in different fields such as biomedicine, cosmetics and industrial biotechnology (Abdel-Mawgoud and Stephanopoulos, 2018; Faivre and Rosilio, 2010). They have low toxicity, high bio-degradability and biocompatibility (Banat et al., 2000), becoming attractive environmentally-friendly biosurfactant substitutes of non-ionic petroleum-based surfactants such as ethylene-oxide derivatives (Bahia et al., 2018)(Chong and Li, 2017). Because of their ability to self-organize into complex supramolecular structures, GL are potential candidates for drug-delivery applications (Corti et al., 2007; He et al., 2019; Muthusamy et al., 2008). Since GL facilitate cellular recognition and are key components of the immune responses, they are used as vaccine adjuvants (*i.e.* lipoarabinomannan as Th1 immunological response enhancers (Behren and Westerlind, 2019)), as immunostimulators (*i.e.* glycoglycerolipids from *M. taiawnensis* in the treatment of autoimmune diseases since they modulate cytokine production and activate/inhibit iNKT cells via Th1 and Th2 responses (Carreño et al., 2014; Ghosh et al., 2013; Liu and Guo, 2017)), and are promising compounds for inhibiting tumor growth (Akasaka et al., 2016, 2013; Maeda et al., 2013).

The classification of glycolipids is based on the lipidic moiety due to the large diversity of glycosylation patterns (Hölzl and Dörmann, 2007; Kalisch et al., 2016; Malhotra, 2012). This paper focuses on glycoglycerolipids (GGL) as potential GL for drug delivery applications and

* Corresponding author.
** Corresponding author.
E-mail addresses: magda.faijes@iqs.edu (M. Faijes), antoni.planas@iqs.edu (A. Planas).

https://doi.org/10.1016/j.ymben.2020.05.010
Received 21 December 2019; Received in revised form 4 May 2020; Accepted 25 May 2020
Available online 31 May 2020

immunotherapy treatments (Abdel-Mawgoud and Stephanopoulos, 2018; Bahia et al., 2018; Chen et al., 2016; Kopitz, 2017). GGL are abundant membrane components in the photosynthetic tissues of plants and cyanobacteria, where they are mainly galactolipids (Andrés et al., 2012; Hölzl and Dörmann, 2007). Wider structural diversity is found in Gram-positive bacteria such *Mycobacterium sp.* or *Lactobacillus plantarum* (Rakhuba et al., 2009; Wicke et al., 2000) where the head group linked to diacylglycerol contains glucosyl, galactosyl or mannosyl units with diverse types of glycosidic linkages (Hölzl and Dörmann, 2007). GGL located in the bacterial cell membrane are capable of exposing their hydrophilic moiety to the surface and be recognized by hormones, antibodies, toxins, virus or other bacteria (Faivre and Rosilio, 2010).

The use of GGL and in general GL in different promising applications is currently limited by the production of these compounds. Different strategies such as chemical synthesis, microbial fermentation or enzymatic synthesis are being addressed for the production of diverse GL derivatives. Despite the progress in chemical synthesis, complex protection and deprotection steps are required to achieve the desired regio- and stereospecificity of the glycosidic linkage, which consequently lower the yield and efficiency of the process (Du et al., 2007). On the other hand, the low amounts extracted from plants or produced by fermentation of microbial natural producers are not suitable for large-scale production (Andrés et al., 2012; Hölzl and Dörmann, 2007; Wicke et al., 2000; Yunoki et al., 2009). We considered metabolic engineering a potential strategy for the production of GL in general and we aimed at building up a metabolic engineering platform for GGL in *E. coli* to achieve these complex structures of interest. Many efforts are being devoted to the biotechnological production of lipids for biofuel applications (Dong et al., 2017; Janßen and Steinbüchel, 2014; Saini et al., 2017; Srirangan et al., 2016) as well as for the production of high-value glycosides (Behren and Westerlind, 2019; Jia et al., 2019; Wang et al., 2015). In our case, since GGL are glycoconjugates with diacylglycerol as the lipidic part, we face the complexity of modulating the diacylglycerol biosynthetic pathway, in which phosphatidic acid (PA) is a shared intermediate with phospholipids biosynthesis and PA availability is affected by the general reactions involving acylCoA. In addition, a suitable glycolipid synthase using diacylglycerol as acceptor has to be introduced to synthesize GGL. MG517 glycosyltransferase from *Mycoplasma genitalium* is a processive (or sequentially-acting) glycolipid synthase from glycosyltransferase family 2 (GT2 according to the Carbohydrate Active Enzyme classification, www.cazy.org (Lombard et al., 2014), which is able to synthesize different GGL from UDP-glucose (UDPGlc) or UDP-galactose (UDPGal) and diacylglycerol (DAG) (Andrés et al., 2012, 2011) (Fig. 1). Our group reported that MG517 was functional when recombinantly expressed in *E. coli*. The new compounds were mono-, di- and triglucosyldiacylglycerols (MGDAG, DGDAG and TriGDAG) that accumulated in the *E. coli* membrane (Andrés et al., 2011). Then, a first generation of engineered strains co-expressing MG517 and enzymes involved in the biosynthetic pathways of UDPGlc and DAG precursors demonstrated that the availability of the DAG was the key bottleneck in GGL production (Mora-Buyé et al., 2012).

Here we report different metabolic engineering strategies to enhance the availability of DAG modulating different precursors, acyl donor, phosphatidic acid and phospholipids, as well as the availability of UDPGlc with the aim of providing a metabolic engineering platform for complex glycoconjugates as GGL.

2. Material and methods

2.1. Bacterial strains and plasmids

The *E. coli* strains and plasmids used in this study are listed in Table 1. DH5α and BL21 Star (DE3) cells (Thermo Fisher) were used as host strains for routine cloning experiments. Genes *pls*C and *gal*U were obtained from *E. coli* JM109 (Promega), *pls*B was from *E. coli* MG1655 (Genbank accession NC_000913) and *E. coli* BL21 Star (DE3) (Genbank accession CP001509) was the source of *fad*R, *cdh*, and *pgp*B genes. BL21 Star (DE3) and its knockouts were used as expression strains. Plasmids pET44b(+), pRSF-1b and pCDF-1b (Novagene), and p10T7 and p5T7 (Ajikumar et al., 2010) were used as expression vectors. Plasmids pCP20, pKD3, pKD4 and pKD46 (Coli Genetic Stock Center (CGSC)) were used to prepare the knockout models following the Datsenko and Wanner method (Datsenko and Wanner, 2000).

2.2. Plasmid construction

Primers for subcloning the genes into the expression plasmids are listed in Table S1, *Supporting Information*. The plasmids containing *pls*B, *pls*C, *gal*U, *pls*C·*pls*B *and mg*517·*cdh* genes were obtained by following the Gibson Assembly approach using primers with 16 bp overlapping regions (Gibson et al., 2009). For the plasmids containing the *pls*C·*pls*B polycistronic gene (both high and low copy number plasmids), an intergenic region with an *Eco*RI site and an RBS sequence (5' GAATTCA AGAAGGAGATATACAT 3') was included. The same cloning procedure was used to prepare the plasmid containing the *mg*517·*cdh* polycistronic gene, with the intergenic sequence containing a new RBS 5' GGATCC CCGGCCTAGGCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGGAG GTGACTA 3'. The plasmid expressing the PlsCxPgpB fusion protein was designed removing the TAA stop codon of the *pls*C gene, adding a sequence encoding the $(PT)_7P$ linker (Kavoosi et al., 2007) and the sequence of *pgp*B (with the ATG start codon). The *pgp*B gene in pRSF-1b plasmid was PCR-amplified (adding the linker sequence) and introduced into the pRSF1b-plsC plasmid by CPEC assembly (Quan and Tian, 2014). For the plasmid expressing the MG517xmCherry fusion protein, first the mCherry gene was amplified from a plasmid under the cytomegalovirus promoter (Shaner et al., 2004) and subcloned into the *Eco*RI and *BamH*I sites of pET44b(+)-*mg*517 by restriction/ligation, introducing a linker (for amino acid sequence EFTLAQPDA) between MG517 (without the TAA stop codon) and mCherry (with ATG start codon) in the fusion protein.

All constructs were verified by DNA sequencing. *E. coli* BL21 Star (DE3) and the different knockouts were then transformed with the different plasmid combinations to obtain the engineered strains (Table 1).

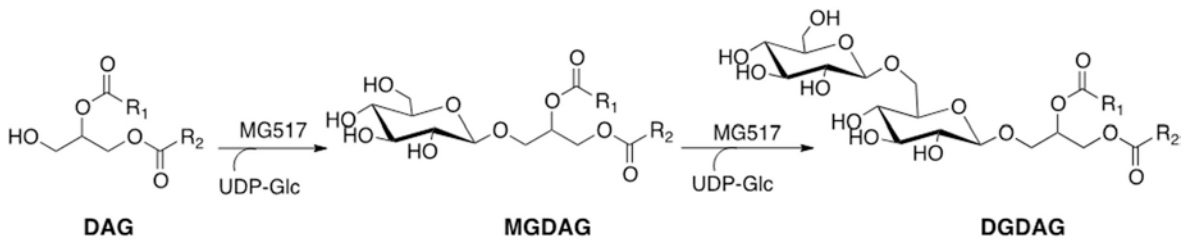

Fig. 1. Reaction catalysed by the glycosyltransferase MG517 from *Mycoplasma genitalium*. DAG = diacylglycerol, MGDAG = monoglycosyldiacylglycerol, DGDAG = diglycosyldiacylglycerol, UDPGlc = UDP-glucose.

Table 1
Engineered strains constructed and used in this work.

	Engineered strain name [a]	Chassis (host strain)	Plasmids[b]
#0	WT/mg517 [c]	BL21 Star (DE3)	pET44b-*mg*517
#1	WT/mg517-plsC^{H} [c]	BL21 Star (DE3)	pET44b-*mg*517 + pRSF1b-*pls*C
#2	ΔtesA/mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C
#3	ΔfadE/mg517-plsC^{H}	BL21 Star (DE3) Δ*fad*E	pET44b-*mg*517 + pRSF1b-*pls*C
#4	ΔtesA ΔfadE/mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fad*E	pET44b-*mg*517 + pRSF1b-*pls*C
#5	ΔtesA ΔfabR/mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fab*R	pET44b-*mg*517 + pRSF1b-*pls*C
#6	ΔtesA/fadR-mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A	p5T7-*fad*R + pET44b-*mg*517 + pRSF1b-*pls*C
#7	ΔtesA ΔfabR/fadR-mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fab*R	p5T7-*fad*R + pET44b-*mg*517 + pRSF1b-*pls*C
#8	ΔtesA/mg517	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517
#9	ΔtesA/mg517-plsC^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*pls*C
#10	ΔtesA/mg517-plsB^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*pls*B
#11	ΔtesA/mg517-plsC^{L}·plsB^{L}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + p10T7-*pls*C·*pls*B
#12	ΔtesA/mg517-plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*B
#13	ΔtesA/mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C·*pls*B
#14	ΔtesA/mg517·cdh-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517·cdh + pRSF1b-*pls*C·*pls*B
#15	ΔtesA/mg517-plsCxpgpB^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C*xpg*pB
#16	ΔtesA/mg517-plsC^{H}-galU	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C + pCDF1b-*gal*U
#17	ΔtesA/mg517-plsCxpgpB^{H}-galU	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517 + pRSF1b-*pls*C*xpg*pB + pCDF1b-*gal*U
#18	ΔtesA ΔushA/mg517-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C
#19	ΔtesA ΔushA/mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C·*pls*B
#20	ΔtesA ΔushA/mg517-plsC^{H}-galU	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C + pCDF1b-*gal*U
#21	ΔtesA ΔushA/fadR-mg517-plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	p5T7-*fad*R + pET44b-*mg*517+ pRSF1b-*pls*C·*pls*B
#22	ΔtesA ΔushA/mg517·cdh -plsC^{H}·plsB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517·*cdh* + pRSF1b-*pls*C·*pls*B
#23	ΔtesA ΔushA/mg517-plsCxpgpB^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517 + pRSF1b-*pls*C*xpg*pB
#24	WT/mg517xmCherry-plsC^{H}	BL21 Star (DE3)	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C
#25	ΔtesA/mg517xmCherry-plsC^{H}	BL21 Star (DE3) Δ*tes*A	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C
#26	ΔfadE/mg517xmCherry-plsC^{H}	BL21 Star (DE3) Δ*fad*E	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C
#27	ΔtesA ΔfadE/mg517xmCherry-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fad*E	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C
#28	ΔtesA ΔfabR/mg517xmCherry-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*fab*R	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C
#29	ΔtesA ΔushA/mg517xmCherry-plsC^{H}	BL21 Star (DE3) Δ*tes*A Δ*ush*A	pET44b-*mg*517x*mCherry* + pRSF1b-*pls*C

[a] Nomenclature: WT or knockout chassis strain/expressed enzymes from plasmids; "·" for proteins expressed from a polycistronic gene; "x" for fusion proteins; "H" for proteins expressed from a high copy number plasmid; "L" for proteins expressed from a low copy number plasmid.
[b] pET44(+) (Amp^{r}), pRSF1b (Kan^{r}) and pCDF1b (Sm^{r}) are high copy number plasmids, while p10T7 (Cm^{r}) and p5T7 (Sm^{r}) are low copy number plasmids.
[c] Reported in (Mora-Buyé et al., 2012). Amp^{r}: ampicillin resistance, Cm^{r}: chloramphenicol resistance, Kan^{r}: kanamycin resistance, Sm^{r}: spectinomycin resistance.

2.3. Knockout strains

All the knockout strains were obtained by following the Datsenko and Wanner guidelines and verified by DNA sequencing (Datsenko and Wanner, 2000). This procedure was based on using helper plasmids to remove genomic regions. In the case of *tes*A it was necessary to remove only the central part of the gene because it overlaps with the two flanking genes. The overlapping sequences for homologous recombination were 50 base long for *fad*E, whereas they were 100 bases long for *ush*A and *fab*R. The doubles knockouts were performed on the BL21 Star (DE3) Δ*tes*A strain following the same procedure. Primers are listed in Table S1.

2.4. Culture conditions and growth media

For DNA manipulations and strain constructions, cultures were grown in lysogeny broth (LB) medium at 37 °C (or 30 °C when indicated in the preparation of knockouts by the Datsenko and Wanner procedure). Ampicillin (100 mg·L^{-1}), kanamycin (50 mg·L^{-1}), chloramphenicol (34 mg·L^{-1}) and streptomycin (100 mg·L^{-1}) were added when required. For GL analysis and determination of enzymatic activities, cultures were grown in minimal medium supplemented with glucose and the corresponding antibiotics, induced when the optical density at 600 nm (OD_{600}) reached between 0.15 and 0.2 with 1 mM final of Isopropyl-β-D-1-thiogalactopyranoside (IPTG) and grown for 24 h at 37 °C and 200 rpm. The minimal medium supplemented with glucose was adapted from Yee and Blanch (1993), containing 25 g·L^{-1} glucose (20%), 4.10 g·L^{-1} $(NH_4)SO_4$, 10.79 g·L^{-1} Na_2HPO_4, 2.34 g·L^{-1} NaH_2PO_4, 2.04 g·L^{-1} NaCl, 0.5 g·L^{-1} $MgSO_4$, 0.01 g·L^{-1} thiamine, 0.114 mg·L^{-1} $AlCl_3{\cdot}6H_2O$, 0.458 mg·L^{-1} $CoCl_2{\cdot}6H_2O$, 0.029 mg·L^{-1} H_3BO_3, 0,029 mg·L^{-1} $NiCl_2{\cdot}6H_2O$, 2.49 mg·L^{-1} $ZnSO_4{\cdot}7H_2O$, 4.43 mg·L^{-1} $CuSO_4{\cdot}5H_2O$, 4.06 mg·L^{-1} $MnCl_2{\cdot}4H_2O$, 0.057 mg·L^{-1} $NaMoO_4$, 4.12 mg·L^{-1} $CaCl_2{\cdot}2H_2O$ and 0.026 mg·L^{-1} $FeCl_3$.

2.5. Sample preparation for glycolipids analysis

Overnight cultures (300 mL) were harvested by centrifugation (13450g 15 min, 4 °C). The cell pellets were washed three times with 0.9% NaCl and freeze-dried. 40 mg of freeze-dried pellets were treated with 800 µL of chloroform and 640 µL of methanol following the Folch extraction methodology (Folch et al., 1957). This mixture was disrupted in an ultrasounds bath for 30 min and filtered. 480 µL of chloroform and 480 µL of deionized water were added to the organic phase and disrupted again for 10 min. After centrifugation (3890g 15 min, 4 °C), 800 µL of the organic phase were transferred to a new tub and the solvent evaporated under a stem of nitrogen.

2.6. Glucose quantification

The glucose content of GGL in the lipid fraction was determined by the anthrone assay (Bailey, 1958; Leyva et al., 2008; Yemm and Willis, 1954). 1 mL of deionized water was added to the previously evaporated lipid fraction and treated with 2 mL of phosphoric acid (85%) at 100 °C for 10 min. After cooling the solution at room temperature, 4 mL of 0.2% (w/v) anthrone reagent in sulphuric acid (96%) were added and the solution boiled for 14 min at 100 °C. Absorbance was measured at 625 nm. Quantification was based on a glucose standard curve ranging 0–150 µg glucose. The quantification limit was 3.1 µg glucose, and the coefficient of variation in all the range was < 12%.

2.7. GGL analysis

The distribution in mono-, di-, tri- and tetra-glycosyldiacylglycerols of the lipid fraction was analyzed by TLC (Silicia gel 60 F_{254}, Merck). TLC plates were developed with chloroform/methanol/water 65:25:4 (v/v) and stained with α-naphtol (0.5% w/v in methanol/water 1:1 (v/v)). After drying, plates were immersed in a mixture of sulphuric acid/methanol/water 45:45:10 (v/v/v) and heated at 100 °C in an oven. Spots were quantified by densitometry using the ImageJ software (Schneider et al., 2012). The total amount of GGL was determined from the glucose content (anthrone assay) and the GGL distribution (TLC analysis) of the lipid fraction according to equation (1):

$$[GGL]_T = \frac{[Glc]_T}{\%\text{MGDAG}+2\cdot\%\text{DGDAG} + 3\cdot\%\text{TriGDAG} + 4\cdot\%\text{TetraGDAG}} \quad (1)$$

where $[GGL]_T$ is the total amount of glycoglycerolipids in mmol/mg dry cells, $[Glc]_T$ is the total amount of glucose in nmol/mg dry cells, and %MGDAG, %DGDAG, %TriGDAG and %TetraGDAG are the percentage distribution of mono-, di-, tri and tetra-glycosyldiacylglycerols.

2.8. Fatty acid analysis

The acyl chain composition was determined as fatty acid methyl esters by gas chromatography. The lipid fraction was dissolved in 2.5 mL hexane and transesterified by adding 25 µL of KOH 2N in methanol. Samples were shaken vigorously and the hexane phase was transferred into a glass tube. Acyl chain methyl esters were analyzed by gas chromatography (Agilent 7890 GC; TRB-WaxOmega (30m x 0.25 mm, 0.25 µm) column; FID detector). Peaks were compared with Supelco 37 component FAME® MIX standard (Merck) featuring 37 fatty acid esters from 14 to 22 carbons and confirmed by mass spectrometry (Agilent 6890N HRGC, HP-5MS UI column, Agilent 5973 Inert Mass Selective detector) using the NIST14 mass spectral library.

2.9. Lipid analysis

The lipid components (phosphatidylethanolamine (PE), phosphatidylglycerol (PG), cardiolipin (CL) and GGL) were quantified by ^{14}C-labelling of the acyl chains, 2D TLC separation and autoradiography. 50 mL cultures were grown in minimal medium containing 10 µCi of [1,2-^{14}C] sodium acetate (3.8 µM) and 20% glucose. After IPTG induction, cultures were grown for 24 h, and cells were harvested and rinsed with 0.9% NaCl. The pellet was suspended in 1 ml chloroform/methanol 2:1 (v/v) mixture, vortexed and disrupted by sonication for 10 min. After centrifugation, the organic phase was separated and evaporated under a nitrogen steam. Then, samples were dissolved with chloroform and the lipid mixture was separated on 2D TLC plates (TLC Silica gel 60 F_{254}, Merck) eluted with ethyl acetate/isopropyl alcohol/chloroform/methanol/0.25% KCl in water (25:25:25:11:9) in the first dimension and chloroform/methanol/water (65:25:4) in the second dimension. Visualization and quantification in mol% were done by electronic autoradiography (Bio-rad Imager) after 16 h of exposition. Lipids were identified by retention times relative to known lipids standard and by positive staining with 0.5% α-naphtol as above (2.8 GGL analysis). Lipids standard used were phosphatidylethanolamine (Sigma), monoglucosyldistearylglycerol (Matreya LLC), diglucosyldistearylglycerol (Matreya LLC), phosphatidylglycerol and cardiolipin (Sigma).

2.10. Preparation of cells extracts for determination of enzymatic activities

50 mL of cell culture grown in minimal medium and induced when OD_{600} was around 0.15, were harvested by centrifugation at 24 h post induction (5000 g, 10 min, 4 °C), washed twice with 0.9% NaCl and suspended in 1.5 mL of extraction buffer containing 20 mM phosphate buffer (pH 6.5), 50 mM NaCl, 10 mM $MgCl_2$ and 1 mM dithiotreitol (Mora-Buyé et al., 2012) for PlsB and PlsC acyltransferase activities, or a buffer containing maleic acid-Tris (25 mM) at pH 7, 20% of glycerol and 50 mM of $MgCl_2$ for PgpB activity. For MG517 activity, an extraction buffer containing 20 mM $MgCl_2$, 20 mM HEPES, 10 mM CHAPS, 20% glycerol, 0.5% NaCl and 1 mM PMSF at pH 8 was used, and the suspension was incubated overnight at 4 °C with agitation. Cells were disrupted by sonication at 0 °C, the cell debris was removed by centrifugation (12470g, 5 min at 4 °C), and the supernatant was used for the enzymatic assays. The total protein was quantified by the BCA assay (Pierce™, ThermoFisher).

2.11. Enzymatic activities determination

PlsC and PlsB acyltransferase activities were determined by monitoring CoA release from oleylCoA acyl donor with the Ellman's reagent (5,5'-dithio-bis-2-nitrobenzoic acid, DTNB) (Mora-Buyé et al., 2012). The reaction mixture contained 0.1 M Tris-HCl (pH 9), 0.5 mM $MgCl_2$, 1 mg·mL^{-1} BSA, 0.1 mM DTNB, 50 mg L^{-1} oleoylCoA and freshly prepared cell extract. The reaction was started by adding 0.1 mg·mL^{-1} of oleoyllysophosphatidic acid for PlsC acyltransferase, or 0.1 mg·mL^{-1} glycerol-3-phosphate for acyltransferase PlsB and the formation of 2-nitro-5-thiobenzoate (TNB) was monitored by absorbance (405 nm) in a microplate reader every 30 s for 10 min at 30 °C. Specific activity values are the average of at least two independent duplicate measurements.

PgpB activity in the PlsCxPgpB fusion protein was determined by monitoring phosphate release in the coupled two-steps PlsC and PgpB reaction from oleoyllysophosphatidic acid as substrate. The reaction mixture contained 0.1 mM of oleoyllysophosphatidic, 1 mM Triton X-100, 2 mM $MgCl_2$ and 10 mM ß-mercaptoethanol and freshly prepared cell extracts (Dillon et al., 1996). The reaction was started by adding oleoyllysophosphatidic and incubated at 30 °C. Aliquots of 80 µL were taken every 2 min, mixed with 20 µL Malachite green reagent (Sigma) and the absorbance was measured at 595 nm.

The MG517 glycosyltyransferase activity was determined using 10 mM HEPES, 5 mM CHAPS, 10% glycerol, 0.25 mM $MgCl_2$ in the presence of 25 µM C6-NBD-ceramide (N-[6-[(7-nitro-2-1,3-benzoxadiazol-4-yl)amino]hexanoyl]-D-erythro-sphingosine), 25 µM BSA, and fresh cell extract. The reaction was started by adding 1 mM UDPGal and incubated at 30 °C. Aliquots were withdrawn every 2 min for 14 min, the reaction was stopped with methanol/water (8:2) and analyzed by HPLC (1200 Agilent equipment, Nova-Pack C18 column, flow rate 1 mL·min^{-1}, eluent 75% acetonitrile in water, fluorimeter detector with excitation wavelength at 458 nm and emission at 530 nm). Chromatographic peaks were assigned by co-injection with independent standards. Initial rates were obtained from the linear progress curve of total product formation.

3. Results and discussion

3.1. Metabolic engineering strategies

The first generation of engineered strains contained different combinations of the *mg*517 with *gal*U and *pls*C genes to overexpress the GalU and PlsC enzymes involved in the biosynthesis of UDPGlc and DAG precursors (Mora-Buyé et al., 2012) (Fig. 2). We showed that the strain overexpressing MG517 and PlsC gave the highest GGL yield and that the DAG pool was limiting in the production of GGL. In this work, different strategies have been developed to impact the availability of the DAG lipidic precursor and, if needed, the sugar nucleotide precursor, with the aim of increasing GGL production yields.

The first strategy was focused on increasing the acyl donor availability by either blocking competing reactions or promoting their biosynthesis (Fig. 2). Acyltransferases PlsB and PlsC, responsible in *E. coli* for the transfer of the acyl groups to *sn*-glycerol-3-phosphate (Gly3P) to give lysophosphatidic (LPA) and phosphatidic acid (PA), respectively,

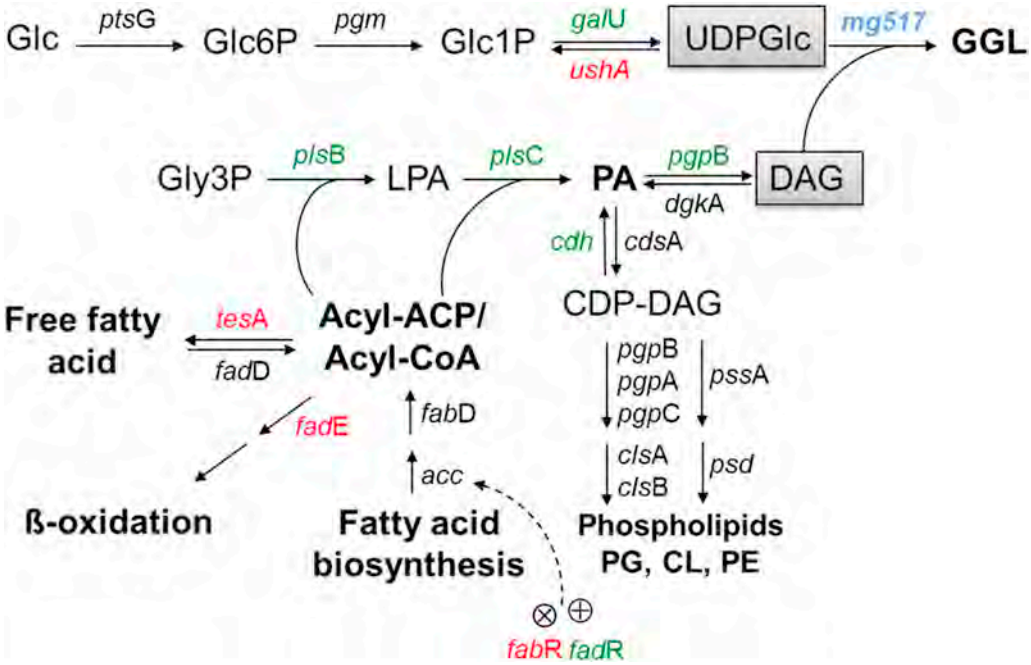

Fig. 2. Metabolic engineering strategies to increase the production glycoglycerolipids (GGL) from UDP-Glucose (UDPGlc) and diacylglycerol (DAG) precursors. In green and in red, overexpression and knockout of different genes, respectively. In blue, *mg517*, common to all strains. Enzymes encoded by the gene presented in the figure: *pts*G: glucose-specific PTS enzyme IIBC component (EC 2.7.1.199); *pgm*: phosphoglucomutase (EC 5.4.2.2); *gal*U: UTP-glucose-1-phosphate uridylyltransferase (EC: 2.7.7.9); *ush*A: UDP-sugar diphosphatase (EC 3.1.3.5.3.6.1.45); *pls*B: glycerol-3-phosphate *O*-acyltransferase (EC 2.3.1.15); *pls*C: 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51); *pgp*B: phosphatidylglycerophosphatase B (EC 3.1.3.27 3.1.3.81, 3.1.3.4, 3.6.1.27); *dgk*A: undecaprenol kinase (EC 2.7.1.66); *cds*A: phosphatidate cytidyltransferase (EC 2.7.7.41); *cdh*: CDP-diacylglycerol diphosphatase (EC 3.6.1.26); *pgp*A: phosphatidylglycerophosphatase A (EC 3.1.3.27); *pgp*C: phosphatidylglycerophosphatase C (EC 3.1.3.27); *cls*A: cardiolipin synthase A (EC 2.7.8.-); *cls*B: cardiolipin synthase B (EC 2.7.8.-); *pss*A: phosphatidylserine synthase (EC 2.7.8.8); *psd*: phosphatidylserine decarboxylase (EC 4.1.1.65); *tes*A: thioesterase I (EC 3.1.2.2); *fad*D: fatty acyl-CoA synthetase (EC 6.2.1.3); *fad*E: acylCoA dehydrogenase (EC 1.3.8.7); *fab*D: malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39); *acc*: acetyl-CoA carboxyltransferase (EC 2.1.3.15); *fab*R: DNA-binding transcriptional repressor FabR; and *fad*R: DNA-binding transcriptional dual regulator FadR. Glc: glucose; Glc6P; glucose-6-phosphate; Glc1P: glucose-1-phosphate, Gly3P: *sn*-glycerol-3-phosphate; LPA: lysophosphatidic acid; PA: phosphatidic acid; ACP: acyl carrier protein; PE: phosphatidylethanolamine; PG: phosphatidylglycerol; CL: cardiolipin (CL). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

use the acyl-acyl carrier protein (acylACP) as well as acyl-CoA thioesters generated by the activation of exogenous fatty acids as acyl donors (Yao and Rock, 2013). Since these acyl donors can be degraded via fatty acid β-oxidation or hydrolyzed to free fatty acids, we first propose to increase the acyl donor pool through the deletion of the *fad*E and/or *tes*A genes involved in these competing reactions (Fig. 2).

AcylCoA dehydrogenase (FadE) is the key enzyme that catalyzes the oxidation of acyl-CoA to 2-enoyl-CoA as the first step of β-oxidation (He et al., 2014). Previous studies in biofuels research reported the importance of knocking out the gene encoding this enzyme to increase the production of free fatty acids (FFA) (He et al., 2014; Janßen and Steinbüchel, 2014; Lu et al., 2008; Steen et al., 2010). Moreover, when combining the *fad*E knockout with overexpression of thioesterases such as *tes*A, FFA productivity was further increased (Bentley et al., 2016; Steen et al., 2010).

Thioesterase I encoded by *tes*A is responsible of catalyzing the hydrolysis of acylated carriers during fatty acid synthesis in order to release FFA. Specifically, this gene encodes a thioesterase type I that is specific for long chain fatty acids (C_{12} – C_{18}) (Lee et al., 1997). It was reported that redirecting the enzyme from the periplasm to the cytosol by removing the leader sequence affected FFA production (Cho and Cronan, 1995). Also, when TesA was overexpressed alone or in combination with a *fad*D knockout or ACC overexpression, an increase of FFA was observed (San and Li, 2017; Steen et al., 2010; Xu et al., 2013). Therefore, we propose that the elimination of the thioesterase activity on acyl donors (Δ*tes*A) together with the acylCoA dehydrogenase (Δ*fad*E), which initiates the β-oxidation pathway after activation of fatty acids by the acylCoA synthethase FadD, would cause an increase in the acyl donor pool and, therefore, the PA precursor to produce DAG.

Secondly, this strategy intended to increase fatty acid availability by modulation of important transcriptional regulators that control both the expression of fatty acid degradation and synthesis genes. The FabR regulator is known to repress *fab*A and *fab*B genes regulating expression of type II fatty acid synthases and, therefore, the synthesis of unsaturated fatty acids (Marrakchi et al., 2002; Zhang et al., 2002). FadR is an activator of fatty acids biosynthesis (*fab* operon) and an inhibitor of their degradation (*fad* operon) (Fujita et al., 2007; Karp et al., 2014). Different studies showed that *fad*R overexpression resulted in a significant increase of FFA production in comparison to overexpression of individual genes such as *fab*A or *fab*B (He et al., 2014; Zhang et al., 2012). Moreover, *fad*R overexpression in combination with *fab*R deletion produced a significant improvement in fatty acid titer (San and Li, 2017). In this context, we propose knocking out *fab*R and overexpressing *fad*R with the aim of increasing the acyl donor availability by transcriptional modulation in the Δ*tes*A/mg517 background.

The second strategy aimed at increasing the conversion of acyl donors to PA through the overexpression of PlsC and PlsB acyltransferases. The first reaction to produce PA is catalysed by PlsB, which converts Gly3P into LPA. Once this product is obtained, acyltransferase PlsC is able to add a second acyl moiety to form PA, which is the precursor of DAG and the initial precursor of phospholipids biosynthesis (Parsons and Rock, 2013; Yao and Rock, 2013) (Fig. 2). Our group had previously reported an increase in the production of GGL when PlsC acyltransferase was overexpressed along with MG517 glycosyltransferase (Mora-Buyé et al., 2012). These acyltransferases have also shown to be key targets to increase the production of other lipid derivatives such as triglycerides and polydroxyalkanoates in *Yarrowia lipolytica* and *Rhodosporidium toruloides* (Celińska et al., 2019; Czerwiec et al., 2019; Park et al., 2018; Rigouin et al., 2019). In this approach, both high and low copy number plasmids expressing PlsC and PlsB were used in order to evaluate the effect of the expression levels in the Δ*tes*A/mg517 background.

The third strategy intended to increase PA production from phospholipids and use it for DAG synthesis. The overexpression of the hydrolase CDP-diacylglycerol pyrophosphatase Cdh, enzyme that catalyzes the conversion of CDP-DAG to PA, could increase PA availability for conversion to DAG at expenses of decreasing CDP-DAG for the

synthesis of phosphatidylethanolamine (PE) and other essential phospholipids (Fig. 2). This is especially interesting since we and others have shown that it is possible to interchange PE and monoglucosyldiacylglycerol (MGDAG) in the membranes (Mora-Buyé et al., 2012; Xie et al., 2006) thus being a potential strategy to further increase the production of GGL. Alternatively, PA can be pushed towards DAG by the overexpression of the phosphatidic phosphatase PgpB, a key regulatory enzyme in lipid metabolism, responsible for the conversion of PA to DAG. However PgpB is also involved in the last step of phosphatidylglycerol biosynthesis (Wikstrom et al., 2009), which is a competing reaction for our purpose (Fig. 2). Therefore, we propose the overexpression of a fusion protein plsCxPgpB in order to redirect the flux towards DAG and not to phospholipids biosynthesis.

If these strategies significantly enhance DAG availability, the UDPGlc precursor might become the limiting factor for GGL biosynthesis. Complementary strategies will be considered in section 3.3.

3.2. Improvement of DAG synthesis pathways

New *E. coli* strains were constructed according to the three different strategies to modulate FFA, acyl donor and PA biosynthesis/degradation in order to increase the DAG pool or availability for GGL production (Table 1). All them expressed the MG517 glycosyltransferase and were grown in minimal medium supplemented with glucose and the proper antibiotics, and induced with IPTG when the OD_{600} was around 0.15–0.2 at 37 °C. This induction caused a new transitory stationary phase that lasted up to 8 h after which cells continued growing. The phenotypes of the recombinant strains were characterized by determination of GGL production, and distribution, fatty acid profile and enzymatic activities at 24 h.

3.2.1. First strategy: increase of acyl donor

As first strategy, knockouts of *fad*E and *tes*A were expected to have an impact on the acyl donor pool and increase DAG and GGL production by avoiding fatty acid ß-oxidation and acyl donor hydrolysis. Knockouts of *tes*A, *fad*E and the double knockout co-expressing MG517 and PlsC (ΔtesA/mg517-plsC^H, ΔfadE/mg517-plsc^H, ΔtesA ΔfadE/mg517-plsC^H, Table 2, strains #2 to #4) were obtained and compared to the wild-type strain (WT/mg517-plsC^H, strain #1). The growth rates of these strains without induction were similar, with specific growth rates (μ_{max}) in the range of 0.21 h^{-1} (WT) to 0.26 h^{-1} (Δ*tes*A strains). Table 2 summarizes the production of GGL in nmol·mg^{-1} of dry cells and the GGL distribution in mono-, di-, tri- and tetra-glycosyldiacylglycerols of the different engineered strains. The WT strain co-expressing MG517 and PlsC (reference strain #1) produced 3.2 nmol·mg^{-1} of GGL, a higher production than that obtained when MG517 was expressed alone as previously reported (strain #0) (Mora-Buyé et al., 2012). A significant 1.8-fold increase was observed in the Δ*tes*A strain (#2) compared to the WT strain (#1), meaning that acyl donor availability has been significantly increased by removing the thioesterase activity as a competing reaction (Fig. 3). By contrast, the Δ*fad*E strain (#3), with blocked aerobic fatty acids ß-oxidation, only gave a slight improvement of GGL production, which indicates that little fatty acid ß-oxidation occurs when using aerobic growth conditions in minimal medium with glucose as carbon source. However, the double knockout, Δ*tes*A Δ*fad*E strain (#4), had a lower productivity than the WT, suggesting that knocking out simultaneously the initial steps of two AcylCoA-utilizing pathways perturbs the regulation of this central metabolic node with a highly detrimental effect on our target GGL productivity.

In addition, this strategy intended to increase the acyl donor availability through transcriptional regulation by knocking out *fab*R or/and overexpressing *fad*R, which were expected to increase acyl donor biosynthesis and consequently DAG availability. Fig. 3 shows that the Δ*tes*A Δ*fab*R strain co-expressing MG517 and PlsC (Table 2, #5) did not improve GGL production compared to Δ*tes*A (#2) indicating that the deletion of acylCoA biosynthesis repression has not an impact on GGL production. However, the overexpression of the *fad*R regulator, which represses fatty acid degradation and activates biosynthesis, in the Δ*tes*A and Δ*tes*A Δ*fab*R strains (#6, #7) resulted in a slight increase in GGL production.

3.2.2. Second strategy: increase of acyl donor conversion to phosphatidic acid

In this strategy aimed at increasing PA biosynthesis and hence DAG availability, the overexpression of the first acyltransferase PlsB, the second acyltransferase PlsC, and co-expression of both enzymes was tested in the best producer Δ*tes*A strain. PlsC and PlsB were expressed in low and high copy number plasmids to compare the effect of their expression levels. Acyltransferase activities measured at 24 h indicated that the expression of these enzymes in the Δ*tes*A strains was increased 2- and 30-fold for PlsB^L and PlsB^H, respectively, and about 300-fold for PlsC^L and PlsC^H compared to the genome-encoded enzymes (untransformed BL21(DE3) cells). Fig. 4 summarizes GGL production by these strains. In agreement with previous results (Mora-Buyé et al., 2012), the overexpression of acyltransferases is a key point to increase GGL production suggesting that PA availability from acyl donor and Gly3P is increased and has an effect on DAG. Whereas overexpression of any of both acyltransferases in high copy number plasmids (as well as co-expression of both) raises GGL production up to 5 nmol·mg^{-1} (a 2-fold increase relative to the ΔtesA/mg517 strain (#8)), both enzymes need to be co-expressed to achieve similar productivity in low copy number plasmids, (Fig. 4).

3.2.3. Third strategy: increase of DAG by enhancing PA availability

Since PA is the direct precursor of DAG, either an increase of the PA level by hydrolysis of phospholipid precursors through CDP-diacylglycerol pyrophosphatase (Cdh) overexpression or pulling PA towards DAG by phophatidic acid phosphatase (PgpB) overexpression might positively impact the DAG level or availability for GGL synthesis (Fig. 2). Cdh was overexpressed in a polycistronic construction along with MG517. Fig. 5 shows that the Δ*tes*A strain co-expressing Cdh, MG517 and both acyltransferases did not have a significant effect in the total GGL production compared to the strain where Cdh was not overexpressed (Table 2, #14 *vs.* #13), but presented a different GGL profile, with DGDAG accounting for 86% of all GGL products (see section 3.5). PgpB was overexpressed as a fusion protein with PlsC to promote a proximity effect of the two consecutive metabolic steps (Table 2, strain #15). The PgpB activity (coupled to PlsC activity in the fusion protein) in this Δ*tes*A strain was almost 8-fold higher than in the parental strain (strain #2). As shown in Fig. 5, overexpression of the PlsCxPgpB fusion protein resulted in a 1.7-fold increased GGL production relative to the parental strain (up to 9.9 nmol·mg^{-1}), which is 3-fold higher than the WT strain (#1). This is the best GGL producer from the engineered strains, indicating that the combined effect of the acyltransferase PlsC and the phosphatase PgpB on a ΔtesA/mg517 background drives the pathway to DAG, which becomes more readily available as substrate for the glycosyltransferase MG517, thus increasing GGL production.

In conclusion, among the different strategies intended to pull the synthesis towards DAG, the elimination of FFA hydrolysis competing reactions (Δ*tes*A) along with the overexpression of acyltransferases and PgpB are the key points to increase acyl donor, PA and DAG availabilities, respectively. The impact of Δ*tes*A and acyltransferase overexpression is in agreement with other works on free fatty acids and triglycerides production (Bentley et al., 2016). Remarkably, PgpB overexpression as PlsCxPgpB fusion protein is first reported in this work and has a large impact in GGL production. As compared to co-expression of both acyl transferases (PlsB and PlsC, strain #13), only the second is required to achieve this high productivity when combined with PgpB. The other strains here studied, even though all improved GGL production as compared to the initial WT strain (#0) only expressing MG517, did not surpass the 6 nmol·mg^{-1} in total GGL products.

Table 2
Glycoglycerolipids production and composition by the engineered strains. GGL: glycoglycerolipids.

	Strain (genotype) [a]	OD [b]	$[Glc]_T$ [c] (μg/mg)	%GGL composition [d]				$[GGL]_T$ [e] (nmol/mg)	GGL [f] strain/wt(#1)
				%M	%D	%Tri	%Tetra		
#0	WT/mg517 [g]	3.6 ± 0.1	0.84	36	31	15	18	2.17 ± 0.50	0.7
#1	WT/mg517-plsC^H [g]	1.7 ± 0.5	1.61	32	26	18	34	3.26 ± 0.60	1.0
#2	ΔtesA/mg517-plsC^H	1.7 ± 0.5	2.47	9	55	22	14	5.69 ± 1.40	1.7
#3	ΔfadE/mg517-plsC^H	3.1 ± 0.3	1.51	20	47	20	13	3.72 ± 0.70	1.1
#4	ΔtesA ΔfadE/mg517-plsC^H	3.1 ± 0.7	0.92	31	39	8	22	2.31 ± 0.70	0.7
#5	ΔtesA ΔfabR/mg517-plsC^H	2.1 ± 0.3	2.04	18	50	12	20	4.84 ± 0.70	1.5
#6	ΔtesA/fadR-mg517-plsC^H	1.7 ± 0.2	2.40	12	60	12	16	5.75 ± 0.50	1.8
#7	ΔtesA ΔfabR/fadR-mg517-plsC^H	1.9 ± 0.1	2.29	14	64	17	6	5.86 ± 0.50	1.8
#8	ΔtesA/mg517	3.2 ± 0.8	0.92	32	42	17	6	2.65 ± 0.50	0.8
#9	ΔtesA/mg517-plsC^L	2.7 ± 0.3	0.92	16	40	31	14	2.11 ± 0.50	0.6
#10	ΔtesA/mg517-plsB^L	2.0 ± 0.6	1.14	16	47	24	13	2.72 ± 0.70	0.8
#11	ΔtesA/mg517-plsC^L-plsB^L	1.5 ± 0.3	2.30	11	50	29	11	5.35 ± 0.50	1.6
#12	ΔtesA/mg517-plsB^H	1.7 ± 0.2	2.06	17	54	17	12	5.11 ± 1.20	1.6
#13	ΔtesA/mg517-plsC^H-plsB^H	1.9 ± 0.3	2.72	4	45	40	10	5.95 ± 0.60	1.8
#14	ΔtesA/mg517-cdh-plsC^H-plsB^H	1.7 ± 0.1	2.20	4	86	6	5	5.71 ± 1.40	1.8
#15	ΔtesA/mg517-plsCxpgpB^H	1.6 ± 0.2	4.83	4	45	29	23	9.93 ± 1.20	3.0
#16	ΔtesA/mg517-plsC^H-galU	1.7 ± 0.2	4.30	25	21	20	34	9.08 ± 0.90	2.8
#17	ΔtesA/mg517-plsCxpgpB^H-galU	1.9 ± 0.3	2.45	10	50	18	21	5.44 ± 1.50	1.7
#18	ΔtesA ΔushA/mg517-plsC^H	2.2 ± 0.6	3.41	6	71	13	9	8.63 ± 1.50	2.6
#19	ΔtesA ΔushA/mg517-plsC^H-plsB^H	1.7 ± 0.1	4.02	2	63	15	18	9.13 ± 0.70	2.8
#20	ΔtesA ΔushA/mg517-plsC^H-galU	1.8 ± 0.2	2.85	33	32	19	16	7.28 ± 0.60	2.2
#21	ΔtesA ΔushA/fadR-mg517-plsC^H-plsB^H	2.9 ± 0.1	2.28	26	34	16	25	5.29 ± 0.50	1.6
#22	ΔtesA ΔushA/mg517-cdh -plsC^H-plsB^H	1.8 ± 0.2	1.43	6	82	9	3	3.81 ± 0.30	1.2
#23	ΔtesA ΔushA/mg517-plsCxpgpB^H	1.7 ± 0.1	3.09	6	63	17	13	7.22 ± 0.70	2.2

[a] Strains and plasmids description in Table 1.
[b] Optical density at 600 nm of the cultures after 24h growth when harvested to evaluate GGL production.
[c] Total sugars in the lipid extract quantified by the anthrone method and expressed in μg glucose equivalents per mg of dry cells.
[d] %GGL composition of the lipid extract by TLC densitometry. M: mono-, D: di-, Tri: tri-, Tetra: tetra-glycosyldiacylglycerols.
[e] Total glycoglycerolipids productivity in nmol GGL per mg of dry cells.
[f] GGL productivity relative to the reference WT strain #1.
[g] Strains reported in (Mora-Buyé et al., 2012).

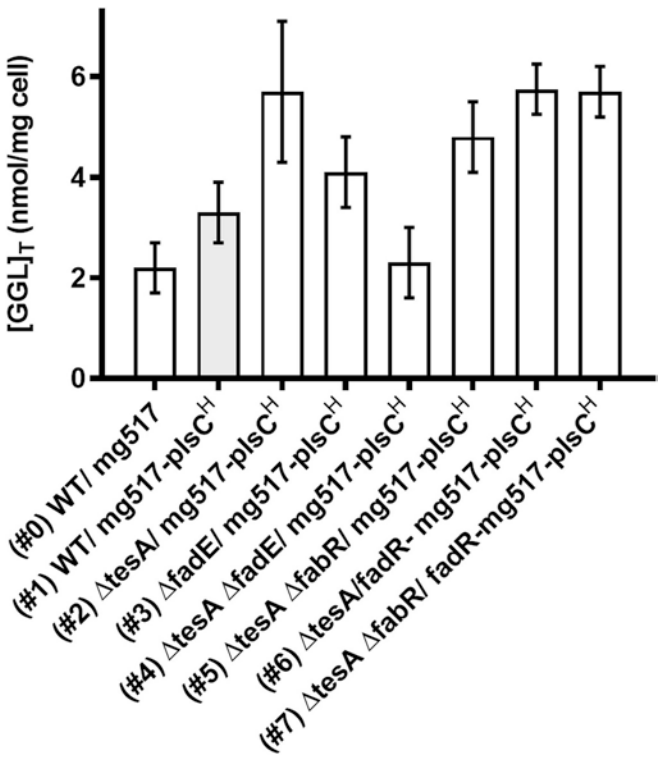

Fig. 3. Glycoglycerolipids (GGL) production from the engineered strains with the different knockouts Δ*tes*A, Δ*fad*E, Δ*fad*E Δ*tes*A and Δ*tes*A Δ*fab*R (1st strategy).

Elimination of the fatty acid competing reactions with *fad*E (in the β-oxidation pathway) or *fab*R (transcriptional regulator) knockouts or stimulation of fatty acid biosynthesis with *fad*R overexpression did not have such an impact of GGL production. Likewise, overexpression of the pyrophosphatase Cdh, which may divert phospholipid precursors to PA, did not significantly contribute to increase GGL levels, although it changed membrane glycolipids distribution in the engineering strain. As far as we know, no previous studies have evaluated the effects of Cdh overexpression in engineered cells. The closest example is the overexpression of a related enzyme involved in ceramide biosynthesis in yeast, the inositol phosphosphingolipid-phospholipase Isc1, which showed a 4-fold increase in ceramide levels (Murakami et al., 2015).

3.3. Increasing UDPGlc availability

DAG and UDPGlc are the substrates of the glycosyltransferase GT MG517 to produce GGL. Once improvement of DAG availability by modulation of the lipidic pathway resulted in higher GGL production, UDPGlc availability may become limiting. The impact of overexpressing the UDP-glucose pyrophosphorylase GalU was studied on different DAG-engineered, GGL producing strains. First, the Δ*tes*A strain overexpressing MG517, PlsC and GalU in high copy number plasmids (Table 2, strain #16) was characterized. While GalU overexpression had no effect on GGL production in the WT strain co-expressing MG517 and PlsC (Mora-Buyé et al., 2012), it did have a major impact in the Δ*tes*A strain, meaning that UDPGlc became the limiting substrate (Fig. 6). This strain reached 9 nmol·mg^{-1} GGL products, which is 1.6-fold higher than the parental strain without GalU overexpression (#2). Then, the obvious step forward was to test if combining GalU with the best previous strain with regard to DAG availability could further increase GGL productivity. In this way, introduction of the plasmid encoding GalU into the ΔtesA/mg517-plsCxpgpB^H strain (#15) gave

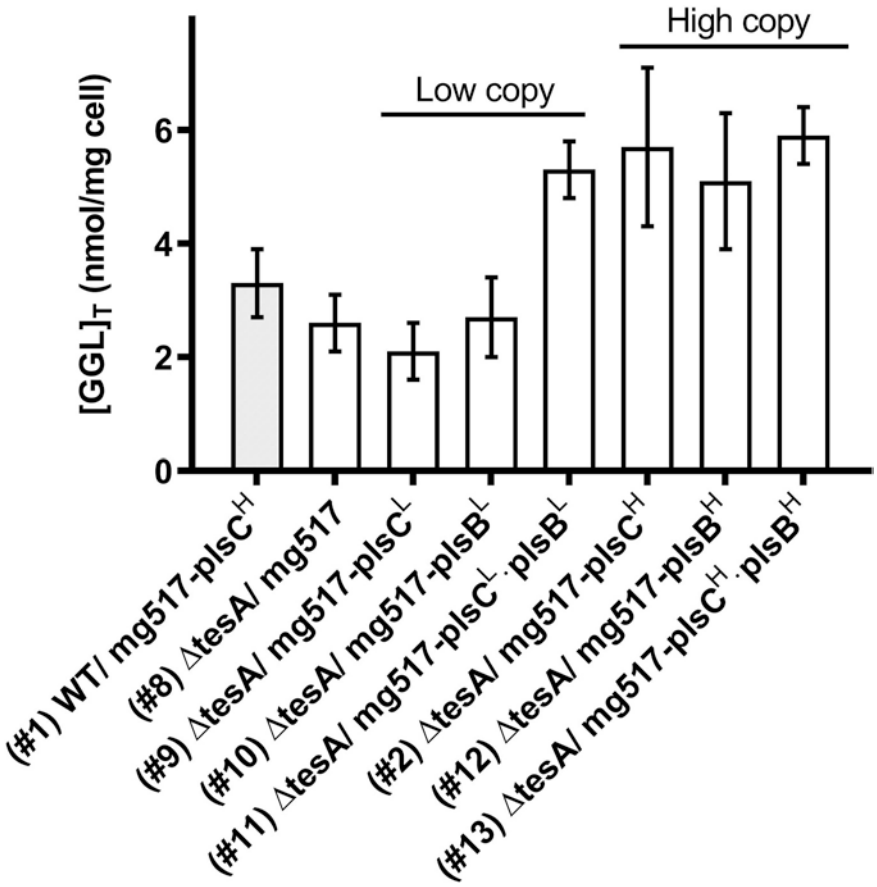

Fig. 4. Glycoglycerolipids (GGL) production in the engineered Δ*tes*A strains when cells are co-expressing MG517 with PlsB, PlsC or both acyltransferases in low or high copy number plasmids (2nd strategy).

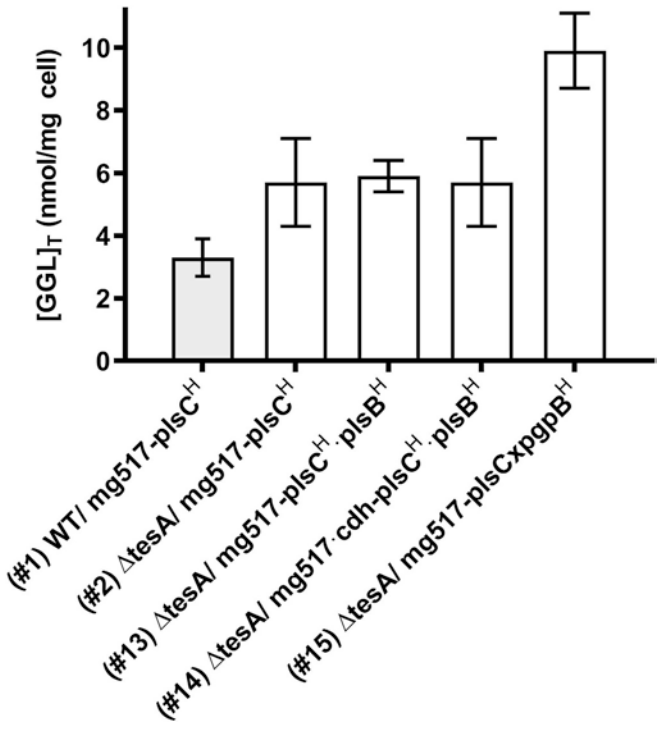

Fig. 5. Effect of Cdh and PgpB overexpressions on glycoglycerolipids (GGL) production in the Δ*tes*A strains (3rd strategy).

strain #17. However, no synergic effects were observed, but rather GGL productivity significantly decreased to 5.4 nmol·mg^{-1}, which was still 1.7-fold higher that the WT strain #1.

A different strategy to impact the UDPGlc level or availability is the inactivation of the enzyme catalyzing the reverse step, the UDP-sugar hydrolase encoded by *ush*A (Fig. 2), since it has been reported that a *ush*A knockout increased the UDPGlc pool in *E. coli* (De Bruyn et al., 2015; Pandey et al., 2014). The Δ*tes*A Δ*ush*A strains overexpressing MG517 and the acyltransferases (Table 2, strains #18, #19) provided a similar effect as GalU overexpression in the Δ*tes*A strain (#16), obtaining around 9 nmol·mg^{-1} of GGL (Fig. 6). This means that prevention of UDPGlc hydrolysis also improves UDPGlc availability for GGL synthesis. However, overexpression of GalU in this Δ*tes*A Δ*ush*A background (strain #20) did not further increase GGL production. The lack of synergic effect of GalU overexpression and *ush*A knockout suggests that UDPGlc reached a maximum level.

As it happened with the Δ*tes*A strains, overexpression of FadR or Cdh in the Δ*tes*A Δ*ush*A background (strains #21 and #22, respectively) did not increase (or slightly decreased) GGL production, meaning that Cdh or FadR overexpression did not affect directly the DAG availability, no matter the UDPGlc limitation (Fig. 7). Again, the strain overexpressing Cdh (strain #22) gave a different GGL profile, with 82% of the total GGL being DGDAG (see section 3.5).

Finally, overexpression of the fusion protein PlsCxPgpB that gave the best results in the Δ*tes*A strain (#15) was checked in the new Δ*tes*A Δ*ush*A background (Fig. 6, strain #23). It did not reach the same high GGL production but 7.2 nmol·mg^{-1} was obtained, being still a good producer strain.

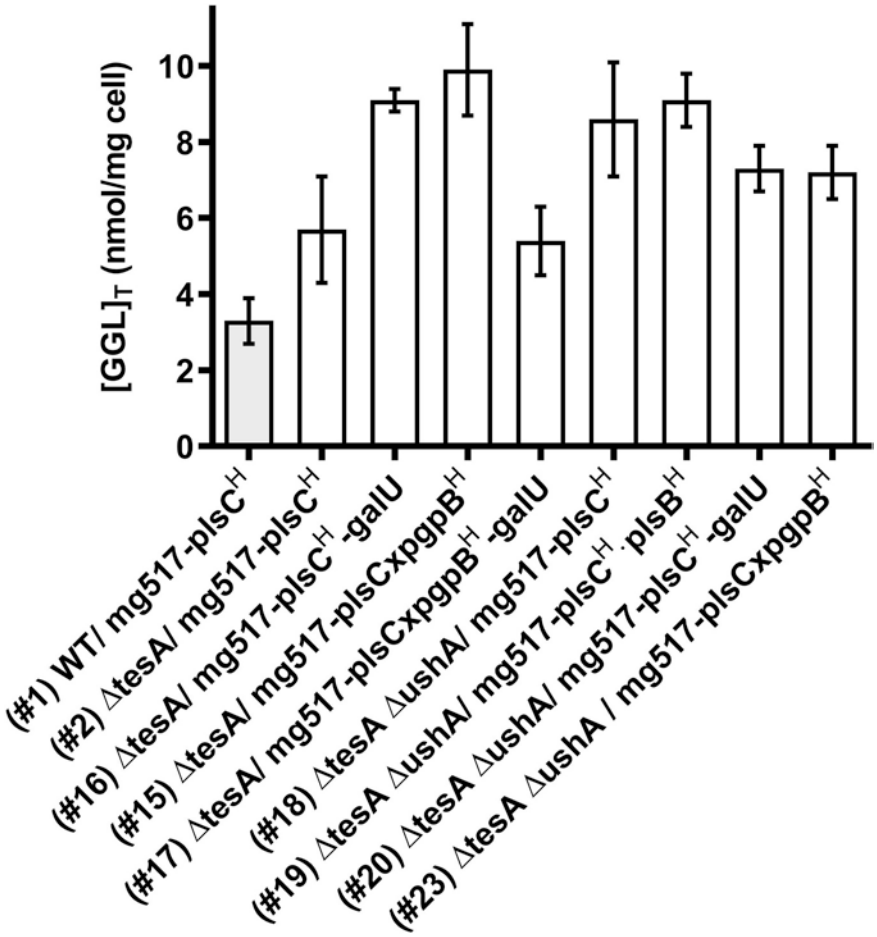

Fig. 6. Effect of GalU overexpression and *ush*A knockout on glycoglycerolipids (GGL) production in the Δ*tes*A strains.

3.4. Fatty acid profile

Since PA present in *E. coli* cells has a diverse acyl chain composition (Parsons and Rock, 2013), the new GGL were diverse not only in the glycosyl units but also in the acyl groups. It was shown in the first generation of engineered strains expressing MG517 and PlsC (Mora-Buyé et al., 2012) that the fatty acid composition of the GGL products was similar to the total lipid fraction, mainly constituted of palmitic (C16:0), palmitoleic (C16:1), 2-hexylcyclopropanoctanoic (C17:Δ) and oleic (C18:1) acids. It means that the glycosyltransferase had no preference for the acyl chains of its DAG acceptor. Here, the fatty acid profile of total lipid extracts from the different knockout strains expressing MG517 and PlsC was analyzed (Fig. 8). Δ*tes*A, Δ*fad*E, Δ*tes*A Δ*fad*E and Δ*tes*A Δ*ush*A strains (#2, #3, #4 and #18) had a similar lipidic profile compared with the WT strain and their lipid membrane is composed of about 45% palmitic acid (C16:0) and up to 10% of palmitoleic (C16:1), 2-hexylcyclopropanoctanoic (C17:Δ) and oleic (C18:1) acids. By contrast, the Δ*tes*A Δ*fab*R strain (#5) synthesized higher amounts of unsaturated fatty acids, mainly C16:1 and C18:1 and less palmitic acid, which only represented 29% of the total fatty acids.

FabR is a transcriptional repressor of *fab*A and *fab*B genes, which are responsible for the synthesis of unsaturated fatty acids. Previous studies showed that FabR mutants increased up to 2-fold the expression of the FabA and FabB proteins with the concomitant increase of unsaturated fatty acids (Parsons and Rock, 2013; Zhang et al., 2002). Our results show that knocking out *fab*R (Δ*tes*A Δ*fab*R strain) did not yield higher GGL titer but resulted in higher proportion of unsaturated acyl chains, making this strain of interest when a different lipidic profile is desired.

FadR overexpression would have a similar effect than the *fab*R knockout in the fatty acids profile, since it is also involved in the degradation of unsaturated fatty acids (Bacik et al., 2015; Fujita et al., 2007; Kim et al., 2018). However, the Δ*tes*A strain co-expressing FadR with MG517 and PlsC (#6) showed only a slight increase in the production of unsaturated fatty acids. This might be due to low protein expression level from a low copy number plasmid, but it was no further studied since FadR had not an impact on GGL production.

3.5. Glycolipids profile

All strains produced a mixture of glycosylated DAG products as the result of the possessive (or sequentially acting) activity of MG517. Mono-, di- and tri-glycosyldiacylglycerols (MGDAG, DGDAG and TriGDAG) were previously identified (Andrés et al., 2011; Mora-Buyé et al., 2012) but another non-identified product, tentatively assigned to TetraGDAG, was also observed. Fig. 9A summarizes the %GGL distribution for all the strains here studied. DGDAG was the major product in most of the strains with few exceptions (strains #16 and #20).

For the Δ*tes*A and Δ*fad*E strains expressing the acyltransferases (strains #2 to #13), there seems to be a correlation between total GGL production and %DGDAG (Fig. 9B). Likewise, a similar correlation is also observed for the strains containing the *ush*A knockout (strains #17 to #22). Strains #14, #15 and #16 deviate from the general trend.

Strain #14, co-expressing the hydrolase Cdh together with the acyltransferases, is a medium GGL producer but reaches the highest % DGDAG value up to 86% of total GGL. On the other hand, strain #15 is the highest GGL producer but %DGDAG is in the mid-range in favor to the larger GGL products, tri- and tetraGDAG. Even more apparent is strain #16, also one of the highest GGL producers

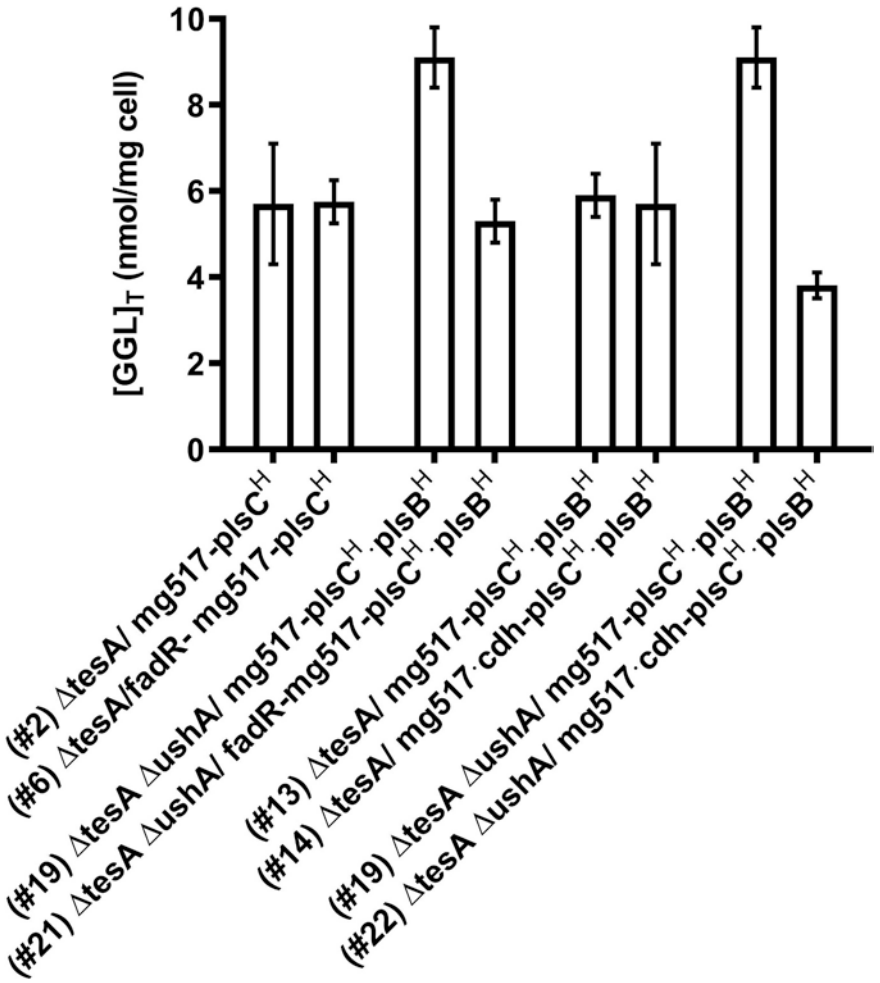

Fig. 7. Effect of FadR and Cdh overexpression on glycoglycerolipids (GGL) production in Δ*tes*A and Δ*tes*A Δ*ush*A strains.

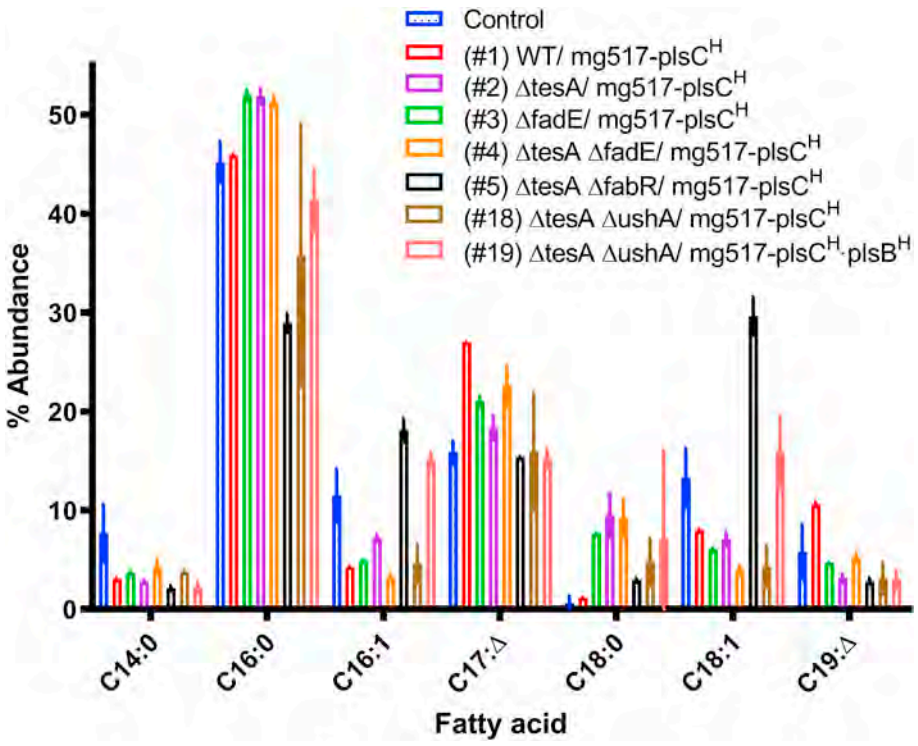

Fig. 8. Relative percentage of fatty acids from lipid extracts of engineered strains with *tes*A, *fad*E, *fab*R, and *ush*A knockouts. Except control (*E. coli* BL21 Star (DE3) cells), engineered strains co-express MG517 and PlsC. Fatty acids: myristic acid C14:0, palmitic acid C16:0, palmitoleic acid C16:1, 2-hexylcyclopropanoctanoic acid, C17:Δ, stearic acid C18:0, oleic acid C18:1 and 2-octylcyclopropanoctanoic acid, C19:Δ.

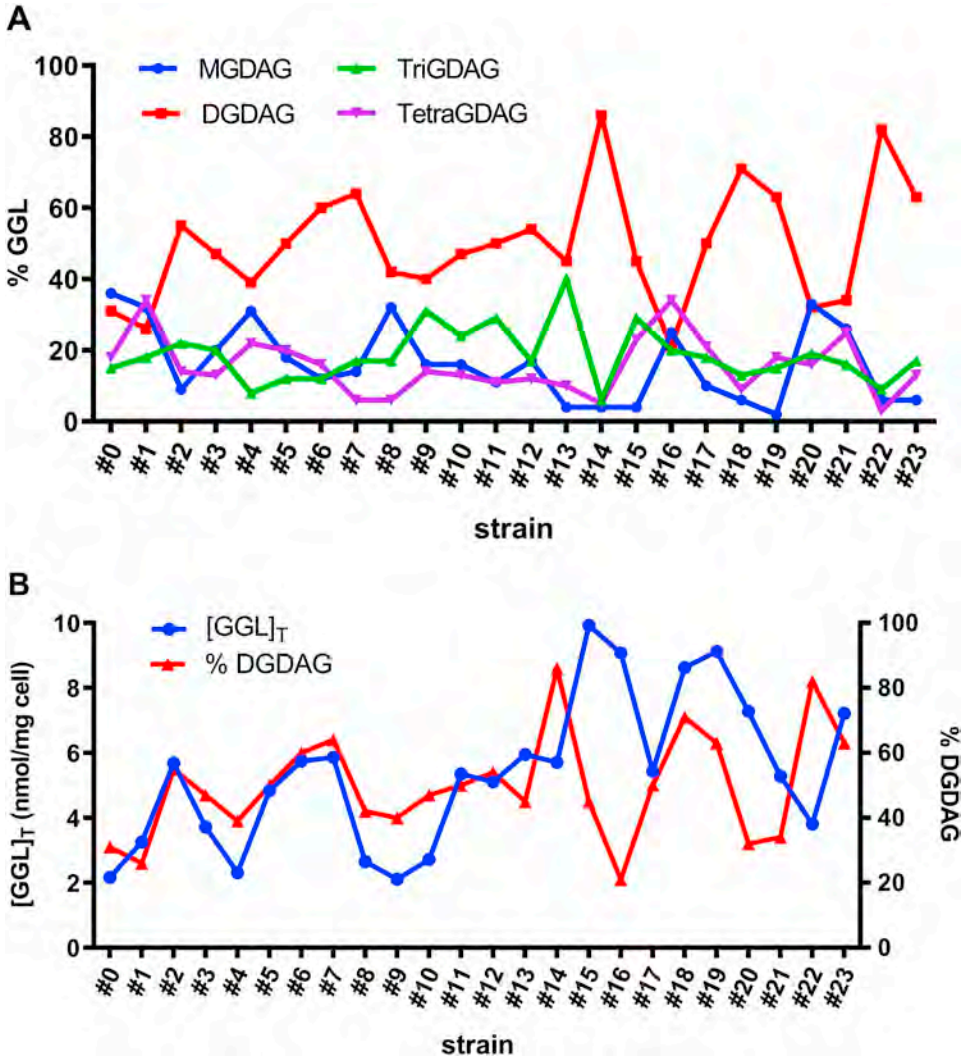

Fig. 9. A. Percentage glycoglycerolipids (%GGL) distribution among all strains studied in this work. **B.** Correlation between total GGL production and %DGDAG (diglycosyldiacylglycerol) for all strains. Data from Table 2.

([GGL]$_T$ > 9 nmol·mg^{-1} of cells), where the %tetraGDAG is increased to its highest value while %DGDAG is the lowest among all strains.

The general trend for most of the strains (the larger the GGL production the higher the %DGDAG) relates with the possessive behavior of the glycosyltransferase MG517, where DGDAG in the main product in *in vitro* enzyme studies (Andrés et al., 2011). The three outliers (strains #14, #15 and #16) differ in PgpB expression or GalU expression under a Δ*tes*A background or lack of PlsB overexpression under a Δ*ush*A background. It suggests a change in the regulation of the processive behavior of MG517 due to a different but yet unknown environment. A deeper understanding of the regulation mechanism of this processive glycolipid synthase is required to rationalize the unexpected results here observed.

To date, it is still unclear which GGL composition is the best as bioactive compounds for biomedical applications. It is known that monogalactosylglycerolipid and derivatives present antitumoral activities (Akasaka et al., 2016; Colombo et al., 2011; Maeda et al., 2013) (and that the GGL from *Ficus microcarpa* (mainly mono- and digalactosylglycerolipid) and from *Meithermus taiwanensis* (tetrasaccharide derivative) show inhibitory effect on TNF-α-induced IL-8 secretion and cytokine production of monocytes respectively (Ghosh et al., 2013; Kiem et al., 2012). Also, GGL are analogs of mono- and diglycosylceramides which are considered interesting marine bioactive compounds because of their immunostimulators properties (Blunt et al., 2018; Costantino et al., 2005; Rocha-Martin et al., 2014). The different GGL composition of these engineered strains could be assayed as vaccine adjuvants.

3.6. Membrane lipids composition

The main lipids in the *E. coli* membrane are phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL). PE (neutral zwitterionic lipid) can reach up to 75% of the total phospholipids depending on the growth conditions, whereas PG and CL (anionic lipids) account for 20% and 5% of the phospholipids composition (Wikstrom et al., 2009, 2004). GGL are neutral polar lipids that perturb the membrane properties. We previously reported for the 1st generation strains expressing MG517 and PlsC (*i.e.* strains #0 and #1 in Table 2) that GGL replaced PE, which was reduced between 10-20% relative to the control (no GGL) strain to compensate the membrane charge (Mora-Buyé et al., 2012). Here we analyzed the membrane lipid composition of the engineered strains to evaluate the effect of GGL overproduction. Cells were grown in the presence of ^{14}C-acetate and the lipid extract was analyzed by 2D TLC and autoradiography. Fig. 10 shows that the WT strain (#1) contained around 31% GGL, whereas the engineered strains expressing MG517 and PlsC, Δ*tes*A, Δ*tes*A Δ*ush*A or Δ*tes*A in combination with PgpB reached GGL levels between 40 and 50%. Compared to a control *E. coli* strain (untransformed BL21 Star (DE3)

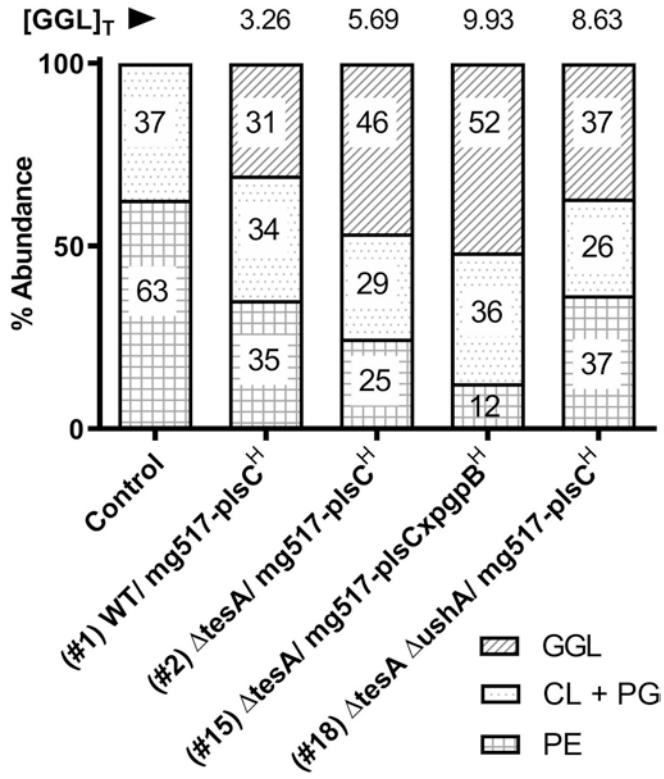

Fig. 10. Membrane lipid composition of the engineered strains. The lipid extract of cells grown in the presence of ^{14}C-acetate was analyzed by 2D TLC and autoradiography. GGL: glycoglycerolipids, PE: phosphatidylethanolamine, PG: phosphatidylglycerol, CL: cardiolipin. Numbers inside columns: % of each lipid. Numbers on top of each column: $[GGL]_T$ in nmol/mg dry cells.

cells), the amount of PE declined from 63 to 12% in the best producer strain, ΔtesA/mg517-pgpBxplsC^H (#15) (Fig. 10). Therefore, GGL is replacing PE for membrane charge compensation. However, the DGDAG, the major GGL component, is a bilayer-forming lipid that confers rigidity to the membrane, whereas PE is a non-bilayer forming lipid, and their exchange will affect membrane properties. Optical microscopy showed that the engineered cells are larger, elongated in a rod-like shape, where septation is hampered (Fig. S3, *Supporting Information*).

3.7. MG517 glycosyltransferase activity

The MG517 glycosyltransferase activity in the cell extracts was determined using UDPGal and the fluorogenic C6-NBD-ceramide as substrates, since the enzyme has shown to accept both DAG and ceramide as acceptor substrates, allowing an easier determination of the activity by HPLC and fluorescence detection (unpublished). Since MG517 is expressed from the same plasmid in all the strains, it was initially assumed that the expression levels and the enzyme activities would be similar for all strains. Surprisingly, the engineered strains presented large differences in enzyme activity (Fig. 11A). As example, the best GGL producer strain (ΔtesA/mg517-pgpBxplsC^H (#15)) and the strain without PgpB co-expression that gave almost half GGL production (#2), both presented the highest specific activity values. On the other hand, strains with high GGL production such as #15, #18, #23 showed variable specific activities. To investigate whether these significant variations of enzyme activity were due to different protein expressions levels or changes in the specific activity depending on the enzyme regulation by the variable lipidic environment, MG517 was expressed as a fusion protein with the fluorescent mCherry protein in some representative engineered strains. As shown in Fig. 11B, both protein expression and specific activity follow the same trend but again significantly varied among strains. Although the variability of MG517 expression levels needs to be studied further, it has no significant effect on GGL productivity and MG517 activity is not limiting in the high GGL producer strains.

4. Conclusions

Among the different strategies studied to enhance the availability of

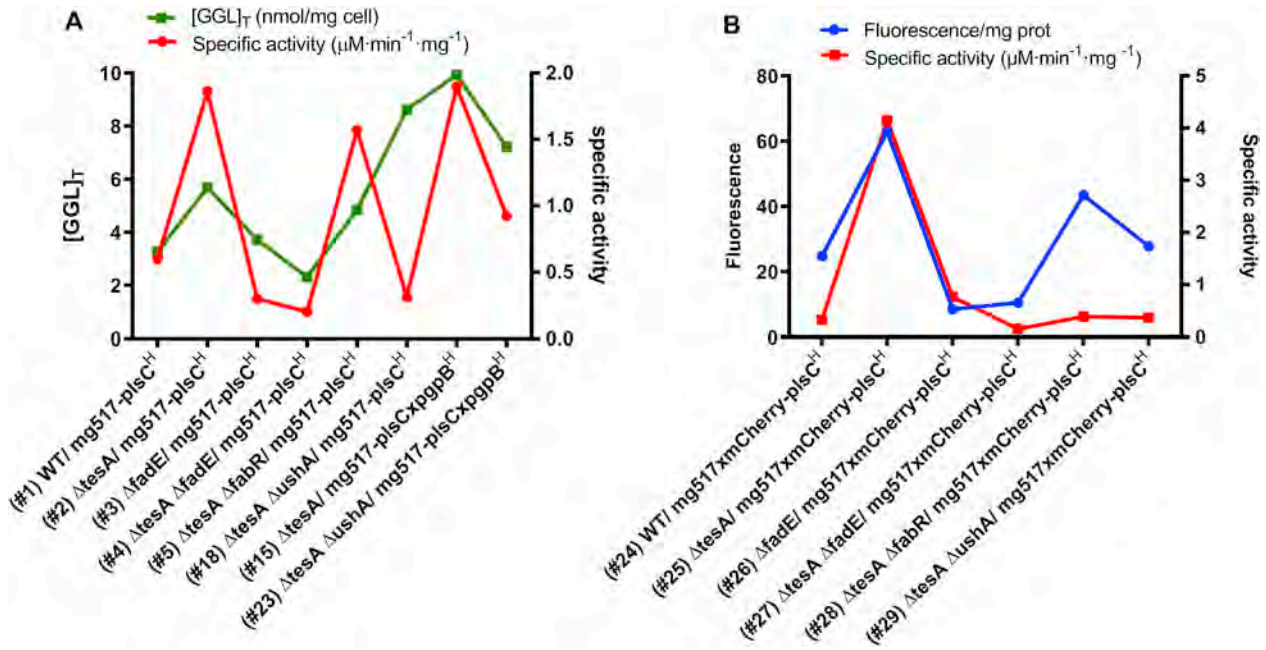

Fig. 11. A. GT MG517 specific activity and total glycoglycerolipids (GGL) production for selected strains. Specific activity of the cell-free extracts expressed as μM product per min and mg of total protein using UDP-Galactose (UDPGal) and ceramide-NBD ((N-[6-[(7-nitro-2-1,3-benzoxadiazol-4-yl)amino]hexanoyl]-D-erythro-sphingosine) as substrates. **B.** Expression and specific activity of MG517 as fusion protein with mCherry for selected strains. Protein expression determined by mCherry fluorescence of the cell-free extract and expressed as fluorescence intensity per mg of total protein.

diacylglycerol and UDPGlc as precursor substrates of the MG517 glycosyltransferase for the production of GGL, the thioesterase *tes*A gene knockout and the co-expression of acyltransferases have been key issues for improving GGL titers. Combining the *tes*A knockout strain expressing MG517 and the acyltransferase PlsC with co-expression of the phosphatidic acid phosphatase PgpB as fusion protein with PlsC, or with the uridyltransferase GalU overexpression or the diphosphatase *ush*A gene knockout, resulted in the best GGL producer strains with a 5-fold higher production yield than the parental strain expressing the MG517 glycosyltransferase alone. Most of the strains produced a mixture composed of di-, tri- and tetra-glycosyldiacylglycerols. Interestingly, the strains co-expressing the pyrophosphatase Cdh modified the GGL composition, shifting the percentage of diglucosyldiacylglycerol up to 80% of the total GGL. Although it is not known which GGL composition is best for eliciting antitumor or immunostimulator activities in biomedical applications, the different GGL profiles from these engineered strains could be assayed as potential candidates. Likewise, strains with the *fad*R gene knockout changed the acyl composition to higher proportion of unsaturated acyl chains in the GGL products. As a consequence of producing GGL, the lipidic composition of the *E. coli* membrane was modified, observing that GGL products replaced the phosphatidylethanolamine (both neutral polar lipids) to maintain the overall membrane charge.

In conclusion, exploring the different strategies has led us to a novel metabolic engineering *E. coli* platform for GGL production. The replacement of the here used MG517 glycolipid synthase by other glycosyltransferases will enable the access to other glycosylated diacylglycerol families with different sugar compositions. In addition, upstream and downstream conditions such as studies with different carbon sources and glycolipid extraction methods, respectively, could be explored to increase GGL yield.

Author statement

Nuria Orive-Milla: Methodology, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing- Original draft preparation. **Tom Delmulle**: Methodology, Validation. **Marjan de Mey**: Supervision, Resources, Writing- Reviewing and Editing. **Magda Faijes**: Conceptualization, Methodology, Validation, Data curation, Writing-Original draft preparation: **Antoni Planas**: Conceptualization, Methodology, Resources, Writing- Reviewing and Editing, Visualization, Supervision, Project Administration, Funding acquisition

Acknowledgements

Work supported by grant BFU2016-77427-C2-1-R from MINECO, Spain., and AGAUR grant 2017SGR-727 from the Generalitat de Catalunya (to A. P.). We acknowledge Carlos Santiveri and Eulalia Segur for their contribution in preparing knockout strains, and Mireia Solà and Marta Ballester for lipid extractions and quantifications. We also acknowledge IQS School of Engineering for the predoctoral fellowship (to N.O).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymben.2020.05.010.

References

Abdel-Mawgoud, A.M., Stephanopoulos, G., 2018. Simple glycolipids of microbes: chemistry, biological activity and metabolic engineering. Synth. Syst. Biotechnol. 3, 3–19. https://doi.org/10.1016/j.synbio.2017.12.001.

Ajikumar, P.K., Xiao, W.-H., Tyo, K.E.J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T.H., Pfeifer, B., Stephanopoulos, G., 2010. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. Science (80-.) 330, 70–74. https://doi.org/10.1126/science.1191652.

Akasaka, H., Mizushina, Y., Yoshida, K., Ejima, Y., Mukumoto, N., Wang, T., Inubushi, S., Nakayama, M., Wakahara, Y., Sasaki, R., 2016. MGDG extracted from spinach enhances the cytotoxicity of radiation in pancreatic cancer cells. Radiat. Oncol. 11, 153. https://doi.org/10.1186/s13014-016-0729-0.

Akasaka, H., Sasaki, R., Yoshida, K., Takayama, I., Yamaguchi, T., Yoshida, H., Mizushina, Y., 2013. Monogalactosyl diacylglycerol, a replicative DNA polymerase inhibitor, from spinach enhances the anti-cell proliferation effect of gemcitabine in human pancreatic cancer cells. Biochim. Biophys. Acta 1830, 2517–2525. https://doi.org/10.1016/j.bbagen.2012.11.004.

Andrés, E., Biarnés, X., Faijes, M., Planas, A., 2012. Bacterial glycoglycerolipid synthases: processive and non-processive glycosyltransferases in mycoplasma. Biocatal. Biotransform. 30, 274–287. https://doi.org/10.3109/10242422.2012.674733.

Andrés, E., Martínez, N., Planas, A., 2011. Expression and characterization of a Mycoplasma genitalium glycosyltransferase in membrane glycolipid biosynthesis: potential target against Mycoplasma infections. J. Biol. Chem. 286, 35367–35379. https://doi.org/10.1074/jbc.M110.214148.

Bacik, J.-P., Yeager, C.M., Twary, S.N., Marti-Arbona, R., 2015. Modulation of FadR binding capacity for acyl-CoA fatty acids through structure-guided mutagenesis. Protein J. 34, 359–366. https://doi.org/10.1007/s10930-015-9630-1.

Bahia, F.M., De Almeida, G.C., De Andrade, L.P., Campos, C.G., Queiroz, L.R., Da Silva, R.L.V., Abdelnur, P.V., Corrêa, J.R., Bettiga, M., Parachin, N.S., 2018. Rhamnolipids production from sucrose by engineered Saccharomyces cerevisiae. Sci. Rep. 8, 1–10. https://doi.org/10.1038/s41598-018-21230-2.

Bailey, R.W., 1958. The reaction of pentoses with anthrone. Biochem. J. 68, 669–672.

Banat, I.M., Makkar, R.S., Cameotra, S.S., 2000. Potential commercial applications of microbial surfactants. Appl. Microbiol. Biotechnol. 53, 495–508.

Behren, S., Westerlind, U., 2019. Glycopeptides and -mimetics to detect, monitor and inhibit bacterial and viral infections: recent advances and perspectives. Molecules 24, 1004. https://doi.org/10.3390/molecules24061004.

Bentley, G.J., Jiang, W., Guamán, L.P., Xiao, Y., Zhang, F., 2016. Engineering Escherichia coli to produce branched-chain fatty acids in high percentages. Metab. Eng. 38, 148–158. https://doi.org/10.1016/j.ymben.2016.07.003.

Blunt, J.W., Carroll, A.R., Copp, B.R., Davis, R.A., Keyzers, R.A., Prinsep, M.R., 2018. Marine natural products. Nat. Prod. Rep. 35, 8–53. https://doi.org/10.1039/c7np00052a.

Carreño, L.J., Kharkwal, S.S., Porcelli, S. a, 2014. Optimizing NKT cell ligands as vaccine adjuvants. Immunotherapy 6, 309–320. https://doi.org/10.2217/imt.13.175.

Celińska, E., Borkowska, M., Białas, W., Kubiak, M., Korpys, P., Archacka, M., Ledesma-Amaro, R., Nicaud, J.-M., 2019. Genetic engineering of Ehrlich pathway modulates production of higher alcohols in engineered Yarrowia lipolytica. FEMS Yeast Res. 19. https://doi.org/10.1093/femsyr/foy122.

Chen, N., Yu, Z.-H., Zhou, D., Hu, X.-L., Zang, Y., He, X.-P., Li, J., Xie, J., 2016. N-Oxyamide-linked glycoglycerolipid coated AuNPs for receptor-targeting imaging and drug delivery. Chem. Commun. (Camb). 52, 2284–2287. https://doi.org/10.1039/c5cc09749e.

Cho, H., Cronan, J.E., 1995. Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. J. Biol. Chem. https://doi.org/10.1074/jbc.270.9.4216.

Chong, H., Li, Q., 2017. Microbial production of rhamnolipids: opportunities, challenges and strategies. Microb. Cell Factories 16, 1–12. https://doi.org/10.1186/s12934-017-0753-2.

Colombo, D., Tringali, C., Franchini, L., Cirillo, F., Venerando, B., 2011. Glycoglycerolipid analogues inhibit PKC translocation to the plasma membrane and downstream signaling pathways in PMA-treated fibroblasts and human glioblastoma cells. U87MG. Eur. J. Med. Chem. 46, 1827–1834. https://doi.org/10.1016/j.ejmech.2011.02.043.

Corti, M., Cantù, L., Brocca, P., Del Favero, E., 2007. Self-assembly in glycolipids. Curr. Opin. Colloid Interface Sci. 12, 148–154. https://doi.org/10.1016/j.cocis.2007.05.002.

Costantino, V., D'Esposito, M., Fattorusso, E., Mangoni, A., Basilico, N., Parapini, S., Taramelli, D., 2005. Damicoside from Axinella damicornis: the influence of a glycosylated galactose 4-OH group on the immunostimulatory activity of alpha-galactoglycosphingolipids. J. Med. Chem. 48, 7411–7417. https://doi.org/10.1021/jm050506y.

Czerwiec, Q., Idrissitaghki, A., Imatoukene, N., Nonus, M., Thomasset, B., Nicaud, J.-M., Rossignol, T., 2019. Optimization of cyclopropane fatty acids production in Yarrowia lipolytica. Yeast 36, 143–151. https://doi.org/10.1002/yea.3379.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U.S.A. 97, 6640–6645. https://doi.org/10.1073/pnas.120163297.

De Bruyn, F., De Paepe, B., Maertens, J., Beauprez, J., De Cocker, P., Mincke, S., Stevens, C., De Mey, M., 2015. Development of an in vivo glucosylation platform by coupling production to growth: production of phenolic glucosides by a glycosyltransferase of Vitis vinifera. Biotechnol. Bioeng. 112, 1594–1603. https://doi.org/10.1002/bit.25570.

Dillon, D.A., Wu, W.I., Riedel, B., Wissing, J.B., Dowhan, W., Carman, G.M., 1996. The Escherichia coli pgpB gene encodes for a diacylglycerol pyrophosphate phosphatase activity. J. Biol. Chem. 271, 30548–30553. https://doi.org/10.1074/JBC.271.48.30548.

Dong, H., Zhao, C., Zhang, T., Zhu, H., Lin, Z., Tao, W., Zhang, Y., Li, Y., 2017. A systematically chromosomally engineered Escherichia coli efficiently produces butanol. Metab. Eng. 44, 284–292. https://doi.org/10.1016/J.YMBEN.2017.10.014.

Du, W., Kulkarni, S.S., Gervay-Hague, J., 2007. Efficient, one-pot syntheses of biologically active α-linked glycolipids. Chem. Commun. 2336. https://doi.org/10.1039/b702551c.

Faivre, V., Rosilio, V., 2010. Interest of glycolipids in drug delivery: from physicochemical properties to drug targeting. Expet Opin. Drug Deliv. 7, 1031–1048. https://doi.org/10.1517/17425247.2010.511172.

Folch, J., Lees, M., Sloane Stanley, G.H., 1957. A simple method for the isolation and purification of total lipides from animal tissues. J. Biol. Chem. 226, 497–509. https://doi.org/10.1371/journal.pone.0020510.

Fujita, Y., Matsuoka, H., Hirooka, K., 2007. Regulation of fatty acid metabolism in bacteria. Mol. Microbiol. 66, 829–839. https://doi.org/10.1111/j.1365-2958.2007.05947.x.

Ghosh, B., Lai, Y.-H., Shih, Y.-Y., Pradhan, T.K., Lin, C.-H., Mong, K.-K.T., 2013. Total synthesis of a glycoglycerolipid from Meiothermus taiwanensis through a one-pot glycosylation reaction and exploration of its immunological properties. Chem. Asian J. 8, 3191–3199. https://doi.org/10.1002/asia.201300933.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C. a, Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345. https://doi.org/10.1038/nmeth.1318.

He, H., Lu, Y., Qi, J., Zhu, Q., Chen, Z., Wu, W., 2019. Adapting liposomes for oral drug delivery. Acta Pharm. Sin. B 9, 36–48. https://doi.org/10.1016/j.apsb.2018.06.005.

He, L., Xiao, Y., Gebreselassie, N., Zhang, F., Antoniewicz, M.R., Tang, Y.J., Peng, L., 2014. Central metabolic responses to the overproduction of fatty acids in Escherichia coli based on 13C-metabolic flux analysis. Biotechnol. Bioeng. 111, 575–585. https://doi.org/10.1002/bit.25124.

Holst, O., 2008. Glycolipids: occurrence, significance, and properties. In: Fraser-Reid, B.O., Tatsuta, K., Thiem, J. (Eds.), Glycoscience: Chemistry and Chemical Biology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1603–1627. https://doi.org/10.1007/978-3-540-30429-6_39.

Hölzl, G., Dörmann, P., 2007. Structure and function of glycoglycerolipids in plants and bacteria. Prog. Lipid Res. 46, 225–243. https://doi.org/10.1016/j.plipres.2007.05.001.

Janßen, H.J., Steinbüchel, A., 2014. Fatty acid synthesis in Escherichia coli and its applications towards the production of fatty acid based biofuels. Biotechnol. Biofuels 7, 7. https://doi.org/10.1186/1754-6834-7-7.

Jia, Y., Akache, B., Deschatelets, L., Qian, H., Dudani, R., Harrison, B.A., Stark, F.C., Chandan, V., Jamshidi, M.P., Krishnan, L., McCluskie, M.J., 2019. A comparison of the immune responses induced by antigens in three different archaeosome-based vaccine formulations. Int. J. Pharm. 561, 187–196. https://doi.org/10.1016/j.ijpharm.2019.02.041.

Kalisch, B., Dormann, P., Holzl, G., 2016. DGDG and Glycolipids in Plants and Algae., Subcellular Biochemistry. United Stateshttps://doi.org/10.1007/978-3-319-25979-6_3.

Karp, P.D., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., Krummenacker, M., Subhraveti, P., Weerasinghe, D., Gama-Castro, S., Huerta, A.M., Muñiz-Rascado, L., Bonavides-Martinez, C., Weiss, V., Peralta-Gil, M., Santos-Zavaleta, A., Schröder, I., Mackie, A., Gunsalus, R., Collado-Vides, J., Keseler, I.M., Paulsen, I., 2014. The EcoCyc database. EcoSal Plus 6. https://doi.org/10.1128/ecosalplus.ESP-0009-2013.

Kavoosi, M., Creagh, A.L., Kilburn, D.G., Haynes, C.A., 2007. Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in Escherichia coli. Biotechnol. Bioeng. 98, 599–610. https://doi.org/10.1002/bit.21396.

Kiem, P. Van, Minh, C. Van, Nhiem, N.X., Cuong, N.X., Tai, B.H., Quang, T.H., Anh, H.L.T., Yen, P.H., Ban, N.K., Kim, S.H., Xin, M., Cha, J.-Y., Lee, Y.-M., Kim, Y.H., 2012. Inhibitory effect on TNF-alpha-induced IL-8 secretion in HT-29 cell line by glyceroglycolipids from the leaves of Ficus microcarpa. Arch Pharm. Res. (Seoul) 35, 2135–2142. https://doi.org/10.1007/s12272-012-1210-8.

Kim, J., Yoo, H.-W., Kim, M., Kim, E.-J., Sung, C., Lee, P.-G., Park, B.G., Kim, B.-G., 2018. Rewiring FadR regulon for the selective production of omega-hydroxy palmitic acid from glucose in Escherichia coli. Metab. Eng. 47, 414–422. https://doi.org/10.1016/j.ymben.2018.04.021.

Kopitz, J., 2017. Lipid glycosylation: a primer for histochemists and cell biologists. Histochem. Cell Biol. 147, 175–198. https://doi.org/10.1007/s00418-016-1518-4.

Lee, Y.-L., Chen, J.C., Shaw, J.-F., 1997. The thioesterase I of Escherichia coli has arylesterase activity and shows stereospecificity for protease substrates. Biochem. Biophys. Res. Commun. 231, 452–456. https://doi.org/10.1006/bbrc.1997.5797.

Leyva, A., Quintana, A., Sánchez, M., Rodríguez, E.N., Cremata, J., Sánchez, J.C., 2008. Rapid and sensitive anthrone-sulfuric acid assay in microplate format to quantify carbohydrate in biopharmaceutical products: method development and validation. Biologicals 36, 134–141. https://doi.org/10.1016/j.biologicals.2007.09.001.

Liu, Z., Guo, J., 2017. NKT-cell glycolipid agonist as adjuvant in synthetic vaccine. Carbohydr. Res. 452, 78–90. https://doi.org/10.1016/j.carres.2017.10.006.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495. https://doi.org/10.1093/nar/gkt1178.

Lu, X., Vora, H., Khosla, C., 2008. Overproduction of free fatty acids in E. coli : implications for biodiesel production. Metab. Eng. 10, 333–339. https://doi.org/10.1016/j.ymben.2008.08.006.

Maeda, N., Kokai, Y., Hada, T., Yoshida, H., Mizushina, Y., 2013. Oral administration of monogalactosyl diacylglycerol from spinach inhibits colon tumor growth in mice. Exp. Ther. Med. 5, 17–22. https://doi.org/10.3892/etm.2012.792.

Malhotra, R., 2012. Membrane glycolipids: functional heterogeneity: a review. Biochem. Anal. Biochem. 1, 1–5. https://doi.org/10.4172/2161-1009.1000108.

Marrakchi, H., Zhang, Y.-M., Rock, C.O., 2002. Mechanistic diversity and regulation of Type II fatty acid synthesis. Biochem. Soc. Trans. 30, 30–33. https://doi.org/10.1042/BST0301050.

Mora-Buyé, N., Faijes, M., Planas, A., 2012. An engineered E.coli strain for the production of glycoglycerolipids. Metab. Eng. 14, 551–559. https://doi.org/10.1016/j.ymben.2012.06.001.

Murakami, S., Shimamoto, T., Nagano, H., Tsuruno, M., Okuhara, H., Hatanaka, H., Tojo, H., Kodama, Y., Funato, K., 2015. Producing human ceramide-NS by metabolic engineering using yeast Saccharomyces cerevisiae. Sci. Rep. 5, 16319. https://doi.org/10.1038/srep16319.

Muthusamy, K., Gopalakrishnan, S., Ravi, T.K., Sivachidambaram, P., 2008. Biosurfactants: properties, commercial production and application. Curr. Sci. 94, 736–747.

Pandey, R.P., Parajuli, P., Koirala, N., Lee, J.H., Park, Y. Il, Sohng, J.K., 2014. Glucosylation of isoflavonoids in engineered Escherichia coli. Mol. Cell 37, 172–177. https://doi.org/10.14348/molcells.2014.2348.

Park, Y.-K., Dulermo, T., Ledesma-Amaro, R., Nicaud, J.-M., 2018. Optimization of odd chain fatty acid production by Yarrowia lipolytica. Biotechnol. Biofuels 11, 158. https://doi.org/10.1186/s13068-018-1154-4.

Parsons, J.B., Rock, C.O., 2013. Bacterial lipids: metabolism and membrane homeostasis. Prog. Lipid Res. 52, 249–276. https://doi.org/10.1016/j.plipres.2013.02.002.

Quan, J., Tian, J., 2014. Circular polymerase extension cloning. In: Clifton, N.J. (Ed.), Methods in Molecular Biology, pp. 103–117. https://doi.org/10.1007/978-1-62703-764-8_8.

Rakhuba, D., Novik, G., Dey, E.S., 2009. Application of supercritical carbon dioxide (scCO2) for the extraction of glycolipids from Lactobacillus plantarum B-01. J. Supercrit. Fluids 49, 49–51.

Rigouin, C., Lajus, S., Ocando, C., Borsenberger, V., Nicaud, J.M., Marty, A., Averous, L., Bordes, F., 2019. Production and characterization of two medium-chain-length polydroxyalkanoates by engineered strains of Yarrowia lipolytica. Microb. Cell Factories 18, 99. https://doi.org/10.1186/s12934-019-1140-y.

Rocha-Martin, J., Harrington, C., Dobson, A.D.W., O'Gara, F., 2014. Emerging strategies and integrated systems microbiology technologies for biodiscovery of marine bioactive compounds. Mar. Drugs 12, 3516–3559. https://doi.org/10.3390/md12063516.

Saini, M., Lin, L.-J., Chiang, C.-J., Chao, Y.-P., 2017. Synthetic consortium of Escherichia coli for n -butanol production by fermentation of the glucose–xylose mixture. J. Agric. Food Chem. 65, 10040–10047. https://doi.org/10.1021/acs.jafc.7b04275.

San, K.-Y., Li, M., 2017. Bacteria and method for synthetizing fatty acids. US9598696B2. https://doi.org/10.1145/634067.634234.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9 (7), 671–675. https://doi.org/10.1038/nmeth.2089.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G., Palmer, A.E., Tsien, R.Y., 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. 22, 1567–1572. https://doi.org/10.1038/nbt1037.

Srirangan, K., Liu, X., Akawi, L., Bruder, M., Moo-Young, M., Chou, C.P., 2016. Engineering Escherichia coli for microbial production of butanone. Appl. Environ. Microbiol. 82, 2574–2584. https://doi.org/10.1128/AEM.03964-15.

Steen, E.J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., del Cardayre, S.B., Keasling, J.D., 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature 463, 559–562. https://doi.org/10.1038/nature08721.

Wang, X., Zou, X., Yin, J., Qiu, L., Lu, M., Hu, J., 2015. Carbohydrate-based vaccine adjuvants – discovery and development. Expet Opin. Drug Discov. 10, 1133–1144. https://doi.org/10.1517/17460441.2015.1067198.

Wicke, C., Huners, M., Wray, V., Nimtz, M., Bilitewski, U., Lang, S., 2000. Production and structure elucidation of glycoglycerolipids from a marine sponge-associated microbacterium species. J. Nat. Prod. 63, 621–626. https://doi.org/10.1021/np990313b.

Wikstrom, M., Kelly, A.A., Georgiev, A., Eriksson, H.M., Klement, M.R., Bogdanov, M., Dowhan, W., Wieslander, A., 2009. Lipid-engineered Escherichia coli membranes reveal critical lipid headgroup size for protein function. J. Biol. Chem. 284, 954–965. https://doi.org/10.1074/jbc.M804482200.

Wikstrom, M., Xie, J., Bogdanov, M., Mileykovskaya, E., Heacock, P., Wieslander, A., Dowhan, W., 2004. Monoglucosyldiacylglycerol, a foreign lipid, can substitute for phosphatidylethanolamine in essential membrane-associated functions in Escherichia coli. J. Biol. Chem. 279, 10484–10493. https://doi.org/10.1074/jbc.M310183200.

Xie, J., Bogdanov, M., Heacock, P., Dowhan, W., 2006. Phosphatidylethanolamine and monoglucosyldiacylglycerol are interchangeable in supporting topogenesis and function of the polytopic membrane protein lactose permease. J. Biol. Chem. 281, 19172–19178. https://doi.org/10.1074/jbc.M602565200.

Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A.G.W., Collins, C.H., Koffas, M. a G., 2013. Modular optimization of multi-gene pathways for fatty acids production in E. coli. Nat. Commun. 4, 1409. https://doi.org/10.1038/ncomms2425.

Yao, J., Rock, C.O., 2013. Phosphatidic acid synthesis in bacteria. Biochim. Biophys. Acta 1831, 495–502. https://doi.org/10.1016/j.bbalip.2012.08.018.

Yee, L., Blanch, H.W., 1993. Defined media optimization for growth of recombinant Escherichia coli X90. Biotechnol. Bioeng. 41, 221–230.

Yemm, E.W., Willis, A.J., 1954. The estimation of carbohydrates in plant extracts by anthrone. Biochem. J. 57, 508–514.

Yunoki, K., Sato, M., Seki, K., Ohkubo, T., Tanaka, Y., Ohnishi, M., 2009. Simultaneous quantification of plant glyceroglycolipids including sulfoquinovosyldiacylglycerol by HPLC-ELSD with binary gradient elution. Lipids 44, 77–83. https://doi.org/10.1007/s11745-008-3248-4.

Zhang, F., Ouellet, M., Batth, T.S., Adams, P.D., Petzold, C.J., Keasling, J.D., Mukhopadhyay, A., 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. Metab. Eng. 14, 653–660. https://doi.org/10.1016/j.ymben.2012.08.009.

Zhang, Y.-M., Marrakchi, H., Rock, C.O., 2002. The FabR (YijC) transcription factor regulates unsaturated fatty acid biosynthesis in Escherichia coli. J. Biol. Chem. 277, 15558–15565. https://doi.org/10.1074/jbc.M201399200.

SUPPORTING INFORMATION

Metabolic engineering for glycoglycerolipids production in *E. coli*: tuning phosphatidic acid and UDP-glucose pathways

Nuria Orive[1], Tom Demulle[2], Marjan de Mey[2], Magda Faijes[1,*], Antoni Planas[1]

[1] Laboratory of Biochemistry, Institut Químic de Sarrià, University Ramon Llull, Via Augusta 350, E-08017 Barcelona, Spain. [2] Centre for Synthetic Biology (CSB), Ghent University, Coupure Links 653, B-9000 Ghent, Belgium

Table S1. Oligonucleotide primers for preparation of knockouts and expression plasmids.

Table S2. Glycosyltransferase GT MG517 specific activity and total glycoglycerolipids (GGL) production for selected strains.

Table S3. Percentage of fatty acids composition from lipids extracts of engineered strains.

Figure S1. Membrane lipids composition. 1D TLD and autoradiography of ^{14}C-labelled lipids.

Figure S2. Membrane lipids composition. 2D TLD and autoradiography of ^{14}C-labelled lipids.

Figure S3. Morphology of engineered cells by optical microscopy.

Table S1. Oligonucleotide primers for preparation of knockouts and expression plasmids.

Primer name	Sequence 5' --> 3'	Use
oMEMO4468	AGCGGATAAAGAAACGGAGCCTTTCGGCTCCGTTATTCATTTACGCGGCTTCAGTGTAGGCTGGAGCTGCTTC	FadE KO
oMEMO4480	TACATCCACTACAACCATATCATCACAAGTGGTCAGACCTCCTACAAGTAAGGCATATGAATATCCTCCTTAG	
oMEMO4464	CTTTGGCGAGTTTGGGGTAAATGGCGCTAAAGGCTTCATTATAACGGCGAGTGTAGGCTGGAGCTGCTTC	TesA KO
oMEMO4465	CGGCCTGGCCTGCCTTGTTGAATGATAAGTGGCAGAGTAAAACGTCGGTACATATGAATATCCTCCTTAG	
fabR_dats_fwC	AGAACCGGCCAAAGAATTGCAGTAAATATGTTTTATTGCGTTACCGTTCATTCACAATACTG-GAGCAATCGTGTAGGCTGGAGCTGCTTC	FabR KO
comp C	ATACTCTCCGTTTAAGCGGCAGGTTTCCGCTGTACGTAAAAGAACCGGCCAAAGAATTGCAGTAAATATG	
fabR_dats_rvD	ATTACTCGTCCTTCACATTTCCCGGAATAATTGCGGTTTTCTCTTGTTCACGGCGATACCAG-TAATAAGCCATATGAATATCCTCCTTAG	
comp_D	AGCAACTAACGCCAGCAGCAGCGTACCTCTATCTTGATTTGCTTGTTTCATTACTCGTCCTTCACATTTC	
ushA_dats_fwA	ACAGAATTTCTAATCTGGATGCAGATTTATCTTCACCGGACGCAGACTTGTC-TATGATGTCGCGTCATACGTGTAGGCTGGAGCTGCTTC	UshA KO
ushA_dats_rvB	AAATTTGCTGATATCGCCCGCCGCGATTAAGCATTGTGCCGGATGCAAACATCCGGCAC-TTTCGGATTACCATATGAATATCCTCCTTAG	
comp_A	ATTACCAGACTAACATACCTGTATGCGTCGTCTGAAGGAAGTCTCAACGCCGAATACAGAATTTCTAATC	
comp_B	TGGCGGCAGGCGATCTGGCAAAGATCCTCGATGCCAAATTTGCTGATATC	
oMEMO4560	CAATTCCCCTGTAGAAATAATTTATGCTATATATCTTTCGTCTTATTATTACC	p10T7-plsC
oMEMO4561	CTCAGCGGTGGCAGCAGCCTTTAAACTTTTCCGGCGGCTTC	
oMEMO4562	GAAGCCGCCGGAAAAGTTTAAAGGCTGCTGCCACCGCTGAG	
oMEMO4563	GGTAATAATAAGACGAAAGATATATAGCATAAATTATTTCTACAGGGGAATTG	
oMEMO4552	CAATTCCCCTGTAGAAATAATTTATGACTTTCTGCTATCCTTG	p10T7-plsB
oMEMO4553	GTTATTGCTCAGCGGTGGCATTACCCTTCGCCCTGCGTCGCACTC	
oMEMO4554	GAGTGCGACGCAGGGCGAAGGGTAATGCCACCGCTGAGCAATAAC	
oMEMO4555	CAAGGATAGCAGAAAGTCATAAATTATTTCTACAGGGGAATTG	
oMEMO4554	GAGTGCGACGCAGGGCGAAGGGTAATGCCACCGCTGAGCAATAAC	p10T7-plsC-plsB
oMEMO4684	CATATGTATATCTCCTTCTTGAATTCTTAAACTTTTCCGGCGGCTT	
oMEMO4553	GTTATTGCTCAGCGGTGGCATTACCCTTCGCCCTGCGTCGCACTC	
oMEMO4685	GAATTCAAGAAGGAGATATACATATGATGACTTTCTGCTATCCTTG	

Table S1 (cont). Oligonucleotide primers for preparation of knockouts and expression plasmids.

Primer name	Sequence 5' --> 3'	Use
fw_plsB_pRSFB	TATACCATGGCTTTCTGCTATCCTTG	pRSF1b-plsB
rv_plsB_pRSFB	CTTAGAATTCTTACCCTTCGCCCTGC	
fw_pRSF_pRSFB	ATGCGAATTCCCTAGGCTGCTGCCACCGCTGAG	
rv_pRSF_pRSFB	CATACCATGGTATATCTCCTTATTAAAGTTAAAC	
fw_vector_Gibson CB	GACGAAAGATATATAGCATAGTTAAACAAAATTATTTCTACA	pRSF1b-plsC-plsB
rv_vector_Gibson CB	GAGAGTGCGACGCAGGGCGAAGGGTAACTGCCACCGCTGAGCAATAAC	
fw_CB_Gibson	CCGCTGAGCAATAACATGCTATATATCTTTCGTCTTATTATT	
rv_CB_gibson	AAAATTATTTCTACAGGTTACCCTTCGCCCTGCGTCGCACTC	
fw_vector_pRSFCpgpB	GCGCAACGAGAACAAGAAAGTTAACTGCCACCGCTGAGCAATAAC	pRSF1b-plsC-pgpB
rv_vector_pRSFCpgpB	GGGTCGGGGTCGGCGTCGGCGTCGGGGTCGGAACTTTTCCGGCGGCTTCGCGTTC	
fw_pgpB_pRSFCpgpB	GACGCCGACCCCGACCCCGACGCCGACCCCGATGCGTTCGATTGCCAGACG	
rv_pgpB_pRSFCpgpB	GTTATTGCTCAGCGGTGGCAGTTAACTTTCTTGTTCTCGTTGCGCTATTTCG	
fw_mg517_mcherry	ATGCGGATCCCCGGCCTAGGCTG	pET44b-mg517-mCherry
rv_mg517_cherry	ATGCGAATTCGTTATCTGATTTAGATTCCAAAACATG	
fw_mcherry_mcherry	AAGGGAATTCACGCTAGCGCAACCGGAC	
rv_mcherry_mcherry	TTTTGGATCCCTACTTGTACAGCTCGTCCATGCC	
fw_pET44b_policistronic	GTGTGAGATTTTGCGTTAAACGAAAGGCTCAGTCGAAAG	pET44b-mg517-cdh
rv_pET44b_policistronic	GAAGACCCGCTTTTTTCATTAGTCACCTCCGCTACTGCCGCCAGGCAAATTC	
fw_cdh_policistronic	GAATTTGCCTGGCGGCAGTAGCGGAGGTGACTAATGAAAAAAGCGGGTCTTC	
rv_cdh_policistronic	CTTTCGACTGAGCCTTTCGTTTAACGCAAAATCTCACAC	
fw_fadR_p5T7fadR	ACTTTAATAAGGAGATATACCATGGTCATTAAGGCGCAAAG	p5T7-fadR
rv_fadR_p5T7fadR	GGTGGCAGCAGCCTAGGTGTCTTATCGCCCCTGAATGGCTAAATCACCC	
fw_p5T7_p5T7fadR	GATTTAGCCATTCAGGGGCGATAAGACACCTAGGCTGCTGCCACC	
rv_p5T7_p5T7fadR	CTTTGCGCCTTAATGACCATGGTATATCTCCTTATTAAAGT	

Table S2. Glycosyltransferase GT MG517 specific activity and total glycoglycerolipids (GGL) production for selected strains.

	Strain	$[GGL]_T$ (nmol/mg cell)	Specific activity ($\mu M \cdot min^{-1}$/mg P_T) [1]	Fluorescence (F/mg P_T) [2]
#1	WT/ mg517-plsC^H	3.26	0.6	
#2	ΔtesA/ mg517-plsC^H	5.69	1.9	
#3	ΔfadE/ mg517-plsC^H	3.72	0.3	
#4	ΔtesA ΔfadE/ mg517-plsC^H	2.31	0.2	
#5	ΔtesA ΔfabR/ mg517-plsC^H	4.84	1.6	
#18	ΔtesA ΔushA/ mg517-plsC^H	8.63	0.3	
#15	ΔtesA/ mg517-plsCxpgpB^H	9.93	1.9	
#23	ΔtesA ΔushA/ mg517-plsCxpgpB^H	7.22	0.9	
#24	WT/ mg517xmCherry-plsC^H		0.3	25
#25	ΔtesA/ mg517xmCherry-plsC^H		4.2	63
#26	ΔfadE/ mg517xmCherry-plsC^H		0.8	9
#27	ΔtesA ΔfadE/ mg517xmCherry-plsC^H		0.2	11
#28	ΔtesA ΔfabR/ mg517xmCherry-plsC^H		0.4	43
#29	ΔtesA ΔushA/ mg517xmCherry-plsC^H		0.4	28

[1] Specific activity of the cell-free extracts expressed as μM product per min and mg of total protein using UDP-galactose (UDPGal) and C6-NBD-ceramide ((N-[6-[(7-nitro-2-1,3-benzoxadiazol-4-yl)amino]hexanoyl]-D-erythro-sphingosine) as substrates.

[2] Protein expression determined by mCherry fluorescence of the cell-free extract and expressed as fluorescence intensity per mg of total protein.

Table S3. Percentage of fatty acids composition from lipids extracts of engineered strains.

	C14:0	C16:0	C16:1	C17:0Δ	C18:0	C18:1	C19:0Δ
Control	7 ± 3	45 ± 2	11 ± 3	16 ± 1	1.2 ± 0.1	13 ± 3	6 ± 3
(#1) WT/ mg517-plsC^H	3.1 ± 0.1	45.9± 0.2	4.3 ± 0.1	27.0 ± 0.1	1.2 ± 0.1	8.0 ± 0.1	10.6 ± 0.3
(#2) ΔtesA/ mg517-plsC^H	2.8 ± 0.2	51.9 ± 0.8	7.2 ± 0.4	18 ± 1	10 ± 2	7.1 ± 0.7	3.1 ± 0.5
(#3) ΔfadE/ mg517-plsC^H	3.7 ± 0.1	52.0 ± 0.5	4.9 ± 0.1	21 ± 0.5	7.7 ± 0.1	6.0 ± 0.2	4.7 ± 0.1
(#4) ΔtesA ΔfadE/ mg517-plsC^H	4.3 ± 0.8	51.3 ± 0.6	3.3 ± 0.3	22 ± 2	9 ± 2	4.0 ± 0.4	5.2 ± 0.5
(#5) ΔtesA ΔfabR/ mg517-plsC^H	2.1 ± 0.3	29 ± 1	18 ± 1	15.4 ± 0.1	2.9 ± 0.1	29.6 ± 0.3	2.8 ± 0.4
(#18) ΔtesA ΔushA/ mg517-plsC^H	3.8 ± 0.2	39 ± 17	5 ± 2	17 ± 8	5 ± 2	4 ± 2	3 ± 1
(#19) ΔtesA ΔushA/ mg517-plsC^H·plsB^H	2.3 ± 0.5	42 ± 3	15 ± 0.8	15.2 ± 0.9	13.5 ± 0.1	16 ± 4	3.0 ± 0.9

Figure S1. Membrane lipids composition. 1D TLD and autoradiography of ^{14}C-labelled lipids. Elution with chloroform, methanol and water (65:25:4). Quantification in mol% by electronic autoradiography (Bio-rad Imager). nGDAG: mono-, di-, tri, tetra-glycosyldiacylglycerol; PE: phosphatidylethanolamine; CL: cardiolipin; PG: phosphatidylglycerol.

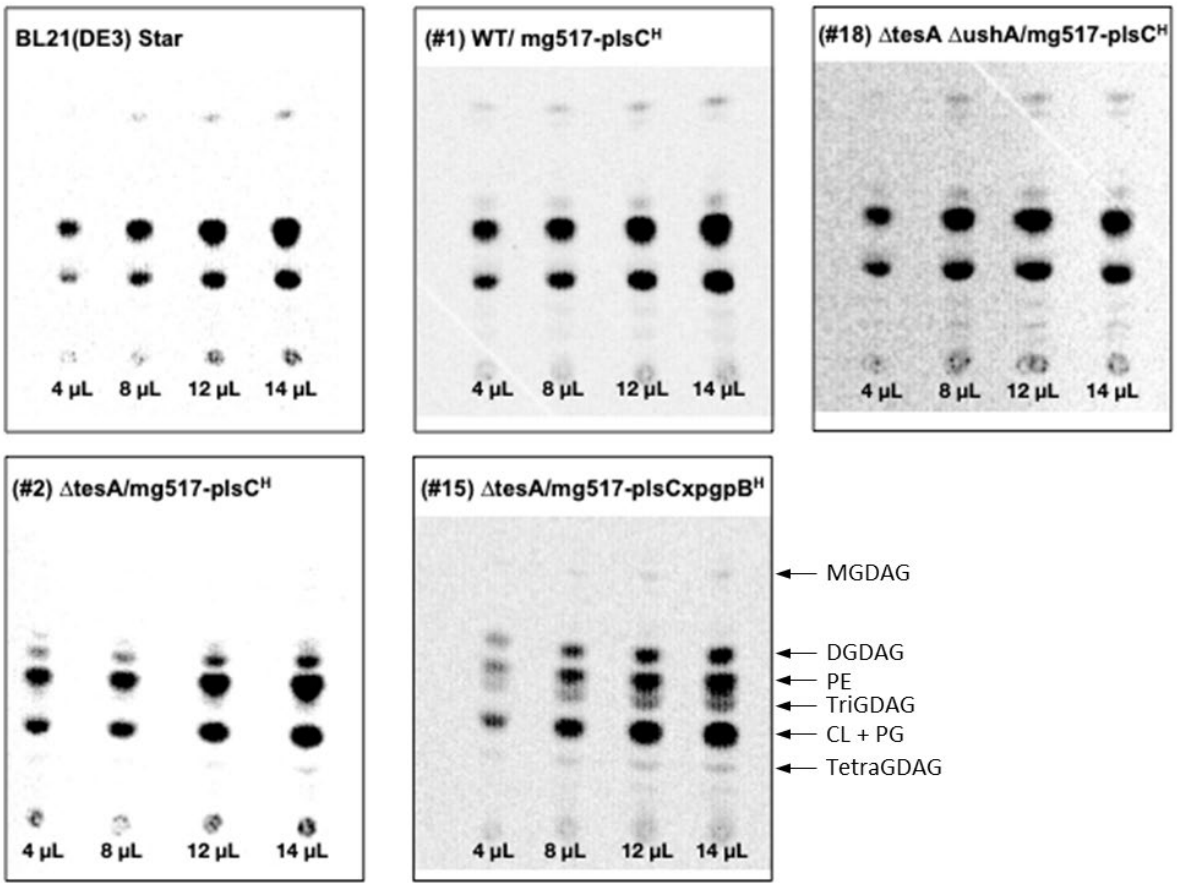

Figure S2. Membrane lipids composition. 2D TLD and autoradiography of ^{14}C-labelled lipids. Elution with chloroform, methanol and water (65:25:4) as first eluent and ethyl acetate/isopropyl alcohol/chloroform/methanol/0.25% KCl in water (25:25:25:11:9) as second dimension eluent. Quantification in mol% by electronic autoradiography (Bio-rad Imager). Autoradiography was exposed for 16 hours and visualized and quantified in mol% by electronic autoradiography (Bio-rad Imager). (1) CL + PG, (2) PE, (3) MGDAG, (4) DGDAG, (5) TGDAG, (6) TetraGDAG. nGDAG: mono-, di-, tri, tetra-glycosyldiacylglycerol; PE: phosphatidylethanolamine; CL: cardiolipin; PG: phosphatidylglycerol.

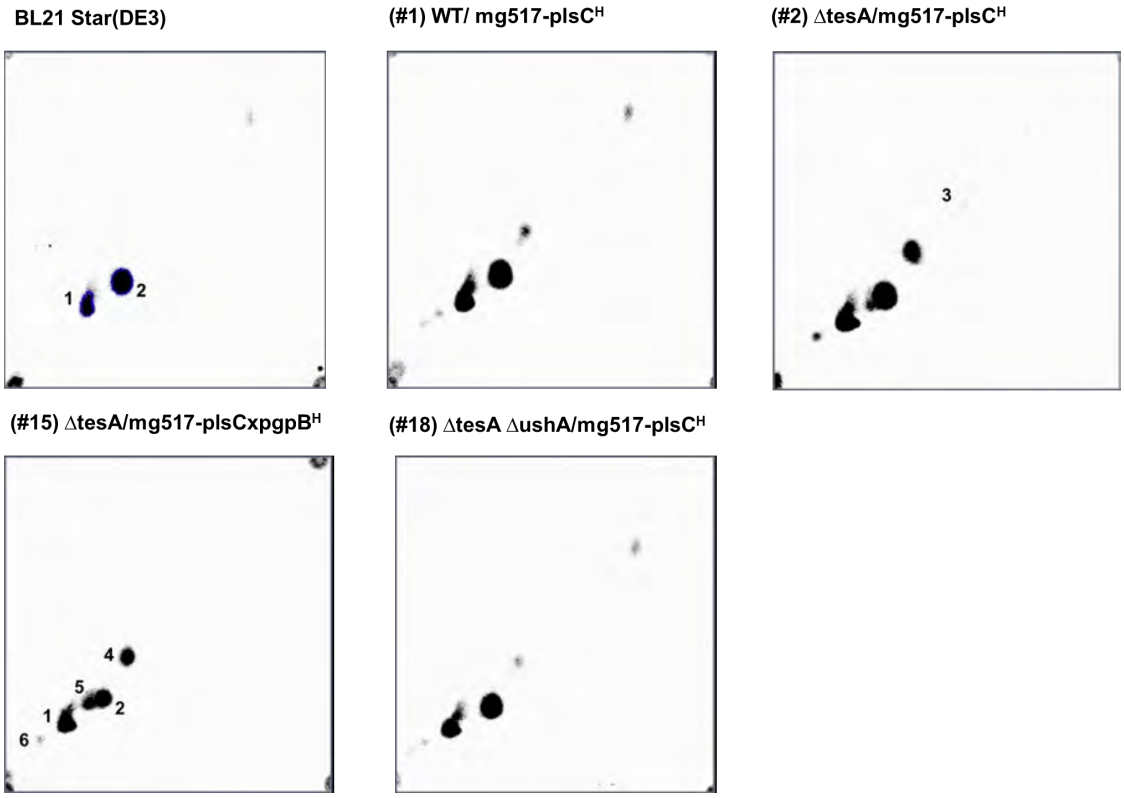

Figure S3. Morphology of engineered cells by optical microscopy.