Université de Lille – Faculté des Sciences et Technologies

Ecole Doctorale Biologie et Santé (ED446)

Thèse rédigée par Louis-David LECLERCQ, soutenue le 24 mars 2023.

pour l'obtention du grade de docteur de l'Université de Lille,

Spécialité : Aspects moléculaires et cellulaires de la biologie.

Etude de la relation structure-fonction des glycolipides de *Mycobacterium abscessus*

Président :	**Pr. Fabrice ALLAIN**, Professeur des Universités, HDR, UGSF, UMR CNRS 8576, Villeneuve d'Ascq, FRANCE
Rapporteurs :	**Dr Anne LEMASSU**, maîtresse de conférences, HDR, Institut de Pharmacologie et Biologie Structurale UMR CNRS 5089, Toulouse, FRANCE
	Dr Thierry FONTAINE, HDR, Institut Pasteur de Paris, FRANCE
Examinateurs :	**Pr. Françoise ROUTIER**, Professeur des Universités, Faculté de médecine de Hanovre, ALLEMAGNE
	Dr Franck BIET, HDR, Infectiologie et Santé Publique UMR INRAe 1282, Tours, FRANCE
Directeur de thèse :	**Dr. Yann Guérardel**, Directeur de recherche, UGSF, UMR CNRS 8576, Villeneuve d'Ascq, FRANCE

Ces travaux ont été réalisés dans l'équipe de Chémoglycobiologie du Pr Christophe BIOT au sein de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF) UMR CNRS 8576, et ils ont bénéficié d'une bourse de financement du Ministère de l'Enseignement Supérieur et de la Recherche.

AVANT-PROPOS

A la question « qu'est-ce que tu voudrais faire quand tu seras grand ? » que je percevais parfois comme une injonction envers les jeunes élèves ni très bêtes, ni très brillants, je répondais banalement « ne pas vraiment savoir » comme si cette préoccupation ne m'eut jamais traversé l'esprit. A table, on peut goûter avant de savoir que l'on n'aime pas, mais au travail, c'est plus compliqué ! J'étais ainsi le fruit d'une vie trop insouciante et trop comblée par mes chers parents. Mon père, enseignant les mathématiques et la physique-chimie, avait alimenté mon attrait pour les raisonnements logiques et les sciences dures, notamment au travers de l'observation du vivant. A l'aube de mon adolescence, je continuais de profiter et d'expérimenter la vie en bravant des interdits qui trouveront leur tombe ici. J'obtins tout de même un baccalauréat scientifique, mention assez bien, selon l'éducation nationale qui était du même avis que ma mère : « Si tu avais travaillé, tu aurais obtenu la mention bien ! »

Mon éducation, pour partie religieuse, et ma personnalité me conduisirent à opter pour un parcours universitaire qui me permettrait d'aider autrui. Sur les bancs de la faculté libre de médecine, mon indépendance, sans être solitude, mon intelligence, toute relative, et mon rejet de toute compétition ne furent pas particulièrement utiles : j'étais « reçu-collé » ! Cette expression espiègle vous rappelle simplement que vous n'aviez pas toutes les réponses aux questions, ou plutôt aux colles. Peu importe la punition, j'avais alors et jusqu'aujourd'hui trouvé la compagnie de Karine qui me comble quotidiennement de son amour, de son sourire et de sa force.

Elle persévérait dans le monde médical pour devenir sage-femme alors que mon manque de courage envers les aiguilles et les seringues, plus que de la véritable bélonéphobie, m'amena à poursuivre l'étude de la biochimie. Après l'obtention de ma maitrise, je n'insistai pas pour intégrer le DEA et je fus sanctionné l'année suivante d'un DESS de protéomique. Je justifiai de deux expériences significatives en contrôle qualité chez ROQUETTE FRERES, et en recherche chez LESAFFRE INTERNATIONAL. Prospectant le marché du travail, je ne réussis pas à convaincre les laboratoires publics ou privés impliqués dans la protéomique de m'accorder leur confiance. Dans le même temps, j'écoutai la parole bienveillante de mon père qui me conseilla d'enseigner en collège et lycée les sciences de la vie et la physique-chimie en tant que vacataire. Après ces 18 mois, je ne m'étais pas imaginé l'implication émotionnelle de ce métier…

Au fur et à mesure de mes recherches, le poste d'ingénieur commercial se profilait comme la solution idéale pour construire une famille, valoriser mon diplôme et demeurer dans un domaine disciplinaire que j'affectionnais. En 2005, lors de notre mariage avec Karine, j'étais embauché par

Roger BIGOT chez INTERCHIM pour promouvoir la gamme chimie, biologie et chromatographie. J'appréciai le contact et l'échange avec les utilisateurs. Nous étions les spéléologues du catalogue et nous recherchions la référence tant désirée parmi plusieurs autres milliers, voire millions. J'appris beaucoup sur des domaines diversifiés satisfaisant ma curiosité. INTERCHIM se révéla être une bonne école pour un débutant qui savait travailler de façon autonome. Le 12 novembre 2006, Garance naquit et j'avais maintenant la responsabilité d'une petite tête blonde : il fallait être plus prudent.

En mars 2009, la prudence m'abandonna puisque je démissionnai pour intégrer la société DIONEX alors que Karine et moi attendions la naissance de Domitille, le 13 septembre 2009. L'entreprise me sollicitait depuis plusieurs mois pour prendre en charge la vente de leurs instruments de chromatographie sur le Nord de la France. J'étais probablement flatté et j'imaginai naïvement rejoindre leur laboratoire d'application à moyen terme. Fabien GUERIN, mon manager, fut aussi naïf que moi car mes premiers résultats démontrèrent que je pratiquai, comme chez INTERCHIM, une forme d'accompagnement totalement désintéressé des résultats commerciaux. Il m'imposa donc le tutorat d'un formateur commercial, Adrien LEGRAND. Sa stratégie de développement se basait sur la connaissance de soi-même : j'étais un travaillomane, empathique et rêveur ! J'aimai son approche mais elle n'avait pour seul but que d'inverser la tendance décroissante de mes ventes. Avec son aide, nous réussîmes à trouver les adaptations nécessaires de mon travail, et acceptables de ma personnalité pour satisfaire aux exigences de résultats : je dus m'adresser davantage à ceux qui possédaient le carnet de chèque, et moins aux utilisateurs ! De plus, les membres de l'équipe commerciale et support de DIONEX m'avaient bien accueilli et entretenaient une ambiance de travail partagée entre solidarité, proximité et émulation intellectuelle qui me stimulaient : nous étions des spécialistes « redouté(e)s » de la chromatographie liquide puisque cette technique était la seule que nous proposions pour survivre. Parmi les projets plus complexes de couplage avec la spectrométrie de masse, ma formation en protéomique et mon cursus universitaire en biochimie devinrent salvateurs pour que je puisse atteindre des objectifs plus ambitieux.

Ce fut probablement prémonitoire car THERMO FISHER SCIENTIFIC, fabricant du fameux Orbitrap, largement utilisé en protéomique, racheta DIONEX plus de deux milliards de dollars en décembre 2010. En France, le rapprochement progressif des sociétés se fit jusqu'en 2013. Pour mener ma petite barque dans la « World Company », j'héritai d'un nouveau manager, José FERREIRA, et les habitudes s'installaient : je chérissais d'abord au mieux les utilisateurs et je les assistais notamment dans l'utilisation du logiciel Chroméléon ; je réalisais des démonstrations d'instrument seul ou avec mes collègues du support auprès des décideurs techniques pour comprendre leurs besoins ; je visitais

des clients installés qui me faisaient confiance en compagnie de prospects pour favoriser entre eux l'échange et l'émulation, ce qui me semblait être pour eux le meilleur moyen de faire un bon choix ; je rédigeais les cahiers des charges pour répondre aux appels d'offres, en particulier sur les nouveaux produits de chromatographie gaz et de spectrométrie de masse inclus dans ma gamme ; j'apprenais comment orienter ma négociation commerciale en position de force ou de faiblesse pour garantir les résultats attendus ; je proposai ma candidature dans des fonctions plus applicatives et en lien avec l'analyse chimique. Bref, j'avais la tête dans le guidon… tant que le vélo restait sur la route ! En même temps, pour travailler chez THERMO FISHER SCIENTIFIC, on vous donnait un beau vélo tout neuf.

Au cours de cette carrière, plusieurs évènements vinrent perturber cette course, à l'approche de la soi-disant crise de la quarantaine. Je ne suis pas certain de les énumérer dans l'ordre. Le premier d'entre eux fut le licenciement d'un collègue de travail, Jonathan DESANNAUX, qui m'avait accompagné chez DIONEX. Non seulement il m'apparut injuste de le renvoyer parce qu'il ne dépassait pas régulièrement un objectif mercantile décidé de façon unilatérale, mais surtout je pense que je m'étais identifié à lui et que je me sentis en danger. Le second se déroula dans le cadre d'une formation commerciale organisé par THERMO FISHER SCIENTIFIC. Un test préliminaire révéla que mes traits conscients et inconscients de personnalité étaient fortement éloignés de ceux attendus d'une fonction commerciale, à l'inverse de l'ensemble de mes collègues de l'époque. J'avouai le savoir instinctivement car je m'intéressai à d'autres postes internes, ou à d'autres professions externes qui se rapprochaient davantage de mes convictions et aspirations. D'ailleurs, une collègue et amie de travail, le Dr Audrey BRAULT, ainsi que le Dr François COUTTE qui me reçut pour un entretien d'embauche m'avaient encouragé à poursuivre un projet de thèse… Je pense que j'ai pris conscience du dernier évènement après son déroulement. En effet, à l'image de ma fille Garance qui avait désiré puis planifié sa future intégration dans une section sportive de natation parce que c'était sa passion, j'avais réfléchi et préparé mon départ de THERMO FISHER SCIENTIFIC avec autant d'abnégation. Inconsciemment, je suis sûr que l'avoir accompagné dans ce cheminement a nourri et initié le mien.

Après négociation avec mon responsable, José FERREIRA, ce dernier convint THERMO FISHER SCIENTIFIC d'accepter la rupture conventionnelle pour la fin du mois de juin 2018. Sur le 1er semestre 2018, je ne travaillai plus que quatre jours par semaine et atteignit l'objectif annuel dans les six mois : j'avoue que l'annonce de mon départ auprès des clients eut un effet auquel je ne m'attendais pas. Dans le même temps, je recherchai les formations susceptibles de m'accueillir en septembre 2018 afin de poursuivre une activité dans le domaine de l'analyse chimique et biochimique. Au lieu de vendre les instruments, je préférai les utiliser. C'est ainsi que je redevins étudiant à l'Université de Lille, en

Master Chimie et Sciences du Vivant, dans le but de conforter mes savoirs théoriques, l'objectif ultime étant de décrocher un projet de thèse pour démontrer mon savoir-faire et mon savoir-être dans un contexte de recherche académique. Je retrouvai d'anciens enseignants : le Dr Christophe MARILLER comprit mon projet et accepta ma candidature ; le Dr Frédéric CHIRAT facilita mon intégration dans la promotion ; les Dr Florence DELPLACE, Elisabeth ELASS-ROCHARD, Isabelle FOURNIER, Tony LEFEBVRE et Fabrice ALLAIN rajeunissaient de quinze ans grâce à mon retour… Recherchant un stage, je rencontrai « mes » anciens clients dont la plupart furent attentifs à mon projet : ici, je les en remercie.

Au fil des rencontres, c'est un entretien avec le Dr Yann GUERARDEL, initié par le Dr Christophe MARILLER, qui fut le point de départ de cette histoire. Yann fut parfaitement compréhensif sur l'idée de ma « reconversion » car sa compagne avait fait de même. Il m'exposa simplement les conditions dans lesquelles il prévoyait de s'engager sur le long terme, dans l'étude de la structure des glycolipides d'une nouvelle espèce de mycobactérie. Il avait fait une demande de financement de thèse, il entretenait depuis longtemps une collaboration fructueuse avec le Dr Laurent KREMER et il deviendrait bientôt directeur de l'unité de glycobiologie structurale et fonctionnelle. Son discours était direct, authentique, réfléchi et fédérateur. Le contexte scientifique et opérationnel me convenait et j'acceptais avec enthousiasme de démarrer mon stage dans son équipe. Je ne fus pas déçu puisque Yann et tous les membres de l'équipe se rendirent disponibles dès le début pour me mettre sur de bons rails. Après 6 mois, j'avais obtenu des premiers résultats encourageants, j'avais pratiqué l'analyse RMN, en plus de la chromatographie et de la spectrométrie de masse, et je décrochai le financement d'une thèse auprès de l'école doctorale biologie et santé. Je fus surpris de ce succès. Jamais je n'avais réussi de concours. Le jury avait certainement entendu ma motivation malgré la singularité de mon profil, ou grâce à elle... Je pus non seulement poursuivre les travaux entrepris durant mon stage au travers du projet de thèse, mais aussi m'engager pleinement dans une vie nouvelle. Adélaïde, ma 3ème fille, allait naître le mois suivant, le 10 juillet 2019.

Durant mes quatres années de thèse, j'ai aimé écouter, apprendre, comprendre, imaginer, tester, expérimenter, répéter, douter, évaluer, construire, illustrer, rédiger, échanger, transmettre, partager, débattre, transiger, conclure… J'aurais également aimé que le fait de construire un socle de connaissances universelles assure un bénéfice partagé au plus grand nombre d'entre nous. Ma contribution demeura modeste mais engagé, et mon bénéfice s'est élevé au-delà. En effet, je ne pourrai pas exprimer plus de gratitude que ce que j'ai reçu de toutes les personnes qui ont croisé mon chemin pendant la réalisation de cette thèse. S'il est une personne qui lit jusque cette dernière ligne, qu'elle se sente visée puisqu'elle y aura participé.

REMERCIEMENTS

Je remercie d'abord le Pr Fabrice ALLAIN, président du jury, ainsi que ses membres : les Dr Anne LEMASSU, Dr Françoise ROUTIER, Dr Thierry FONTAINE et Dr Frank BIET qui ont accepté d'évaluer mon projet de thèse, en lisant ce rapport et en assistant à ma défense. J'en profite pour reconnaitre la précieuse contribution du Dr Emilie LAYRE et du Dr Romain VEYRON-CHURLET et de leurs conseils avisés durant les comités de suivi individuel.

Je remercie le Dr Frédéric CHIRAT qui m'a simplement encouragé au début de mon projet lorsque je m'apprêtais à réintégrer le monde de la recherche universitaire. Il m'a rassuré sur les moyens d'y parvenir en m'expliquant qu'aucun obstacle n'est insurmontable. Plus qu'un enseignant bienveillant, il fut un collègue de travail espiègle et curieux qui participait à embellir notre journée par des interactions tantôt anodines, tantôt essentielles.

J'adresse ma plus sincère reconnaissance auprès du Dr Christophe MARILLER et du Pr Christophe BIOT. En plus de m'avoir accueilli dans le Master CSV, le Dr MARILLER m'a proposé de rencontrer Yann pour démarrer un stage de M2, et le Pr BIOT a recommandé ma candidature lors du concours de l'école doctorale. Ce fut primordial pour la continuité et la réussite de ce projet.

Les membres de la plateforme PAGès ont été les pionniers de ma reconquête de la chimie analytique : le Dr Bernadette CODDEVILLE pour l'utilisation des instruments de GC et GC-MS n'est pas resté assez longtemps pour m'en apprendre davantage, le Dr Nao YAMAKAWA pour l'analyse MALDI-TOF et la LC-MS sans oublier le Dr Emmanuel MAES pour l'initiation à la RMN. Chacun d'eux m'a appris, m'a aidé et m'a conseillé sous quelque forme qui lui convenait. Leur support m'a redonné confiance et leurs enseignements m'accompagneront pour longtemps. Je remercie aussi Charlotte CLARISSE et Louise DESCHANEL de leur gentillesse et de leur flexibilité. Enfin, le Dr Xavier TRIVELLI de la plateforme IR-RMN de Lille a été mon guide dans les arcanes tourmentés de l'analyse RMN. Il a su progressivement m'amener à la compréhension de cet outil exigeant, et m'impliquer dans son utilisation quotidienne au point d'en avoir obtenu le meilleur. Je remercie également le Dr Héléne LAVANANT et le Pr Isabelle SCHMITT-ALONSO pour leur aide dans l'obtention des résultats FT-ICR.

J'aimerai témoigner également toute ma gratitude à l'ensemble des membres passés ou présents de l'équipe de chémoglycobiologie qui ont participé de près ou de loin à me stimuler durant ces quatre années : merci au Dr Elisabeth ELASS-ROCHARD pour avoir participé à ma recherche bibliographique et pour avoir partagé ses connaissances sur les mycobactéries ; merci au Dr Frédéric KRZEWINSKI de m'avoir fait confiance pour le monitorat et l'utilisation du matériel ; merci à Audrey

AUGER pour ces séances de jeu de fin d'année et son dynamisme quotidien ; merci à Vincent RIGOLOT pour son écoute et son flegme qui rendra jaloux tout citoyen britannique ; merci à Yannick MASSON pour son empathie et nos discussions impliquées ; merci à Réginald LEFEBVRE pour son humilité et sa simplicité ; merci au Dr Florence DELPLACE et au Dr Isabelle HUVENT pour leurs encouragements et le partage de leur précieuse expérience ; merci au Dr Yannick ROSSEZ pour avoir été simultanément parmi les premiers et les derniers témoins de l'achèvement de ce projet ; merci au Dr Estelle GARENAUX et au Dr Shin Yi YU pour les moments de détente et de convivialité ; merci au Dr Cédric LION de sa subtil clairvoyance sur les détails superflus ; merci à Clara DESSENNE et Jeanne PAYSANT qui furent des stagiaires volontaires et exemplaires. Merci à toutes les personnes qui ont croisé ma route pendant ces quatre années : Clémence, Lin, Faustine, Alexandre, Angélina, Laura, Zina, Lucie, Clément, Robert, Jessie, Olivier, David, Marc, François-Xavier, Bertrand, Quentin, Adeline, Corentin... et toute l'unité UMR CNRS 8576 !

L'ensemble de ce projet et de ces résultats n'auraient pas été possibles sans la vision du Dr Laurent KREMER de l'Institut de recherche en infectiologie de Montpellier, ainsi que la participation de mon jumeau inopiné, le Dr Wassim DAHER et sa doctorante Morgane ILLOUZ. Malheureusement, le COVID ne nous a autorisé que des contacts distants et professionnels. Néanmoins, je suis certain qu'ils auraient pu être encore plus chaleureux. Ensemble, nos noms sont écrits comme témoins de cette collaboration fructueuse. J'espère qu'ils le seront de nouveau et que nous finirons par nous rencontrer !

Cher Yann, lors de notre première discussion, tu m'as expliqué que tu concevais parfaitement mon souhait d'embrasser les métiers de la recherche même tardivement parce que ta compagne avait également choisi une nouvelle voie professionnelle afin de mieux s'épanouir. Même si ton histoire a influencé ton jugement, je crois que les véritables raisons de ta sollicitude sont davantage liées à ce que tu es. L'enthousiasme et la soif d'apprendre, la tolérance et l'humilité, l'altruisme et le partage, la persévérance et l'engagement sont des valeurs qui correspondent d'après moi à ce que tu es. Il n'est pas question ici de flatter ni la noblesse d'un homme ni celle de son métier mais d'illustrer ce qui nous a rassemblé au travers de ce projet. Je te suis extrêmement reconnaissant de m'avoir épaulé durant ce projet, autant pour m'avoir témoigné ta confiance lors du concours devant le jury de l'EDBSL que pour m'avoir insufflé ton énergie débordante et ton acuité intellectuelle dans la réalisation de ma thèse. Je suis sincèrement heureux de refermer un chapitre de cette histoire sans pour autant la conclure.

LISTE DES ABBREVIATIONS

aac(2)' : aminoglycoside acétyltransférase
AAI : Amino-acid identity
Ac_2TMM : Di-*O*-acyl tréhalose monomycolate
ACP : Acyl carrier protein
ADN : Acide désoxyribonucléique
AftA-D : Arabinofuranosyltransférase
Ag : Antigène
ANI : Average Nucleotide Identity
Araf : Arabinofuranose
AraLAM : Lipoarabinomannane sans coiffe
ARN : Acide ribonucléique
ARN_m : Acide ribonucléique messager
arr : ADP-ribosyl rifampicin transferase
ATCC : American Type Culture Collection
atf : acétyltransférase
ATFA : Acide trifluoroacétique
ATP : Adénosine triphosphate
atpE : ATP synthase
ATS : American thoracic society
BCG : Bacille de Calmette Guérin
BCKADH : Branched-chain keto acids deshydrogenase
bla : β-lactamase
C : Cytosine
CCL : Chemotactic cytokine ligand
CCM : Chromatographie couche mince
CD : Class differentiation
Chp2 : DAT acyltransférase
COSY : Correlation spectroscopy
CmaA :Cyclopropane mycolic acid synthase
CR : Complement receptor
CSP : Conserved Signature Protein
CSI : Conserved Signature Indel
CXCL : Chemotactic cytokine X ligand
DCC : Dominant circulating clone
DC-SIGN : Dendritic Cell-Specific Intercellular Grabbing Non-integrin
DDH : DNA-DNA hybridation
DHB : Di Hydroxy Benzoïc acid
DAT : Di-*O*-acyl tréhalose
DMAG : Dimycolyl diarabinoglycérol
DSMZ : Deutsche Sammlung von Mikroorganismen und Zellkulturen
EIMS : Electronic impact mass spectrometry
eis : enhanced intracellular survival
ELISA : Enzyme linked immuno sorbent assay
embA-C : Arabinofuranosyltransférase (inhibée par l'éthambutol)
erm(41) : erythromycine ribosome methylase
ESX : Système de sécrétion de type VII
fmt : fatty acid methyltransferase
FadD32 : Fatty acyl-AMP ligase
FAMEs : Fatty acid methyl esters
FAS : Fatty acid synthase
FT-ICR : Fourier Transform Ion Cyclotron Resonance
G : Guanine
gap : GPL addressing protein
GC : Gas chromatography
GCMS : Gas chromatography with mass spectrometry
GDP : Guanosine diphosphate
GDND : Glycosyl Di-*O*-acyl Nonadecyl Diol
GlfT : Galactofuranosyltransférase
gtf : glycosyltransferase
GMM : Glucose monomycolate
GPL : Glycopeptidolipide
nsGPL : Glycopeptidolipide non spécifique
gyrA : DNA gyrase sous-unité α
gyrB : DNA gyrase sous-unité β
Had : β-hydroxyacyl-ACP dehydratase
HβD2 : Human β défensin 2
HeLa : Henrietta Lacks
HIV : Human immunodeficient virus
HMBC : Heteronuclear multiple bond correlation
HPLC : High performance liquid chromatography
HSQC : Heteronuclear single quantum correlation
HSP65 : Heat shock protein 65
ICAM-1 : Intercellular adhesion molecule
IDSA : Infectious Disease Society of America
IFN : Interferon
Ig : Immunoglobuline
IHC : Immunohistochimie
IL : Interleukine
Ins : *myo*-inositol
inhA : enoyl-ACP reductase (inhibée par isoniazide)
ino1 : inositol synthase
IRIM : Institut de recherche en infectiologie de Montpellier
ITS1 : internal transcribed spacer sequence
KatG : catalase peroxidase
KasA : β-ketoacyl ACP synthase
LAM : Lipoarabinomannane
LC-MS : Liquid chromatography coupled to mass spectrometry
LM : Lipomannane
LOS : Lipoligosaccharide
LprG : Lipoprotéine G
LpqY : Lipoprotéine Y
lsr2 : leprae serum responsive clone 2
mabA : β-ketoacyl-ACP reductase
MABSC : Complexe Mycobacterium abscessus

MAC : Complexe Mycobacterium avium intracellulare
mAGP : Mycolyl arabino galactane peptidoglycane
MALDI : Matrix asssisted laser with desorption and ionisation
MAMEs : Mycolic acid methyl ester
ManLAM : Lipoarabinomannane avec une coiffe mannose
mbtH : mycobactin H (gène et homologue)
MCE : Mammalian cell entry protein
MgtC : Magnesium transporter C
MIC : Minimum inhibitory concentration
MIRU-VNTR : Mycobacterial identification repetitive unit – variable number tandem repeat
MmaA : Mycobacterial mycolic acid synthase
MMDAG : Méromycolyl diacyglycérol
MmpL : Mycobacterial membrane protein large
MmpS : Mycobacterial membrane protein small
MNT : Mycobactéries non-tuberculeuses
mps : mycobacterial peptide synthase
MR : Mannose receptor
MS : Mass spectrometry
MsrA : methionine sulfoxide réductase A
Mtf : méthyltransférase
MTX : Méthylthioxylose
NCTC : National Collection of Type Cultures
NOESY : Nuclear Overhauser Effect Spectroscopy
NRPS : Non Ribosomal Peptide Synthase
OADC : Oleic acid/Albumin/Dextrose/Catalase
OMS : Organisation mondiale de la santé
Ots : Osmoregulatory trehalose synthesis
PatA : Phosphatidylinositol mannoside acyltransférase A
PAT : Penta-*O*-acyl tréhalose ou polyacyltréhalose
PCR : Polymerase chain reacion
pcaA : proximal cyclopropanation of α mycolic acid
PDIM : Phtiocerol dimycocerosate
PGE2 : Prostaglandine rénale E2
PGL : Phénolglycolipide
PILAM : Lipoarabinomannane avec une coiffe phosphoinositol
PimA-F : mannosyltransférase
PIMs : Phosphatidylinositolmannosides
PIM_2 : Phosphatidylinositoldimannosides
PIM_6 : Phosphatidylinositolhexamannosides
pks : polyketide synthase
PMA : Acide phosphomolybdique
PncA : pyrazinaminidase
POCP : Percentage of conserved proteins
QIT : Qaudrupole Ion Trap
Rf : Rapport frontal
RFLP : Restriction Fragment Lenght Polymorphism
rpoB : RNA polymérase, sous-unité β
RMN : Résonance magnétique nucléaire
rrl : ribosomal RNA large (ARN 23S)
rrs : ribosomal RNA small (ARN 16S)
SARS-COV2 : Severe acute respiratory syndrome

LISTE DES FIGURES

Figures	**Titres**
1	Répartition phylogénétique du genre *Mycobacterium*
2	Aggrégation des mycobactéries.
3	Facteurs de risque des infections pulmonaires associées à *Mabs*
4	Cycle infectieux de *Mabs* chez le poisson zèbre
5	Traitement des infections pulmonaires liées à *Mabs*
6	Schéma de la paroi des mycobactéries
7	Schéma général de la biosynthèse des PIMs
8	Schéma général de l'enchainement oligosaccharidique du LM et du LAM
9	Schéma général de l'arabinogalactane
10	Structure des acides mycoliques de quelques espèces mycobactériennes
11	Loci de biosynthèse des GPLs chez *Msmeg* et *Mabs*
12	Schéma général de la structure chimique du GPL de *Mabs*
13	Comparaison des peptides synthases non ribosomales de *Mav*
14	Devenir de *Mabs* R et S après phagocytose
15	Le contexte du gène MAB_*4690*
15	Profil CCM de Mabs
16	Caractérisation des lipides *apolaires* de *Mabs*
17	Analyse MALDI-TOF du TDM de *Mabs*
18	Analyse RMN du TDM de *Mabs*
19	Analyse RMN de l'Ac2TMM
20	Acylation de l'Ac2TMM
21	Analyse MALDI-TOF de l'Ac_2TMM
22	Fragmentation MALDI-TOF de l'Ac2TMM
23	Contenu en acides mycoliques de *Mabs*
24	Comparaison des acides mycoliques de *Mabs* R et S
25	Caractérisation RMN des acides mycoliques *de* Mabs
26	Acides mycoliques totaux et issus du *TDM*
27	Composition en monosaccharides de la fraction polaire de *Mabs* S et R
28	Purification de la fraction polaire de Mabs
29	Production des DATs selon le milieu *de* culture
30	Caractérisation du tréhalose dans les DATs
31	Analyse RMN du 3,4 DAT
32	Substitution lipidique du 3,4 DAT
33	Analyse RMN du 2,3 DAT
34	Acylation du 2,3 DAT
35	Composition lipidique du 2,3 DAT
36	Profil CCM de la souche clinique GD163
37	Analyse RMN de l'acyl tréhalose de la souche clinique GS163
38	Analyse MALDI-TOF de l'acyl tréhalose de la souche clinique GS163

39	Analyse MALDI-TOF du PI, PIM et PIM2
40	Analyse MALDI-TOF de la fraction 40 contenant l'Ac2PIM2
40	Analyse MALDI-TOF de la fraction 57 contenant les PIM_6
42	Composition en monosaccharides des PIMs
43	Spectres EIMS des monosaccharides issus des PIMs
44	Composition lipidique des AcPIM2
45	Spectres EIMS des FAMEs issus des PIM_2
46	Analyse RMN de l'Ac2PIM2
47	Attribution des signaux RMN de l'Ac_2PIM_2
48	Structure de l'Ac2PIM2 purifié de Mabs
49	Caractérisation du GPL-2a succinylé dans *Δgtf3*
50	Caractérisation du lipopeptide dans Δgtf1,2
51	Analyse des GPLs présents *dans Δgtf1*
52	Analyse des GPLs présents dans *Δgtf1,3*
53	Schéma de glycosylation des GPLs *de Mabs* S
54	1ère purification du GPL4690
55	2ème purification du GPL4690
56	Analyse FT-ICR du GPL4690 et de son dérivé
57	Analyse RMN du GPL4690
58	Signaux aromatiques des acides aminés du GPL4690
59	Signaux aliphatiques des acides aminés du GPL4690.
60	Séquence peptidique du GPL4690 (par HMBC/CO)
61	Liaison des rhamnoses au peptide
62	Influence de la température dans les solvants protiques
63	Séquence peptidique du GPL4690 (par TOCSY/NOESY)
64	Attribution des signaux du GPL4690 dans les solvants protiques
65	Structure intermédiaire du GPL4690 après analyse FT-ICR et RMN
66	Analyse MALDI-TOF du GPL4690
67	Analyse MALDI-TOF d'un variant lipidique du GPL4690
68	Analyse MALDI-TOF d'un dérivé du GPL4690
69	Traitement alcalin du GPL4690 et de ses dérivés
70	Composition en monosaccharides du GPL4690
71	Structures identifiées du GPL4690
72	Prédiction de substrats ajoutés par les NRPS
73	Analyse CCM de Mabs R et de mutants dans la synthèse ou le transport de glycolipides
74	Couplage flash/MALDI-TOF *pour* la recherche du GPL4690
75	Test de complémentation du GPL4690 par couplage flash/MALDI-TOF
76	Analyse LC-MS du GPL4690 et de ses dérivés
77	Confirmation du motif glycannique du GPL4690 par LC-MS
78	Recherche LC-MS du GPL4690 dans différentes souches
79	Structure proposée du GPL4690
80	Schéma hypothétique de biosynthèse du GPL4690

LISTE DES GRAPHIQUES ET TABLEAUX

Graphiques	**Titres**
1	Evolution bibliométrique des articles traitant des espèces mycobactériennes majeures
2	
3	Croissance des espèces mycobactériennes au XXe siècle
4	Accroissement épidémiologique des MNTs
5	
Tableaux	**Titres**
1	Multirésistance antibiotique de *Mabs*
2	Acyl tréhaloses des principales espèces
3	LOSs des principales espèces
4	GPLs de *Mycobacterium avium*
5	GPLs d'autres espèces mycobactériennes
6	PGLs des principales espèces
7	Mélange de solvants utilisés pour la séparation CCM
8	Signaux RMN du tréhalose du TDM
9	Signaux RMN du tréhalose de l'Ac_2TMM
10	Comparaison des acides mycoliques de Mabs avec d'autres espèces.
11	Signaux RMN du tréhalose du 3,4 DAT
12	Distribution lipidique du 3,4 DAT
13	Signaux RMN du tréhalose du 2,3 DAT
14	Signaux RMN du tréhalose de l'acyl tréhalose issus de la souche GS163
15	Contenu glycannique des PIMs identifiés de *Mabs*
16	Contenu lipidique des PIMs identifiés de *Mabs*
17	Signaux RMN de l'Ac2PIM2
18	Formule brute du GPL4690
19	Formule brute du dérivé du GPL4690
20	Signaux RMN des acides aminés du GPL4690
21	Signaux RMN des rhamnoses du GPL4690
22	Recherche des rhamnosyltransférases impliqués dans la biosynthèse du GPL4690
23	Alignement des protéines impliquées dans la méthylation des GPLs
24	Alignement des protéines impliquées dans le transport des GPLs
25	Rôles biologiques des GPLs de *Mabs*

I INTRODUCTION 20
I.A LES MYCOBACTERIES QUI NOUS ENTOURENT 23
I.A.1 CLASSIFICATION TAXONOMIQUE DU GENRE MYCOBACTERIUM 23
I.A.1.a CLASSIFICATION DE RUNYON 23
I.A.1.b CLASSIFICATION SELON LE GENOME 25
I.A.2 DESCRIPTION MICROBIOLOGIQUE DU GENRE MYCOBACTERIUM 29
I.A.2.a MORPHOLOGIE, CULTURE ET CROISSANCE 29
I.A.2.b DETECTION DES MYCOBACTERIES 30
I.A.2.b.i Techniques microscopiques 30
I.A.2.b.ii Techniques microbiologiques 32
I.A.2.b.iii Techniques biochimiques 33
I.A.2.b.iv Techniques de biologie moléculaire 34
I.A.2.c COLONISATION DES NICHES ECOLOGIQUES 35
I.A.2.c.i Le sol 35
I.A.2.c.ii L'eau 35
I.A.2.c.iii L'air 35
I.A.2.c.iv La faune 36
I.A.3 LES MYCOBACTERIES NON TUBERCULEUSES 36
I.A.3.a EVOLUTION CROISSANTE DES MNTs FACE A LA TUBERCULOSE 37
I.A.3.b IMPACT ECONOMIQUE 38
I.A.3.c LA SITUATION EN FRANCE 38
I.A.3.d MYCOBACTERIUM SMEGMATIS : ESPECE MODELE 39
I.A.3.e LE COMPLEXE MYCOBACTERIUM ABSCESSUS 39
I.A.3.e.i Les sous-espèces du complexe MABSC 40
I.A.3.e.ii Epidémiologie des infections associées à *Mabs* 40
I.A.3.e.iii Facteurs de risque 42
I.A.3.e.iv Diagnostic des infections associées à *Mabs* 43
I.A.3.e.v Mécanismes de résistance antibiotique 45
I.A.3.e.vi Facteurs de virulence de *Mabs* 47
Transfert génétique horizontal 47
Systèmes de sécrétion de type VII (ESX) 47
Protéines de type MCE 47
Protéines de transport type MmpL 48
Régulateurs transcriptionnels 48

I.A.3.e.vii Cycle infectieux de *Mabs* 49
Les 2 morphotypes rugueux et lisse de *Mabs* 49
Interaction avec le système immunitaire 51
I.A.3.e.viii Traitement des infections associées à *Mabs* 55
Conduite thérapeutique 55
Nouveaux traitements 56
I.B ORGANISATION DE L'ENVELOPPE MYCOBACTERIENNE 57
I.B.1 LA MEMBRANE PLASMIQUE ET LES LIPOGLYCANNES 58
I.B.1.a LES PHOSPHATIDYL-INOSITOL MANNOSIDES (PIMs) 58
I.B.1.a.i Biosynthèse des PIMs 59
Ajout successif des mannoses 60
Composition de l'ancre lipidique 61
I.B.1.a.ii Rôles biologiques des PIMs 61
I.B.1.b LIPOMANNANE (LM) 62
I.B.1.c LIPOARABINOMANNANE (LAM) 63
I.B.1.d ROLES BIOLOGIQUES DU LM ET DU LAM 64
I.B.2 LE PERIPLASME ET LE COMPLEXE MYCOLYL ARABINOGALACTANE PEPTIDOGLYCANNE (mAGP) 65
I.B.2.a SPECIFICITES DU PEPTIDOGLYCANNE DES MYCOBACTERIES 65
I.B.2.b L'ARABINOGALACTANE (AG) 66
I.B.2.b.i La chaine galactane 66
I.B.2.b.ii La chaine arabinane 67
I.B.2.c ESTERIFICATION PAR LES ACIDES MYCOLIQUES 67
I.B.2.c.i Structure des acides mycoliques 67
I.B.2.c.ii Biosynthèse des acides mycoliques 69
I.B.2.c.iii Liaison à la chaine arabinane 70
I.B.2.c.iv Le dimycolyl diarabino glycerol (DMAG) 70
I.B.3 LA MYCOMEMBRANE ET SES GLYCOLIPIDES 70
I.B.3.a LES GLYCOLIPIDES DERIVES DU TREHALOSE 70
I.B.3.a.i Les acyl-tréhaloses 71
Position et structure des lipides 71
Biosynthèse et transport du di-*O*-acyltréhalose (DAT) 72
Rôles biologiques 72
I.B.3.a.ii Les mycolyl-tréhaloses 73
Tréhalose dimycolate (TDM) et monomycolate (TMM) identifiés selon les espèces 73

Biosynthèse et transport du TMM : découverte du glucose monomycolate (GMM) 73
Rôles biologiques du TDM 74
I.B.3.a.iii Les tréhaloses polyphléates (TPPs) 75
I.B.3.a.iv Les lipooligosaccharides (LOSs) 75
Structures parmi 3 espèces de MNTs et *Mycobacterium tuberculosis* (*Mtb*) 75
Rôles biologiques des LOSs de *Mycobacterium marinum (Mmar)* 76
I.B.3.b LES GLYCOPEPTIDOLIPIDES (GPLs) PORTEURS DE DEOXYHEXOSES 77
I.B.3.b.i *Mycobacterium xenopi* (*Mxen*) : nouveau noyau peptidique 77
I.B.3.b.ii *Mycobacterium avium* (*Mav*) : une grande diversité de structures 77
Groupe 1 : type fucose 78
Groupe 2 : type rhamnose 79
Groupe 3 : type glucose 79
I.B.3.b.iii *Mycobacterium peregrinum* (*Mper*) : absence du 6dTal et présence de sulfate 79
I.B.3.b.iv *Mycobacterium habana* (*Mhab*) : présence d'acides uroniques 80
I.B.3.b.v *Mycobacterium smegmatis* (*Msmeg*) : le GPL modèle 81
I.B.3.b.vi *Mycobacterium abscessus* (*Mabs*) : specificités structurales 83
I.B.3.b.vii Transport des GPLs 85
I.B.3.b.viii Rôles biologiques des GPLs de *Mabs* 86
I.B.3.c LES LIPOPEPTIDES 86
I.B.3.c.i La fortuitine de *Mycobacterium fortuitum* (*Mfor*) 86
I.B.3.c.ii Plusieurs lipopeptides dans le complexe *Mycobacterium avium* (MAC) 87
I.B.3.d LES PHENOLGLYCOLIPIDES (PGLs) 89
I.B.3.d.i Le mycoside A de *Mycobacterium kansasii* (*Mkan*) 90
I.B.3.d.ii Le mycoside B de *Mycobacterium bovis* (*Mbov*) 90
I.B.3.d.iii Le mycoside G de *Mycobacterium marinum* (*Mmar*) 91
I.B.3.d.iv Les PGL de *Mycobacterium leprae* (*Mlep*) 91
II PRESENTATION DES OBJECTIFS 92
III RESULTATS 95
III.A EXTRACTION DES GLYCOLIPIDES 95
III.B CONTROLE DES SOUCHES ET DES PURIFICATIONS PAR CCM 96
III.C PURIFICATION DES EXTRAITS 96
III.D ETUDE DE LA FRACTION APOLAIRE 98
III.D.1 PURIFICATION DE DEUX DERIVES DU TREHALOSE ET D'ACIDES MYCOLIQUES 98
III.D.1.a Analyse MALDI-TOF du TDM natif 100

III.D.1.b Analyse du TDM natif par RMN _____ 101
III.D.1.c Analyse RMN du dérivé di-*O*-acylé natif du TMM _____ 103
III.D.1.d Analyse MALDI-TOF du dérivé di-*O*-acylé natif du TMM _____ 106
III.D.1.e Étude des acides mycoliques de *Mabs*_____ 108
III.D.1.e.i Caractérisation des MAMEs purifiés par RMN_____ 109
Mkan _____ 109
Mabs _____ 110
III.D.1.f Conclusion _____ 113
III.E ETUDE DE LA FRACTION POLAIRE _____ 114
III.E.1 DETERMINATION DE LA COMPOSITION EN MONOSACCHARIDES _____ 114
III.E.2 PURIFICATION DES GLYCOLIPIDES DE LA FRACTION POLAIRE _____ 114
III.E.3 CARACTERISATION DES DI-*O*-ACYL TREHALOSE _____ 116
III.E.3.a Analyse du 3,4 DAT_____ 118
III.E.3.a.i Spectres RMN du 3,4 DAT natif _____ 118
III.E.3.a.ii Spectres MALDI-TOF du 3,4 DAT natif_____ 120
III.E.3.a.iii Composition lipidique du 3,4 DAT natif _____ 120
III.E.3.b Analyse du 2,3 DAT_____ 121
III.E.3.b.i Spectres RMN du 2,3 DAT natif _____ 121
III.E.3.b.ii Spectres MALDI-TOF du 2,3 DAT natif_____ 125
III.E.3.b.iii Composition lipidique du 2,3 DAT _____ 126
III.E.3.c Analyse d'un acyl tréhalose provenant de la souche clinique GS163 ____ 126
III.E.3.c.i Spectres RMN de l'acyl tréhalose natif issu de la souche GS163 ____ 127
III.E.3.c.ii Spectres MALDI-TOF de l'acyl tréhalose natif issud de la souche GS163 129
III.E.3.d Conclusion : 3 acyl tréhaloses identifiés. _____ 130
III.E.4 CARACTERISATION DES PIMs _____ 131
III.E.4.a Purification des PIMs _____ 131
III.E.4.b Spectres MALDI-TOF des PIMs_____ 131
III.E.4.c Composition en monosaccharides des PIMs _____ 135
III.E.4.d Composition lipidique des $AcPIM_2$_____ 137
III.E.4.e Spectres RMN de l'Ac_2PIM_2 _____ 140
III.E.4.f Conclusion : structure résolue de l'Ac_2PIM_2_____ 144
III.E.5 ETUDE DE LA BIOSYNTHESE DES GPLs DE *Mabs* S_____ 145
III.E.5.a Analyse comparative des GPLs de *Mabs* S et du mutant *Δfmt* (1er article) 145

III.E.5.b Etude de la glycosylation des GPLs via l'analyse structurale des GPLs des mutants *Δgtfs* et impact in-vitro sur la primo-infection (2ème article)_____ 189

III.E.5.c Lien entre l'acétylation et la glycosylation des GPLs de Mabs S chez les mutants Δatfs (version soumise du 3ème article) _____ 253

III.E.5.d Conclusion _____ 253

III.E.6 IDENTIFICATION D'UN NOUVEAU GPL CHEZ *Mabs* R _____ 318

III.E.6.a Purification orthogonale du GPL4690 _____ 318

III.E.6.a.i Purification sur colonne de silice _____ 318

III.E.6.a.ii Purification sur colonne de silice greffée C_{18}_____ 319

III.E.6.b Analyse FT-ICR-MS (Fourier Transform Ion Cyclotron Resonance-Mass spectrometry) 321

III.E.6.c Analyse RMN du GPL4690 _____ 324

III.E.6.c.i Spectres dans les solvants deutérés _____ 324

Identification des acides aminés _____ 324

Séquençage du peptide_____ 328

Identification des monosaccharides _____ 330

Liaisons glycosidiques _____ 331

Autres observations_____ 331

III.E.6.c.ii Spectres dans les solvants protiques _____ 333

Séquence peptidique_____ 333

Fonctions alcools _____ 335

III.E.6.c.iii Conclusion _____ 336

III.E.6.d Analyse par MALDI Q-IT TOF _____ 337

III.E.6.d.i Spectres MS des fractions natives _____ 337

III.E.6.d.ii Spectres de fragmentation MS^n des fractions natives _____ 337

III.E.6.d.iii Spectres des fractions après traitement alcalin _____ 340

III.E.6.d.iv Spectres MS^n des fractions après traitement alcalin _____ 340

III.E.6.d.v Conclusion _____ 343

III.E.6.e Détermination de la composition en monosaccharides et en lipides ____ 343

III.E.6.f Conclusion sur les structures identifiées _____ 345

III.E.7 RECHERCHE D'UNE CORRELATION AVEC LE GENE *MAB_4690c* _____ 346

III.E.7.a Analyse bioinformatique des gènes autour de *MAB_4690c* _____ 346

III.E.7.a.i Prédiction des substrats ajoutés par *MAB_4690c* et *MAB_4691c* ___ 346

III.E.7.a.ii Recherche des rhamnosyltransférases potentielles _____ 348

III.E.7.a.iii Alignement avec les autres gènes du locus de biosynthèse des GPLs 348

III.E.7.a.iv Conclusion 349
III.E.7.b Production du GPL4690 par *Mabs* et les mutants NRPS 350
III.E.7.b.i Séparation CCM 350
III.E.7.b.ii Séparation par chromatographie-flash puis détection MALDI-TOF 351
III.E.7.b.iii Analyses LC-MS 353
III.E.7.c Conclusion 356
IV DISCUSSION 357
V MATERIELS ET METHODES 367
V.A EXTRACTION DES GLYCOLIPIDES 367
V.B Chromatographie couche mince (CCM) 367
V.B.1 Analytique 367
V.B.2 Préparative 368
V.B.3 Révélation 368
V.C Purification des glycolipides par chromatographie-flash 368
V.C.1 Séparation par adsorption 368
V.C.2 Séparation phase inverse 369
V.D Dérivation chimique 369
V.D.1 Saponification des glycolipides (β-élimination) 369
V.D.2 Méthanolyse (transestérification) 370
V.D.3 Hydrolyse alcaline suivie d'une méthyl-estérification 370
V.D.4 Obtention des dérivés itols-acétates des monosaccharides 370
V.D.5 Obtention des dérivés perméthylés du tréhalose 370
V.E CHROMATOGRAPHIE GAZ 371
V.E.1 Dérivés itols-acétates 371
V.E.2 FAMEs et dérivés triméthylsilylés 371
V.F SPECTROMETRIE DE MASSE 371
V.F.1 MALDI QIT-TOF 371
V.F.2 MALDI et ESI FT-ICR 372
V.F.3 ESI-Ion Trap 372
V.G Résonance Magnétique Nucléaire 372
V.H Couplage à des billes fluorescentes 373
VI BIBLIOGRAPHIE 374

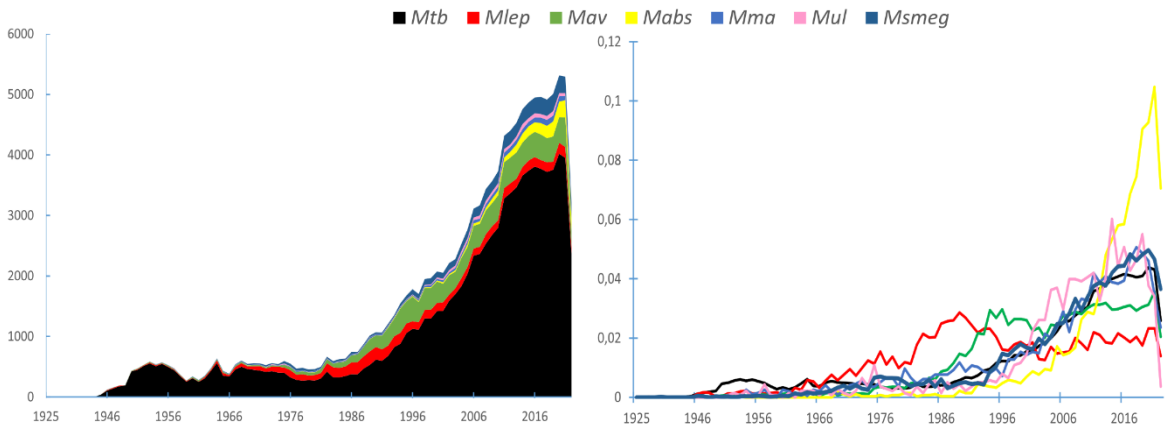
I INTRODUCTION

Tout le monde s'accorde sur le fait que l'épidémie mondiale liée au SARS-COV2 a profondément transformé notre civilisation. S'agissant de la santé publique, l'augmentation rapide de la mortalité chez les personnes à risque, a contraint de nombreux pays à isoler les individus malades et sains pour contrer la contagiosité du virus. L'un des effets collatéraux de cette stratégie fut de limiter l'interaction de l'ensemble des patients suivis par les personnels soignants. En outre, ces derniers redoublaient d'efforts pour contenir la pandémie : nous étions en réaction face à la menace.

Récemment, le rapport 2022 de l'Organisation Mondiale de la Santé (OMS) a montré que le nombre de décès liés à la tuberculose a connu sa 1ère augmentation de 4% depuis 20 ans, avec plus de 1,5 millions de morts[1]. En 2000, cette pathologie était la 7ème maladie la plus meurtrière avec 3,4% d'incidence alors qu'en 2019, elle redescendait à la 13ème place avec 2,2% d'incidence. Cette résurgence succède à une diminution significative du nombre de cas diagnostiqués en 2020 en raison de l'isolement des personnes. Cela démontre que le manque de prévention, le défaut de diagnostic et la prise en charge trop tardive ont impacté défavorablement la survie des populations. En 2020 et 2021, la tuberculose serait la 2ème cause de mortalité liée à un agent infectieux après le SARS-COV2, mais depuis sa découverte à la fin du 19e siècle, il est évident qu'elle le dépasse largement en valeur absolue.

De façon surprenante, le principal facteur de risque lié à la tuberculose est la malnutrition, et le financement mondial pour soutenir les actions destinées à éradiquer cette maladie diminue depuis 2018. C'est pourquoi, certains suggèrent que les pays riches et développés ont préféré dédier leurs investissements à la production d'un vaccin contre le SARS-COV2, alors que le développement de vaccins plus efficaces que le BCG contre la tuberculose est également nécessaire[2]. De plus, des souches résistantes aux antibiotiques tels que la rifampicine, l'isoniazide ou la fluoroquinolone sont de plus en plus fréquemment isolées.

Comme le SARS-COV2 a occulté la tuberculose, il ne faudrait pas que cette dernière occulte d'autres infections mycobactériennes qui pourraient avoir une incidence marquée sur la santé humaine, notamment dans les pays développés. Parmi elles, *Mycobacterium abscessus* (*Mabs*) fait l'objet d'une attention particulière comme le montre l'évolution bibliométrique exponentielle des articles publiés à son sujet.



Graphiques 1 et 2 : Evolution bibliométrique des articles traitant des espèces mycobactériennes majeures. Nombre d'articles en valeur absolue (à gauche) et en proportion de l'espèce mycobactérienne considérée (à droite) à partir d'une requête sur le moteur de recherche www.pubmed.com en août 2022 (*Mtb = tuberculosis ; Mlep = leprae ; Mav = avium ; Mabs = abscessus ; Mma=marinum ; Mul=ulcerans ; Msmeg=smegmatis*). Parmi les espèces pathogènes, *Mtb* est la plus représentée en valeur absolue et *Mabs* révèle la plus forte progression relative depuis 2016.

Cette espèce découverte depuis moins d'un siècle, possède le plus vaste arsenal de résistance antibiotique et la capacité de persister dans des niches écologiques proches de l'homme. Avec les autres mycobactéries non tuberculeuses (MNTs), elle fait l'objet d'une surveillance accrue de la communauté scientifique, notamment chez les patients atteints de mucoviscidose davantage exposés aux infections bactériennes émergentes[3].

Malgré cet engouement, la description des composés non protéiques présents dans l'enveloppe et susceptibles d'interagir avec son hôte n'a jamais fait l'objet d'études aussi complètes que pour l'espèce *Mycobacterium tuberculosis* (*Mtb*)[4]. C'est pourquoi, l'objectif de ce travail de thèse est de décrire le glycolipidome propre à la souche de référence ATCC19977 de *Mabs*, et d'évaluer comment il contribue à sa pathogénicité.

Après un rappel sur la classification taxonomique, la description microbiologique des mycobactéries montrant l'évolution croissante du nombre d'espèces et les milieux qu'elles sont capables de coloniser, j'aborderai le groupe des mycobactéries non tuberculeuses (MNTs) auquel *Mabs* appartient. Leur description épidémiologique face à la tuberculose ainsi que leur impact économique montreront que les pays développés ne sont pas épargnés. Parmi ces MNTs, la description exhaustive des 3 sous-espèces du complexe *Mycobacterium abscessus* (MABSC) révèlera la présence de 3 clones dominants circulants à travers le monde, les nombreux facteurs de risque et de virulence de ce pathogène opportuniste ainsi que le panel de mécanismes de résistance antibiotique qui complique le choix des traitements.

Pour une meilleure compréhension du cycle infectieux de *Mabs* et de son interaction avec notre système immunitaire, nous reviendrons ensuite sur la structure détaillée de la paroi mycobactérienne et de son contenu glycolipidique dont de nombreux membres sont réputés avoir un effet immunomodulateur dépendant de leur structure chimique. Après avoir revu la biosynthèse et les rôles biologiques des lipoglycannes retrouvés majoritairement dans la membrane plasmique, nous décrirons les particularités chez les mycobactéries d'une macromolécule commune à de nombreuses autres bactéries, le peptidoglycanne. Enfin, nous terminerons cette description en listant l'ensemble des structures chimiques des glycolipides libres retrouvés parmi toutes les espèces de mycobactéries, en distinguant ceux assemblés autour du noyau tréhalose, ceux constitués de déoxyhexoses, ceux dépourvus de monosaccharide et ceux construits autour du noyau phénol.

I.A LES MYCOBACTERIES QUI NOUS ENTOURENT

I.A.1 CLASSIFICATION TAXONOMIQUE DU GENRE MYCOBACTERIUM

Etymologiquement, le terme mycobactérie vient du grec *mykes*, qui signifie « champignon », et *bakterion*, qui veut dire « petit bâton ». Finalement écarté du règne des champignons, cet organisme est maintenant pleinement reconnu comme une bactérie. Dès les premières expériences de coloration et d'observation microscopique des bacilles de Hansen et de Koch, en 1873 et 1882, la forme en bâtonnet des mycobactéries est décrite. Les agents étiologiques de la lèpre et de la tuberculose, *Mycobacterium leprae* et *Mycobacterium tuberculosis* ont dans un premier temps été largement étudiées du fait de leur incidence majeure sur la santé publique. Après elles, des mycobactéries saprophytes non pathogènes, *Mycobacterium smegmatis*[5] (*Msmeg*) et *Mycobacterium phlei*[6] (*Mph*) ainsi que des mycobactéries pathogènes chez les oiseaux *Mycobacterium avium*[7] (*Mav*), les rats *Mycobacterium lepraemurium*[8], les ovins *Mycobacterium paratuberculosis*[9] (*Map*) ou les tortues *Mycobacterium chelonae*[5] (*Mche*) furent caractérisées au début du 19ème siècle. En 1926, Aronson étudie les tubercules prélevés sur des poissons d'eau de mer dans la région de Philadelphie. Il parvient à isoler des bacilles, à les mettre en culture et à infecter des pigeons, des souris, des grenouilles et des poissons rouges. Il caractérise ainsi *Mycobacterium marinum*[10] (*Mmar*) et jette les bases d'une systématisation de la description des mycobactéries. Dans les années 1950, après la description de pathogènes opportunistes tels que *Mycobacterium ulcerans*[11] (*Mul*) et *Mycobacterium abscessus*[12] (*Mabs*), le besoin de classer le nombre croissant d'espèces se distinguant de *Mycobacterium tuberculosis* (*Mtb*) s'exprime à travers une classification dédiée, la classification de Runyon.

I.A.1.a CLASSIFICATION DE RUNYON

La classification de Runyon se base sur l'étude de souches cliniques prélevées sur des patients américains[13]. Ces souches diffèrent de *Mtb* du fait de plusieurs critères :

- leur résistance antibiotique vis-à-vis de l'isoniazide ou de la streptomycine ;
- leur activité catalase plutôt élevée ;
- la vitesse de leur croissance plutôt lente à température ambiante ;
- leur incapacité à infecter mortellement le cobaye ;
- la pigmentation des colonies qui possèdent une morphologie souvent lisse plutôt que rugueuse.

Ces mycobactéries « atypiques », terme plus tard contesté par Runyon lui-même[14], sont réparties en 4 groupes du fait de leur phénotype :

- Le groupe I des photochromogènes : elles produisent une pigmentation jaune pâle jusqu'à orange intense uniquement à la lumière ; les colonies sont visibles après 3 à 4 semaines suite à un ensemencement léger ; leur morphologie est généralement lisse.

- Le groupe II des scotochromogènes : elles sont colorées quelle que soit la luminosité ; la croissance et la morphologie des colonies est semblable au groupe I ; elles ne provoquent pas de symptômes pathologiques chez l'animal.

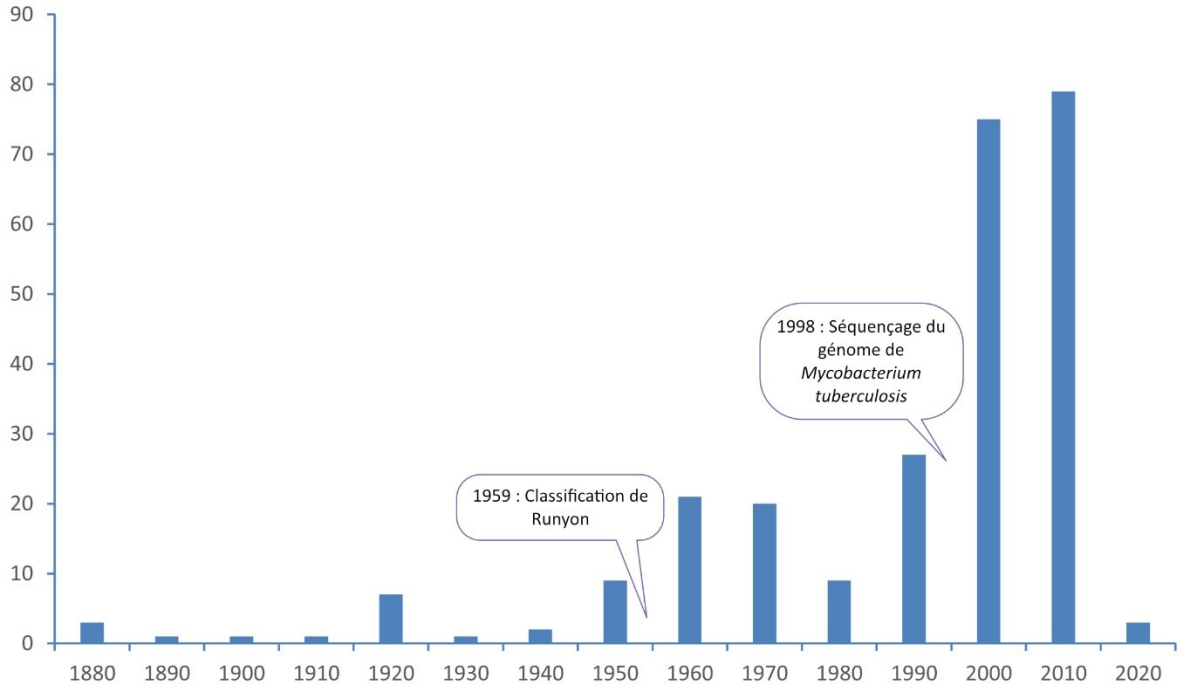
- Le groupe III des non chromogènes : elles ne produisent aucune pigmentation ; la croissance est possible au-delà de 45°C ; les colonies sont lisses et ne forment généralement pas de cordes ; la pathogénicité est variable chez le cobaye.

- Le groupe IV des espèces à croissance rapide : elles produisent des colonies visibles après 48h à 20-25°C à partir d'un inoculum invisible ; c'est le cas de l'espèce *Mycobacterium fortuitum (Mfor)*, découverte en 1928, dont l'inoculation chez la souris est létale.

Parmi ces mycobactéries non tuberculeuses (MNTs), *Mabs* ainsi que certaines espèces déjà proposées ne sont pas reprises dans les travaux publiés en 1959. Dans la décennie suivante, Bojalil ajoute des critères à la classification de Runyon[15] : la température optimale de croissance et la survie des colonies au-delà de 60°C, l'utilisation d'acides organiques ou de monosaccharides comme source de carbone, la production d'acide selon le monosaccharide présent dans le milieu de culture. Tous les critères avaient un poids égal et permettaient de classer 14 espèces. C'est dans l'étude de Tsukamura en 1967 que *Mabs* réapparait. La source d'azote du milieu de culture ou l'activité catabolique enrichissent les précédents critères de sélection. A partir des 95 caractéristiques retenues, la souche type ATCC 19977 est déclarée identique à une souche de *Mycobacterium runyonii* qu'elle remplace en tant que sous-espèce de *Mfor*[16].

La description phénotypique détaillée des MNTs, de leurs conditions de culture et de leur pathogénicité aboutit ainsi à décrire plus d'une quarantaine de nouvelles espèces[17]. Au début des années 1970, les techniques d'analyse de l'ADN ainsi que des lipides ont confirmé la distinction, entre le genre *Mycobacterium* et le genre *Corynebacterium* d'une part, en raison de leur proportion G+C d'au moins 61%[18], et avec le genre *Nocardia* d'autre part, en raison d'une plus grande longueur des chaines d'acides gras[19]. En effet, la description des acides mycoliques de l'enveloppe des

mycobactéries par les techniques de chromatographie fait partie de l'arsenal utilisé pour parfaire leur classification[20,21], voire confirmer leur identification[22]. Certains auteurs proposent également que la mycobactine, retrouvée parmi de nombreuses mycobactéries à croissance rapide, produite pour la captation du fer, soit un outil de classification taxonomique[23]. Néanmoins, à la fin des années 1980, les techniques de biologie moléculaire sont admises pour décrire la phylogénie du règne bactérien à partir de l'ensemble du génome, et les mycobactéries ne font pas exception à la règle.



Graphique 3 : Nombre d'espèces mycobactériennes nouvelles proposées par décennie depuis la 1ère description jusqu'à nos jours (source « DSMZ German collection of microorganisms »).

I.A.1.b CLASSIFICATION SELON LE GENOME

En 1987, un comité d'experts se réunit pour réconcilier les approches de systématique bactérienne[24]. Il consacre les techniques d'appariement de l'ADN comme référence pour déterminer les espèces nouvelles. Ainsi, 2 souches possédant plus de 70% d'ADN apparenté sont membres d'une même espèce. Néanmoins, les phylogénies déduites doivent montrer une cohérence phénotypique. L'hybridation ADN avait été utilisé avant les recommandations du comité notamment pour rassembler *Mtb* et *Mycobacterium bovis* (*Mbov*) au sein d'une même espèce, répartir les sérotypes de *Mav* et *Mycobacterium intracellulare* en 2 groupes ainsi que pour mettre en évidence l'éloignement entre les génomes des MNTs et les espèces du complexe *Mtb*[25,26].

Dans les années 1990, le séquençage de l'ARN ribosomal, plus rapide que les techniques biochimiques, et la comparaison de cette ARN entre espèces devient la technique privilégiée pour classer les taxons. Par exemple, en 1992, Pitulle tire profit du séquençage de l'ARN ribosomal 16S de 9 espèces de mycobactéries à croissance rapide, et de leur alignement avec l'ensemble des ARN 16S déjà séquencés de 35 espèces, pour construire un dendrogramme phylogénétique[27]. Ce dernier rendra compte de la séparation entre les mycobactéries à croissance lente et celles à croissance rapide, plus proches d'un ancêtre commun. Le gène codant pour l'ARN 16S se révèle parfois insuffisant pour distinguer des espèces trop proches. D'autres gènes s'avèrent ainsi nécessaires pour assurer une distinction plus spécifique. Il s'agit notamment du gène *hsp65* codant pour la protéine de choc thermique HSP65, ou du gène *rpoB* codant la sous-unité β de l'ARN polymérase. Dans le cas du gène *hsp65*, les tentatives d'identification d'isolats cliniques à partir d'une base de données dévoilèrent que ce dernier n'était pas strictement discriminant, notamment pour les espèces du complexe *Mche*[28], en raison du manque d'exhaustivité des bases de données. D'autre part, pour le gène *rpoB*, le choix d'une région hypervariable de 723 bp apporte une spécificité supérieure à la séquence de l'ARN 16S dans le cas des mycobactéries à croissance rapide[29]. D'autres séquences d'ADN, codantes ou non, parfois concaténées, seront utilisées pour classer les espèces mycobactériennes. En 2014, Tortoli admet que l'évolution plus récente des mycobactéries à croissance lente par rapport à celles ayant une croissance rapide est confirmée par toutes les méthodologies[30]. Afin de parfaire la classification de plus d'une centaine d'espèces (Graphique 1), la comparaison des génomes s'enrichit de nouvelles stratégies d'alignement.

Suite au premier séquençage complet d'un génome bactérien (*Haemophilus influenzae*) en 1995, de multiples projets se mettent en place pour dévoiler l'intégralité des génomes des procaryotes, et proposer des règles de reconstitution et d'annotation des génomes grâce à l'apport concomitant des outils de bioinformatique[31]. L'ANI (Average Nucleotide Identity) est progressivement adopté pour comparer la totalité des génomes de deux souches bactériennes. Selon une proposition de Goris, deux souches appartiennent à la même espèce si l'ANI dépasse 95%, à relier à la valeur de 70% de similitude obtenue à partir d'ADN hybridés (DDH). Par contre, elles sont considérées comme sous-espèces si l'ANI est compris entre 95 et 98%[32]. En outre, une étude incluant *Mtb* et *Mbov* révèle que le seuil d'identité est de 98,65% dans le cas de l'ARN 16S pour distinguer deux espèces[33].

Tortoli proposa une étude exhaustive de 148 génomes mycobactériens dont il compara les ANI pour construire 2 arbres phylogénétiques[34]. Son travail confirme que les mycobactéries à croissance lente dérivent plus tardivement que les mycobactéries à croissance rapide de leur ancêtre commun.

Le complexe *Mycobacterium terrae* se révèle comme le groupe le plus ancien parmi les mycobactéries à croissance lente, et le complexe *Mycobacterium abscessus-chelonae* (MABSC) est le plus ancien des mycobactéries à croissance rapide (Figure 1). Les espèces pathogènes strictes telles que *Mtb* et *Mycobacterium leprae* (*Mlep*) se retrouvent parmi les mycobactéries à croissance lente. Néanmoins, Tortoli s'accorde une exception à la réunion des espèces sur la base de l'ANI puisque *Mmar* et *Mul* possède des valeurs supérieures à 98%, mais elles restent séparées dans un même complexe du fait de la spécificité des pathologies observés chez l'homme et de leurs phénotypes chromogéniques.

Récemment, Gupta a proposé une étude phylogénétique qui se base sur les signatures moléculaires relatives aux protéines déduites *in-silico* à partir de 150 génomes[35]. C'est une approche qui présente l'avantage d'identifier les groupes monophylétiques et paraphylétiques, mais pas les groupes polyphylétiques. Cette classification préserve la frontière entre mycobactéries à croissance rapide et lente, et définit 5 clades nouvellement nommés et dominés par 2 larges groupes (Figure 1) :

- Le clade historique *Mycobacterium* représenté par les espèces du clade « *Mycobacterium tuberculosis-simiae* » est le plus grand et inclut également *Mlep*, *Mmar* et *Mul* ;
- Le clade *Mycolicibacter* avec les espèces du complexe *Mycobacterium terrae* ;
- Le clade *Mycolicibacillus* avec les 2 espèces *Mycobacterium triviale* et *Mycobacterium koreense* ;
- Le clade *Mycolicibacterium* englobant un très grand nombre de mycobactéries à croissance rapide dont l'espèce *Mfor* ;
- Le clade *Mycobacteroides* avec notamment *Mabs* et *Mche*.

Grâce à des méthodes comparatives toujours basées sur l'analyse du génome, cette division en 5 clades du genre *Mycobacterium* a été remise en question récemment[36]. En effet, en se basant sur plusieurs marqueurs dont l'ARN 16S, l'identité des acides aminés (AAI), le pourcentage de protéines conservées (POCP) ou l'ANI, la similarité des espèces à l'intérieur des 4 nouveaux clades proposés est très, voire trop, supérieure à celle des espèces du genre *Mycobacterium*. De plus, la taille du clade *Mycolicibacillus* n'est pas significativement pertinente et les auteurs préfèrent envisager la possibilité que ces clades deviennent des sous-genres. A l'aube de notre siècle, au sein de la base de données microbiologique DSM, 259 espèces de mycobactéries ont été déposées, 49 d'entre elles ont été rejetées dont 36 depuis l'année 2000 (Graphique 3). La pléiade de techniques d'identification des mycobactéries offre des opportunités de découvrir de nouvelles souches, dont la comparaison incessante consolidera la définition de nouvelles espèces et affinera leurs limites taxonomiques.

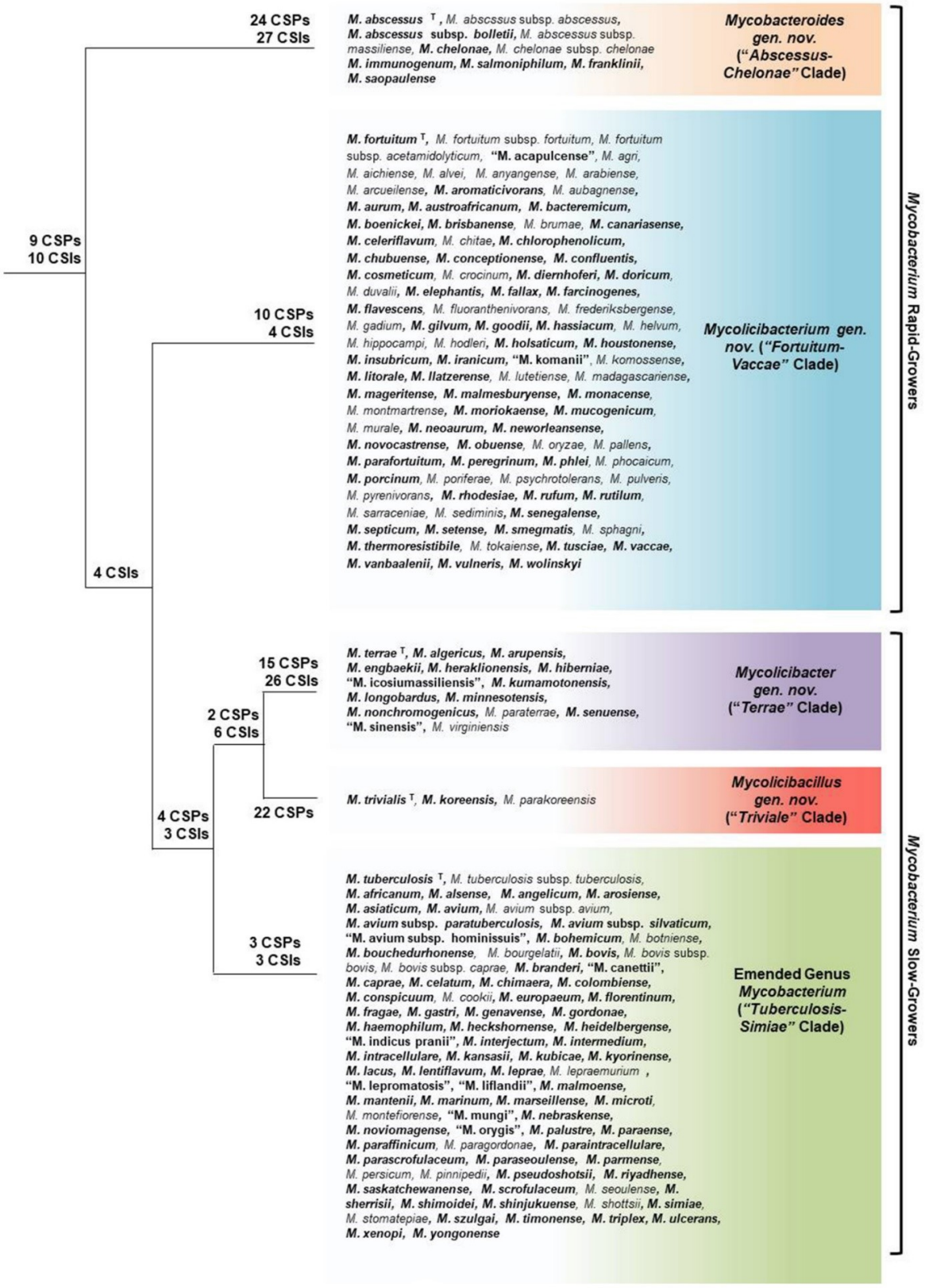


Figure 1 : Répartition phylogénétique des 5 clades du genre Mycobacterium d'après Gupta[35]. Les CSPs sont des homologues protéiques retrouvés uniquement dans les espèces apparentées, alors que les CSIs sont des insertions ou des délétions d'acides aminés de longueur fixe, présentes à une position spécifique d'une région protéique conservée.

I.A.2 DESCRIPTION MICROBIOLOGIQUE DU GENRE MYCOBACTERIUM

Les mycobactéries sont des microorganismes unicellulaires procaryotes appartenant au règne bactérien classés parmi les actinobactéries (ou actinomycètes), ordre des corynébactériales (ou mycobactériales), famille des mycobactériacées et regroupées à minima dans le genre *Mycobacterium*[37].

Les mycobactéries possèdent un seul chromosome. La taille du génome de *Mtb* est de 4,4 Mb ; celui de *Mlep* est le plus court avec 3,2 Mb et celui de *Mycobacterium gordonae* le plus long avec 7,6 Mb[38]. Leur contenu en G+C est compris entre 57 et 73%[17,18]. Les mycobactéries possèdent parfois des plasmides dont l'un d'eux est à l'origine de la production de mycolactone chez *Mul*[39]. Plusieurs furent détectés parmi les sérovars de *Mav*[40]. Par contre, les tentatives d'en caractériser un chez *Mtb* n'ont pas abouti[41].

I.A.2.a MORPHOLOGIE, CULTURE ET CROISSANCE

L'observation au microscope électronique[42] montre des cellules cylindriques de 0,5 µm de diamètre et d'une longueur allant de 1 à 4 µm. Elles ne possèdent ni flagelles, ni cils, et ne produisent pas de spores. Par contre, les mycobactéries possèdent une enveloppe très épaisse dont les lipides représentent 40% de la matière sèche[43].

Les mycobactéries sont des bactéries essentiellement aérobies, mais le cycle infectieux des espèces pathogènes les oblige à s'adapter à l'hypoxie au sein de l'hôte pour produire leur énergie[44]. Dans ce cas, elles se comportent comme des microaérophiles. La culture se pratique sur milieu solide ou liquide. Un milieu nutritif riche et économique de type Lowenstein-Jensen est préparé à partir d'œufs pour la culture sur milieu solide, alors que le milieu de type Middlebrook est utilisé dans les 2 cas[45]. Hétérotrophes, les mycobactéries sont capables de métaboliser plusieurs sources de carbone organique : le glycérol, les acides organiques, les monosaccharides, les polyols ou les lipides. Ces derniers constituent une source privilégiée au sein de l'hôte. La source d'azote peut aussi bien être un acide aminé que la forme minérale ammoniacale même si la croissance est plus lente sur ce dernier substrat[46].

Sur milieu solide, l'observation du caractère morphologique lisse ou rugueux est facilitée. Le premier correspond à une surface plane, luisante et régulière qui semble adhérente alors que la texture du second apparaît plus irrégulière avec une apparence dite « en chou-fleur ». En milieu liquide, la turbidité et la capacité à former des cordes sont évalués[47]. En 2003, Tu et ses collaborateurs

observent plus de 11000 échantillons par microscopie photonique, et ils déterminent 6 structures morphologiques (Figure 2) dans lesquelles les mycobactéries sont isolées ou agrégées entre elles. Dans l'une d'entre elles, les bacilles s'agglutinent dans le sens longitudinal pour former des cordes serpentiformes. C'est la principale structure observée (93%) chez *Mtb* contrairement à *Mav*. Dans le but de relier la formation des cordes au niveau microscopique à l'aspect macroscopique rugueux des colonies, 6 espèces de MNTs à croissance rapide ont été observées par microscopie confocale laser et microscopie électronique[48]. Les auteurs parviennent à établir que l'aspect plus ou moins rugueux des colonies est lié à la profondeur des irrégularités mesurées par la topographie 3D. Ils montrent aussi que toutes les espèces ont formé des cordes, mais ils ne confirment pas un lien direct avec l'aspect rugueux. D'autre part, ils s'étonnent que les espèces non pathogènes soient capables de former des cordes, sachant que ce caractère est reconnu comme contributeur de la virulence de *Mtb* ou de *Mabs*.

Les mycobactéries à croissance rapide forment des colonies visibles au bout de 7 jours contrairement aux mycobactéries à croissance lente. L'automatisation des cultures en milieu liquide permet de mesurer continuellement cette vitesse de croissance par des méthodes radiométriques ou fluorescentes[45]. D'autre part, un grand nombre d'alternatives de culture sur milieu liquide et solide ont été développées également dans le but de tester la susceptibilité aux antibiotiques[49]. Les mycobactéries sont plutôt mésophiles : la température de croissance est de 37°C mais elle est diminuée à 30°C pour la détection des espèces associées aux organismes ectothermes (*Mmar*, *Mul*, *Mche*…) et à 42°C pour des espèces thermophiles (*Mycobacterium xenopi*, *Mycobacterium hassicum*).

I.A.2.b DETECTION DES MYCOBACTERIES

I.A.2.b.i Techniques microscopiques

A la fin du 19ème siècle, les techniques d'observation de microscopie photonique sont associées à des protocoles de coloration afin de mettre en évidence les microbes pathogènes. Par exemple, la coloration de Gram est utilisée pour distinguer les bactéries Gram positives dont l'épaisse couche de peptidoglycanne retient le violet de gentiane, des bactéries Gram négatives qui seront décolorées puis recolorées par la fuschine. En raison de la plus grande complexité de leur enveloppe, les mycobactéries ont nécessité la mise au point d'une procédure spécifique appelée coloration de Ziehl-Neelsen et optimisé dès les années 1940 pour faciliter leur observation[50] :

- D'abord, une coloration par la fuschine est faite à chaud en présence de sel ;
- Ensuite, 2 étapes de décoloration par un acide, puis un alcool se succèdent ;
- Enfin, une coloration de contraste par le bleu de méthylène termine le protocole.

Excepté le chauffage, cette succession d'étapes a aussi été pratiqué dans la coloration du bacille de la diphtérie par Kinyoun en 1914[51]. La similitude des enveloppes de *Corynebacterium* et de *Mycobacterium* expliquent aujourd'hui que les deux protocoles furent équivalents. L'utilisation d'un colorant fluorescent, l'auramine, permet d'améliorer la sensibilité de détection des bacilles. Une revue systématique réalisée en 2006 par Steingart confirme que la sensibilité, variant entre 50 et 95%, est meilleure avec l'auramine sans impact sur la spécificité[52]. Malgré l'ancienneté de la coloration de Ziehl-Neelsen, cette dernière demeure préconisée comme un outil diagnostic de 1ère intention dans les pays en voie de développement pour lesquels la simplicité, le déploiement et le coût facilitent la prise en charge de nombreux patients[53].

Dans les services hospitaliers d'anatomopathologie, l'immunohistochimie fait également partie des techniques microscopiques applicables sur les fluides et les tissus. Récemment, Crothers a montré que l'utilisation d'un anticorps polyclonal sur 267 cas cliniques procure une sensibilité supérieure à la coloration de Ziehl-Neelsen, dans sa version non améliorée[54]. Malgré tout, les colorations ne sont pas propres aux mycobactéries, et s'appliquent à d'autres bacilles acido-alcoolo résistants du genre *Corynebacterium* et *Nocardia*. C'est pourquoi, il s'est avéré indispensable de rechercher des critères plus spécifiques.

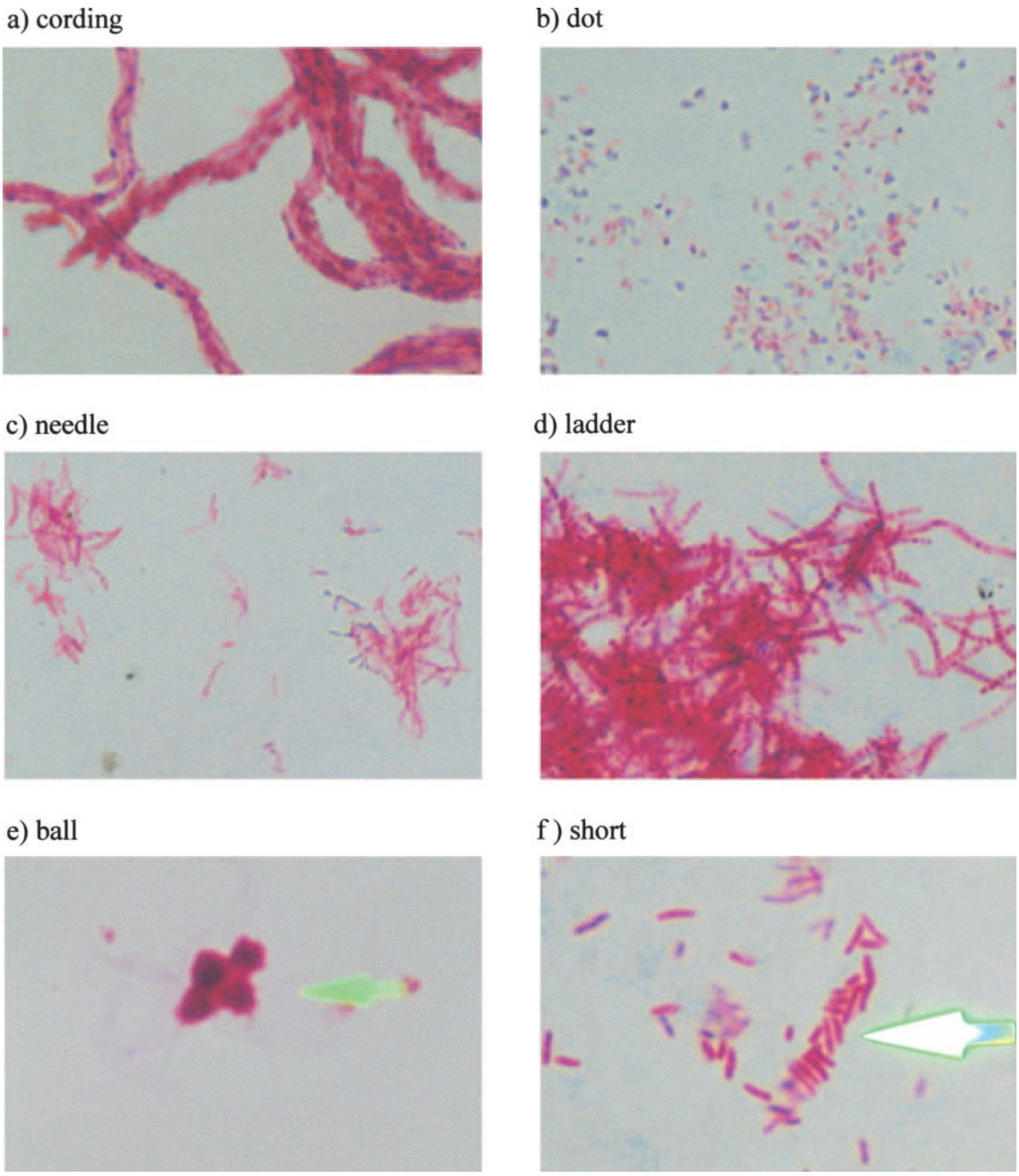


Figure 2 : Coloration de Kinyoun de bacilles mycobactériens montrant leur organisation dans 6 structures macroscopiques : cordes (a), bacilles isolés (b), aiguilles (c), échelle (d), agglomérats courts sphérique (e) et longitudinale (f) d'après Tu[47].

I.A.2.b.ii Techniques microbiologiques

Bien qu'elles aient servies pendant plus d'un siècle à leur classification, les méthodes de détection des mycobactéries basées sur leur culture requièrent les contraintes de manipulation en laboratoire de type PII ou PIII ainsi qu'une attente de plusieurs semaines avant l'obtention de résultats

certains[55]. Par exemple, la durée du test d'activité arylsulfatase varie de 3 à 14 jours selon que l'espèce mycobactérienne possède une vitesse de croissance rapide ou lente. Il en est de même pour la plupart des tests d'activité (catalase, nitrate reductase, pyrazinamidase…) dont la compilation garantira la spécificité après 8 semaines de travail. C'est pourquoi, elles sont progressivement reléguées dans le domaine au profit de méthodes plus rapides et plus précises[56], mais aussi moins coûteuses et moins complexes. La mise en culture est malgré tout inévitable pour obtenir une quantité suffisante de matériel biologique à partir des prélèvements cliniques. De plus, le renseignement des caractères phénotypiques de base apportera un premier niveau de spécificité à l'identification par les techniques qui suivent.

I.A.2.b.iii Techniques biochimiques

Les techniques d'analyse biochimique ont été envisagées pour identifier les mycobactéries en même temps qu'elles enrichissaient leur description taxonomique. Il s'agissait de profiter de l'originalité structurale grandissante des lipides découverts pour distinguer plus facilement les espèces entre elles. La chromatographie phase gazeuse (GC) fut la 1ère technique automatique envisagée mais elle ne parvint qu'à identifier correctement 63% d'isolats parmi 1077 testés[57]. Aujourd'hui, une solution technique de séparation des esters méthyliques d'acides mycoliques par chromatographie liquide haute performance (HPLC) est disponible pour l'identification de *Mtb* et d'une trentaine de MNTs[58]. Les performances supérieures de cette solution atteignent un taux d'identification de 75% sur 370 isolats, avant réévaluation des données. Par contre, elle ne distingue pas les espèces du complexe MABSC.

Dans ce cas, la spectrométrie de masse MALDI-TOF pourrait se montrer plus adaptée car elle permet de différencier les 3 sous-espèces de *Mabs* à partir de leur contenu en glycopeptidolipides (GPLs)[59]. En outre, l'alimentation des bases de données et la préparation préalable à l'analyse MALDI-TOF d'échantillons cliniques se doivent d'être poursuivies et standardisées pour l'améliorer[60,61]. Le travail bibliométrique de Cao sur cette approche rapporte un taux de 85% d'identification correcte du genre et de 71% pour l'espèce parmi 2593 isolats étudiés dans 19 articles[62]. Une 1ère étude multicentrique impliquant 15 laboratoires dans 9 pays européens a été publiée en 2022. Elle laisse entrevoir une proportion supérieure à 95% d'identifications valides[63]. Néanmoins, pour justifier de leur concordance, l'ensemble des travaux précédemment cités s'est appuyé sur les méthodes de référence caractérisant l'espèce mycobactérienne grâce à son génome.

I.A.2.b.iv Techniques de biologie moléculaire

Parce qu'elles présentent l'avantage d'être extrêmement spécifique et que les étapes d'amplification améliorent leur sensibilité, les techniques d'analyse de l'ADN ont bénéficié d'un fort développement technologique. Elles sont ainsi devenues les plus performantes et les plus recommandées pour l'identification de microorganismes.

Sur le marché du diagnostic médical, il existe au moins 3 tests fréquemment rencontrés pour rechercher les mycobactéries[64]. Le premier test, AccuProbe™ (Hologic Genprobe), est fondé sur l'hybridation d'une sonde reconnaissant le gène de l'ARN 16S à partir de culture pure. Il ne nécessite pas d'amplification préalable par PCR. Il existe des versions dédiées de ce test contenant les sondes adaptées aux MNTs. Le test INNOLiPA™ (Fujirebio) cible la région non codante ITS1 située entre les gènes de l'ARN 16S et 23S et donne également accès aux MNTs. Enfin, le test Genotype™ (Hain Lifescience) d'hybridation inversée cible l'ARN 23S ainsi que le gène *gyrB* pour différencier les espèces du complexe *Mtb*. Il nécessite une amplification PCR après l'extraction de l'ADN et existe également dans une version dédiée aux MNTs. Chaque sonde n'étant pas pleinement spécifique, l'utilisation de plusieurs d'entre elles peut affiner la caractérisation.

Des méthodologies de séquençage ciblé autorisent également une bonne spécificité sans pour autant s'intéresser à l'ensemble du génome. Dans un travail datant de 2003, Xiang définit une séquence d'amorce correspondant à une portion du gène *rrs* fortement conservé parmi les bacilles acido-alcoolo résistants. Après amplification de l'ADN mycobactérien à partir de cette amorce, il entame alors le séquençage de la zone hypervariable du gène codant l'ARN 16S. En moins de 5 jours, il peut ainsi confirmer une identification de l'espèce par comparaison avec les génomes des souches types déposés auprès de la GenBank[65].

L'analyse du polymorphisme de séquences génétiques ciblées par l'action d'enzymes de restriction (RFLP), la détection multiple de petites séquences oligonucléotidiques répétitives (spoligotyping), l'analyse de séquences d'ADN aléatoires dispersées dans le génome mycobactérien (MIRU-VNTR) font partie des techniques de typage moléculaire également basées sur le génome. Elles sont compilées dans une revue de 2016 qui se concentre sur l'espèce *Mtb* et l'ensemble des souches de ce complexe, et dans laquelle le séquençage complet du génome est considéré comme la technique la plus discriminante mais aussi la plus coûteuse[66].

Le choix d'une technique de détection doit donc tenir compte du nombre d'échantillons microbiologiques à analyser, du coût de l'analyse d'un échantillon et du budget disponible, du champ

d'application envisagé (phylogénétique, santé publique, sécurité alimentaire, écologie…) ainsi que du champ de compétences nécessaires pour sa mise en œuvre.

I.A.2.c COLONISATION DES NICHES ECOLOGIQUES

Les mycobactéries, dont la majorité des espèces reconnues sont saprophytes colonisent une multitude de niches écologiques plus ou moins éloignées de l'homme[67].

I.A.2.c.i Le sol

Dans une étude concernant les massifs alpins, Thorel a montré que 85% des échantillons de sols prélevés dans les massifs montagneux du Jura, des Vosges et du massif de Belledonne contenaient au moins 8 espèces identifiables de mycobactéries[68]. Les mycobactéries se retrouvent davantage sur des sols frais, humides et riches en acides humiques. A l'inverse, l'espèce *Mycobacterium hassiacum* est capable de croitre jusqu'à une température de 65°C dans un milieu contenant 5% de chlorure de sodium[69], démontrant les possibilités de retrouver les mycobactéries dans des environnements plus extrêmes. Enfin, les mycobactéries peuvent proliférer sur les sols souillées par le bétail [70].

I.A.2.c.ii L'eau

Les mycobactéries ont davantage été observées dans l'eau sous toutes ses formes. Par exemple, elles sont présentes dans les eaux naturelles où elles infectent les amibes[71] et les poissons[72]. A partir de là, elles se retrouvent ensuite dans les systèmes de prélèvement, de traitement, de distribution et de consommation[73] de la ressource hydrique à destination de l'homme. Elles sont alors capables de persister sous forme de biofilm[74], ou de se disséminer sous forme de microgouttelettes formées à partir des tours de réfrigération des systèmes de ventilation[75]. Les eaux usées contiennent également de grande quantité de mycobactéries[76]. Leur persistance dans les eaux de réseau résulte aussi de l'inefficacité des méthodes de désinfection qui ont été adaptées à d'autres pathogènes plus fréquents[77]. L'eau sert donc de véhicule ubiquitaire et passif aux mycobactéries. De plus, elles possèdent une enveloppe fortement hydrophobe qui favorise leur dissémination sous forme d'aérosols en présence d'eau[78], et dont les constituants limiteraient la dessiccation lorsque celle-ci viendrait à manquer[79].

I.A.2.c.iii L'air

La présence des mycobactéries dans l'air est liée à celle de particules de sol ou de microgouttelettes d'eau. Des MNTs ont ainsi été détectés dans les aérosols émis par les eaux de pluie et de rivière [80]. A ce titre, le phénomène aboutissant à la dispersion des MNTs par les microgouttelettes

a été reproduit pour *Mycobacterium intracellulare* et *Mycobacterium scrofulaceum*[81]. L'expulsion de microgouttelettes chargées de bacilles à partir d'individus malades représente également une source importante de mycobactéries dans les environnements clos.

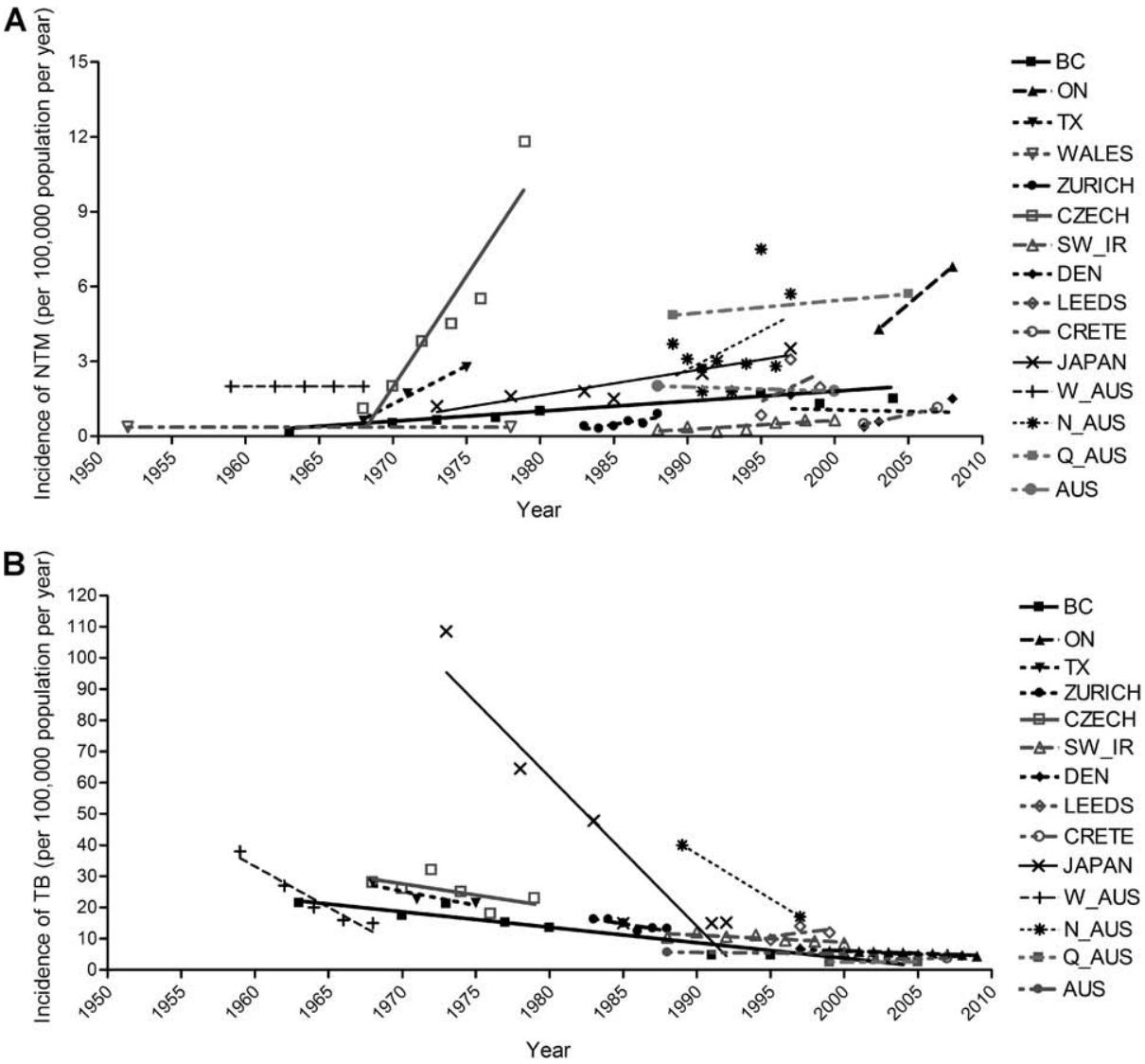
I.A.2.c.iv La faune

Rarement présentes chez les plantes, les mycobactéries sont pathogènes de nombreuses espèces animales qui peuvent parfois être des porteurs sains : poisson, écureuil, tatou, mangouste, opossum, blaireau, sanglier, cochon, cerf, vache, buffle et éléphant comptent parmi les hôtes les plus fréquents[67]. C'est pourquoi, certains d'entre eux sont utilisés comme des modèles biologiques étudiés en laboratoire pour mieux comprendre le cycle infectieux des mycobactéries. C'est le cas du tatou responsable de la transmission de la lèpre chez l'homme[82] ainsi que du poisson zèbre largement utilisé pour l'étude de *Mabs*[83]. Certaines espèces de mycobactéries sont adaptées à plusieurs espèces animales. C'est le cas des variantes du complexe *Mtb* ou du complexe *Mav*. L'infection croisée entre animaux s'opère essentiellement par l'ingestion d'eau ou de nourriture contaminée. Les animaux d'élevage ainsi que sauvage constituent des réservoirs proches de l'homme et susceptibles de le contaminer ensuite.

I.A.3 LES MYCOBACTERIES NON TUBERCULEUSES

Les MNTs regroupent toutes les espèces de mycobactéries qui ne font pas partie du complexe *Mtb* ainsi que de l'espèce *Mlep*. Elles sont non pathogènes, pathogènes opportunistes ou pathogènes strictes. De même que la tuberculose, elles colonisent essentiellement les poumons mais elles se disséminent aussi dans les tissus cutanés[84], les tissus mous sous-jacents, le système lymphatique[85] ou le système nerveux[86]. Une pathologie associée des voies aériennes (asthme, mucoviscidose, bronchoectasie…) ainsi qu'une immunodéficience dont la cause peut être multiple (séropositivité HIV, traitement immunomodulateur, âge…) sont autant de facteurs de risque qui favorise leur apparition[87].

Notre perception de l'impact clinique des MNTs découle de notre capacité à pouvoir les caractériser. La découverte initiale de la lèpre et de la tuberculose a focalisé l'attention sur ces deux espèces, mais il n'est pas exclu qu'entre le 19e siècle et la 1ère partie du 20e siècle l'implication des MNTs dans les pathologies prises en charge ait été sous-estimée.



Graphiques 4 et 5 : Evolution de l'incidence en nombre de cas pour 100 000 personnes des MNTs (A) face à celle de la tuberculose (B) d'après l'étude systématique de Brode[90].

I.A.3.a EVOLUTION CROISSANTE DES MNTs FACE A LA TUBERCULOSE

Dans la province canadienne de l'Ontario, entre 1998 et 2010, le nombre de MNTs isolés chez les patients atteints d'infection pulmonaire a doublé, les auteurs indiquant qu'une telle évolution ne pouvait s'expliquer par l'amélioration des techniques de détection[88]. A Taïwan, entre 2000 et 2008, le nombre de cas de *Mtb* diminue alors que celui de MNTs augmente : leur incidence a été multipliée par 5 atteignant 44,7 cas pour 100000 patients[89]. La même année, 62 centres médicaux issus de 30 pays à travers le monde ont participé à une étude rétrospective de 20000 patients. Une recherche de MNTs dans des échantillons pulmonaires a été réalisée, et 91 espèces différentes ont été dénombrées. Alors

que les espèces du complexe *Mav* sont les plus représentées chez les mycobactéries à croissance lente, *Mabs* et *Mfor* sont les plus fréquentes parmi celles à croissance rapide[90].

En 2014, un inventaire des cas de syndromes tuberculeux survenus dans 9 pays européens tend à montrer que *Mtb* demeure l'espèce dominante dans les pays développés[91]. D'autres travaux tempèrent cette hypothèse : l'étude américaine de Vinnard révèle que le taux d'incidence de décès liés à la tuberculose a baissé de 9 à 3 pour un million entre 1999 et 2014, alors qu'il s'est maintenu autour de 2,6 dans le cas des MNTs[92] ; dans la région de Madrid, le nombre de cas d'infections aux MNTs a même dépassé celui de *Mtb* entre 2013 et 2017[93]. Un travail de revue systématique récent (Graphiques 2 et 3) souligne que les tendances, décroissante pour la tuberculose et ascendante pour les MNTs, se croisent dans la majorité des pays industrialisés anglophones[94]. Enfin, la revue et les travaux d'Adjémian sur la prévalence des MNTs indiquent que celle-ci oscille entre 1,4 et 14 pour 100000 habitants aux Etats-unis[95].

I.A.3.b IMPACT ECONOMIQUE

L'essor des MNTs sur la santé publique conduit les pays économiquement les plus développés à s'intéresser à leur coût. En 2010 aux Etats-Unis, le nombre d'infections pulmonaires associées aux MNTs était estimé à plus de 86000 cas, pour un coût total de 815 millions de dollars[96]. Sur les années 2010-2011, en Allemagne, le coût direct lié à la prise en charge d'un patient atteint d'une infection pulmonaire associés aux MNTs atteint quasiment 40000 euros par an, contre 10 euros pour un patient contrôle avec un taux de mortalité de 22% pour les patients alors qu'il est de 6% dans le groupe contrôle[97].

I.A.3.c LA SITUATION EN FRANCE

Une étude menée en France, de 2009 à 2014, sur 477 cas de maladies pulmonaires associées aux MNTs place les espèces du complexe *Mav* en tête, et celles du complexe MABSC en 4ème position[98]. Elle montre également que l'âge, le choix du traitement et l'identification de *Mabs* sont les trois facteurs pronostics les plus significatifs. Une 2ème étude française sur plus de 5000 patients conduite de 2010 à 2017 révèle que la prévalence des MNTs avoisine 6 pour 100000 habitants, et que le risque de mortalité est plus élevé chez les personnes âgées ou non traités durablement par les antibiotiques. Le coût de prise en charge dépasse les 12000 euros quand il est 10 fois moindre pour un patient contrôle.

L'apparition croissante de nouvelles espèces et leur implication documentée en santé publique justifie de s'intéresser à des espèces non pathogènes pour mieux comprendre la physiologie mycobactérienne.

I.A.3.d MYCOBACTERIUM SMEGMATIS : ESPECE MODELE

Cette espèce, décrite pour la 1ère fois en 1899[5,6], est utilisée comme modèle d'étude du genre *Mycobacterium* du fait de son caractère non pathogène. De plus, c'est une espèce à croissance rapide qui fut un outil précieux pour décrire la physiologie générale des mycobactéries[99]. Depuis la 1ère description de souches mutantes possèdant un taux de transformation de plasmide 100000 fois supérieur à la souche sauvage par Snapper[100], les techniques de biologie moléculaire ont ainsi bénéficié d'une amélioration majeure afin de tirer profit de la souche type mc^2 155[101,102]. La manipulation du génome de *Msmeg* a contribué à la description de gènes essentiels ou non et associés à la croissance, la division, la production de la paroi, la synthèse de ses constituants, la résistance antibiotique... *Msmeg* permet également la production de protéines recombinantes identifiées chez les MNTs et *Mtb* dont les structures ont ainsi pu être décrites [103]. Enfin, *Msmeg* fait partie des espèces capables de produire des glycopeptidolipides (GPLs) en grande quantité dans son enveloppe et dont la structure se rapproche de ceux de *Mabs*.

I.A.3.e LE COMPLEXE MYCOBACTERIUM ABSCESSUS

Mycobacterium abscessus (Mabs) est une mycobactérie non tuberculeuse à croissance rapide qui doit son nom à sa première description en 1953, à partir d'un abcès de genou[12]. A ce titre, la souche type ATCC19977 est regroupée avec la souche type NCTC946 de *Mche*, par Kubica en 1972. Les 2 membres, qui se différencient notamment par leur capacité à utiliser le citrate, constituent ensemble le complexe « *Mycobacterium chelonei* »[104]. En 1992, Kusunoki démontre que les 2 sous-espèces se distinguent effectivement par la culture en présence de 5% NaCl, de 0,2% picrate ou d'allantoïne. Le degré d'homologie réciproque de leur ADN atteint 35%, justifiant que la souche ATCC19977 soit finalement élevée au rang d'espèce[105]. La description ultérieure de *Mycobacterium immunogenum (Mimu)* au sein du complexe *Mfor* met en lumière quelques nouveaux marqueurs moléculaires entre *Mabs* et *Mche*. En particulier, la séquence de leur gène *hsp65* diffère de 29 paires de base, dans une région hypervariable de 441 paires de base[106]. A l'inverse, la région choisie du gène codant l'ARN 16S est identique. En 2009, Nash révèle qu'un autre marqueur moléculaire remarquable qui distingue *Mabs* de *Mche* est l'absence chez ce dernier du gène *erm(41)*, responsable d'une résistance inductible aux macrolides[107]. La même année, le complexe MABSC est créé.

I.A.3.e.i Les sous-espèces du complexe MABSC

En effet, parmi des échantillons cliniques du Brésil, 3 sous-espèces *Mycobacterium subs abscessus (Mabs)*, *subs massiliense (Mmas) et subs bolletii (Mbol)* n'ont pas pu être isolées sur la base de leur phénotype de croissance, contrairement à *Mimu* et *Mche*[108]. De même, l'analyse du polymorphisme du gène de l'ARN 16S ou de leur contenu en acides mycoliques ne les a pas discriminées. Le niveau d'homologie de leur ADN, supérieur à 70%, amène donc les auteurs à rassembler les 3 sous-espèces au sein d'un même complexe.

En fait, *Mmas* avait été préalablement identifié par Adékambi grâce à une série de tests phénotypiques plus exhaustif et à la variabilité d'une séquence non codante entre les gènes de l'ARN 16S et 23S, comparativement à *Mabs*[109]. Il réussit ensuite à caractériser la seconde sous-espèce, *Mbol*, au moyen du séquençage du gène *rpoB*[110]. Le séquençage complet de son génome s'effectue en 2012[111]. Avant que des règles de nomenclature ne fassent disparaitre temporairement *Mmas*[112], c'est le séquençage complet du génome de 53 souches cliniques apparentées à la souche type CIP108297, et la construction de leur arbre phylogénétique à partir de l'ANI, du gène *rpoB* ou de 11 gènes assemblés qui suggèrent de revenir à la division de *Mbol* en 2 sous-espèces[113].

Les 3 sous-espèces redeviendront équivalentes au sein du complexe MABSC à la suite de 2 études bioinformatiques. La 1ère compare les 3 souches types avec des souches cliniques et détaille le contenu en G+C variant de 62,7% à 64,2%, la longueur du génôme de 4,8 Mb à 5,51Mb et le nombre de gènes uniques de 15 à 36. Parmi toutes les souches, plus de 57000 séquences codant des protéines circonscrivent un génome de base qui ne représente que 64% du pangénome, révélant une très faible conservation génétique[114]. A titre d'exemple, le même travail sur les souches connues de *Mtb* montre une couverture de 96% du pangénôme. La 2ème fut publiée en 2016 par Tortoli, sur un ensemble plus large de 43 génomes, prélevées dans la base de données GenBank. D'après l'ANI supérieur à 96%, la distance génétique n'est pas suffisante pour la définition d'espèces, mais elle confirme les 3 sous-espèces[115]. Mise en évidence par Kim[116], la présence du gène *erm(41)* chez *Mbol* qui ne l'est plus chez *Mmas* justifie aussi leur distinction.

I.A.3.e.ii Epidémiologie des infections associées à *Mabs*

Ponctuellement, les infections de la peau et des tissus mous sont consécutives à une intervention chirurgicale superficielle[117],[118]. De la même manière, le système nerveux est affecté après la pose de cathéter[119],[86]. Néanmoins, les pathologies les plus couramment rencontrées sont essentiellement pulmonaires, et les études épidémiologiques tentent d'évaluer la transmission de

l'infection respiratoire à partir de l'environnement ou de patients malades qui seraient porteurs d'une souche naissante fortement contagieuse.

En France, à la fin des années 1990, le suivi de 298 patients atteints de mucoviscidose montre que *Mabs* est la MNT la plus fréquemment détectée, sans qu'une contamination entre les individus ou à partir d'une source environnementale ne soit démontrée[120]. La tendance est confirmée entre 2003 et 2004, par une autre étude française incluant 49 centres français et 1582 patients, et qui révèle un taux de prévalence de *Mabs* égal à 3,2%[121]. En 2006, une étude suédoise sur 28 patients dont seulement 14 sont atteints de mucoviscidose montre que 24 patients sont infectés par une souche unique. Les auteurs en déduisent alors que la transmission de *Mabs* à partir d'une source environnementale est plus probable qu'une contamination entre individus[122].

En 2013, le séquençage du génome complet de 168 isolats cliniques, dans la population britannique, incite Bryant à proposer un modèle de transmission directe entre individus atteints de mucoviscidose[123]. En 2016, il étend son étude à 517 patients britanniques, australiens et américains. Il montre que 74% d'entre eux sont concernés par 2 clusters de *Mabs* et le 3ème de *Mmas*[124]. Ces 3 clones circulants dominants (DCC1, DCC2 et DCC3) étant présents sur tous les territoires, la thèse de la transmission directe se renforce.

En Italie, le clone dominant DCC1 a été retrouvé chez des patients atteints de mucoviscidose, mais les possibilités de transmission directe sont circonscrites au sein de groupes ne dépassant pas 3 individus ayant fréquenté le même centre de soin[125]. De même, les souches appartenant aux 3 clusters majoritaires, DCC1 à DCC3, sont présentes chez 123 patients allemands mais aucun indice de transmission directe n'a pu être récolté[126]. Au sein de l'hôpital australien de Melbourne, une transmission directe est décrite dans un cluster de 3 patients atteints de mucoviscidose dont les souches sont distantes de moins de 7 SNPs[127].

La mucoviscidose s'avère y être une maladie moins fréquente, mais sur le continent asiatique, avec une représentation de plus d'un tiers, *Mabs* dépasse les autres MNTs parmi les syndrômes respiratoires observés à Singapour[128]. Dans la province japonaise d'Okinawa, le complexe MABSC est en tête des MNTs détectés chez les patients victimes de symptômes respiratoires[129]. D'autre part, Taïwan est également concerné puisqu'entre 2010 et 2014, le nombre de nouvelles infections par *Mabs* représente un quart des cas constatés[130]. Finalement, l'analyse de 220 isolats japonais et taïwanais confirme que le plus grand nombre de souches appartient au cluster des clones dominants DCC1 de *Mabs*, et que le 2ème cluster DCC2 est restreint aux patients atteints de mucoviscidose[131].

Il semble donc moins à craindre une vaste expansion épidémique causé par *Mabs* qu'une persistance d'au moins 3 clones ciblant les populations affaiblies. En outre, la spécialisation apparente de *Mabs* envers les patients atteints de mucoviscidose s'accompagne du plus grand arsenal de résistance antibiotique dans le genre *Mycobacterium*.

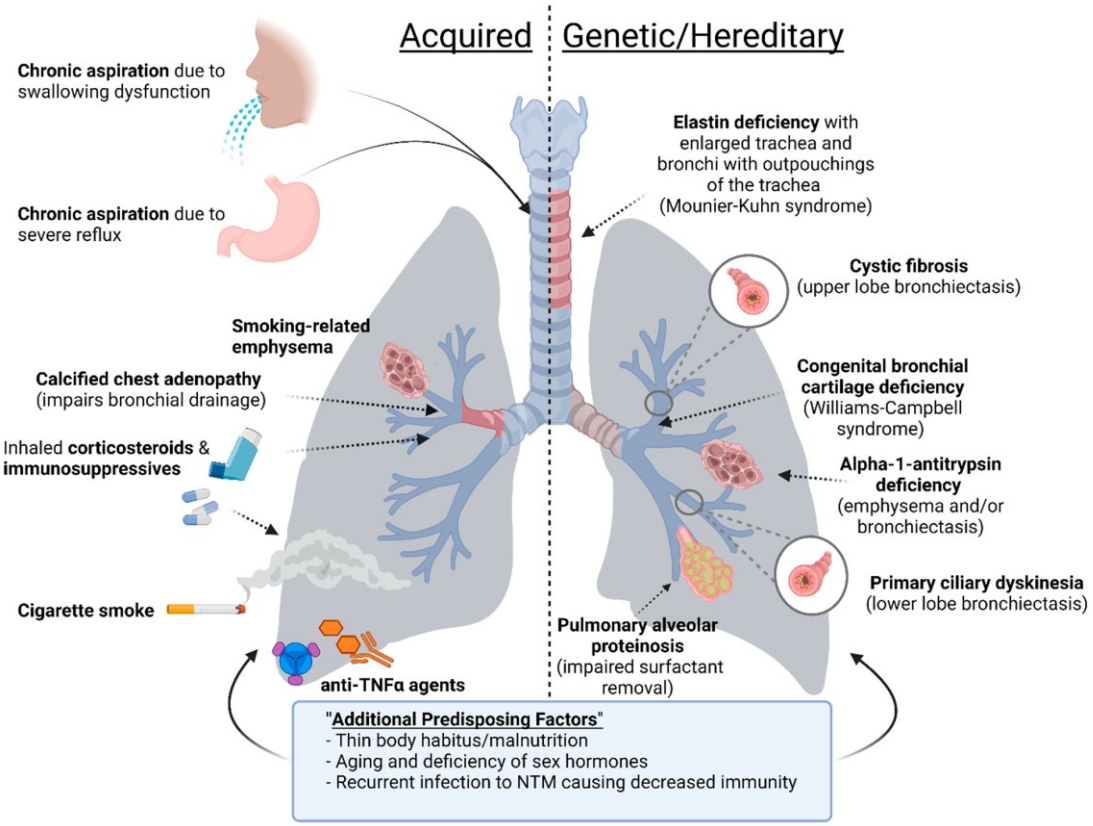


Figure 3 : Facteurs de risque acquis ou héréditaires associés aux infections pulmonaires par *Mabs* d'après la revue d'Abdelaal[132]. Parmi les risques acquis (à gauche), le refflux digestif, la consommation de tabac et certains traitements prédisposent aux infections pulmonaires. Dans les critères héréditaires (à droite), tous concernent la fonction respiratoire.

I.A.3.e.iii Facteurs de risque

Les dysfonctionnements de la fonction respiratoire s'accompagnent régulièrement d'infections par les MNTs, en particulier les désordres de la clairance pulmonaire. Le mucus qui tapisse nos bronches capte les micro-organismes, puis les élimine par excrétion vers l'appareil digestif. De plus, les macrophages alvéolaires contenus dans ce mucus participent à leur reconnaissance et à leur élimination. Toute modification d'origine génétique, anatomique ou médicamenteuse qui diminue la clairance pulmonaire est encline à favoriser les infections notamment par les MNTs. De la même manière les déficiences acquises ou induites du système immunitaire inné limitent le contrôle des infections par les macrophages. Les pathologies associées aux opportunités d'infection sont donc

multiples[132,87]: mucoviscidose, dyskinésie ciliaire primaire, bronchoectasie, syndrome de Lady Windermere, asthme, broncho-pneumopathie chronique obstructive, emphysème, déficit en alpha-1 anti trypsine, déficit en élastine… L'ensemble des risques connus chez les patients est illustré et listé dans la Figure 3 tirée de la revue d'Abdelaal[132]. Il subsiste néanmoins des personnes saines et sans risque apparent qui développent néanmoins une pathologie pulmonaire associée aux MNTs. Dans ce cas, l'historique, la physiologie et la morphologie du patient impliquent des causes plus subtiles et difficilement identifiables : le tabagisme, un indice de masse corporel inférieur à 18, un taux de leptine faible, une légère scoliose, un âge avancé ou le sexe sont des éléments à prendre en compte.

I.A.3.e.iv Diagnostic des infections associées à *Mabs*

Compte tenu de l'existence d'au moins une dizaine d'espèces pathogènes opportunistes parmi les MNTs et de l'absence potentielle de facteurs significatifs de risque, le diagnostic des trois espèces du complexe MABSC exige la conjonction d'indices cliniques, radiologiques, biochimiques et génétiques.

Dans un premier temps, pour les atteintes pulmonaires, l'examen radiographique du tronc nécessite l'expertise du praticien pour distinguer entre des lésions associées à *Mtb* ou aux MNTs qui présentent un aspect en nid d'abeilles[133]. Ensuite, la recherche microscopique dans les prélèvements de sécrétion bronchique doit mettre en évidence des bacilles acido-alcoolo résistants (cf chapitre DETECTION). C'est une approche particulièrement utilisé dans les pays en voie de développement qui doivent contrôler un grand nombre de syndromes tuberculeux. La mise en culture des échantillons doit confirmer la vitesse de croissance, lente dans le cas de *Mtb* ou du complexe *Mav*, et rapide pour MABSC, ainsi que les caractères biochimiques et métaboliques (cf chapitre DETECTION). Enfin, l'amplification et l'analyse de l'ADN des colonies obtenues sont considérées comme les techniques de référence par l'Organisation Mondiale de la Santé (OMS) et l'American Thoracic Society (ATS) pour identifier l'espèce parmi les MNTs. Dans le chapitre sur la détection des mycobactéries 3 kits commerciaux ont été cités mais d'autres versions sont décrites dans la revue de Chin[134].

L'analyse MALDI-TOF est une alternative moins répandue. Parmi 152 échantillons cliniques, elle a permis l'identification correcte de 148 souches comparativement aux 149 identifiées par l'analyse des produits de restriction du gène *hsp65* amplifié[135]. Récemment, une tentative de dosage des anticorps sériques IgA a été évaluée chez des patients atteints de mucoviscidose. Elle a ciblé deux antigènes propres à *Mabs* : une phospholipase C recombinante et une fraction de lipoprotéine activatrice du récepteur TLR-2. En combinant les deux résultats, une sensibilité de 66,7% et une

spécificité de 85,4% ont été calculées[136]. La performance des tests de PCR en temps réel pour la détection de mutations clés dans les gènes de l'ARN 23S et 16S est supérieure avec 83,8% de sensibilité et 97,8% de spécificité[137].

Dans le cas du complexe MABSC, la recherche des sous-espèces est obligatoire pour guider correctement la conduite thérapeutique qui en découle puisque leur susceptibilité aux antibiotiques n'est pas la même.

Tableau 1 : Multirésistance antibiotique de *Mabs*. Liste exhaustive des antibiotiques et des fonctions ciblées chez les mycobactéries. La synthèse d'arabinogalactane et des acides mycoliques ainsi que leur transport sont des fonctionsspécifiques. *Mabs* possède trois mécanismes principaux de résistance : la modification de l'antibiotique, son export ou le changement structurale de la cible visée.

ANTIBIOTIQUE		CIBLE			RESISTANCE		
Classe	*Molécules*	*Fonction*	*Gène*		*Mécanisme*	*Gène*	
	éthambutol	synthèse de l'arabinogalactane	*embB*	*MAB_0185c*	le polymorphisme du gène modifie la cible		*MAB_0185c*
	isoniazide (prodrogue)	synthèse des acides mycoliques	*InhA*	*Rv1484*	pas d'activité catalase pour activer la prodrogue	*KatG*	*MAB_2470c*
	thiacetazone		*HadAB*	*Rv0635/Rv0636*	protéines redondantes assurant la même fonction	*Had*	*MSMEG_6754*
	analogues D6, D15, D17				export de l'antibiotique par une pompe d'efflux	*MmpL5*	*MAB_4382c*
benzimidazoles	EJMCh-6	transport des acides mycoliques	*MmpL3*	*MAB_4508*			
benzothiazoles	CRS400393						
piperidinoles	PIPD1						
indole-2 carboxamides	IC5, IC25						
glycopeptides	oritavancine	synthèse du peptidoglycanne			hydrolyse de l'antibiotique par une β-lactamase	*bla*	*MAB_2875*
β-lactamines	ampicilline, amoxicilline, carbapenem, imipenem, cefoxitine						
Inhibiteurs de β-lactamase	avibactam	activité β-lactamase	*bla*	*MAB_2875*			
	pyrazinamide (prodrogue)	synthèse d'acides gras, ribosome			une mutation inactive la pyrazinamidase	*PncA*	*MAB_1472c*

ANTIBIOTIQUE		CIBLE			RESISTANCE		
Classe	*Molécules*	*Fonction*	*Gène*		*Mécanisme*	*Gène*	
aminoglycosides	gentamicine C, dibekacine, tobramycine capréomycine, kanamycine	ribosome (ARN 16S)	*rrs*	*MAB_r5051*	modification de l'antibiotique par une N-acétyltransférase	*aac(2')*	*MAB_4395*
						eis2	*MAB_4532c*
	amikacine				le polymorphisme du gène modifie la cible		*MAB_r5051*
	streptomycine				inactivation de l'antibiotique par phosphorylation	*str3''*	*MAB_2385*
macrolides	clarithromycine, erythromycine, azithromycine	ribosome (ARN 23S)	*rrl*	*MAB_r5052*	modification de la cible par une méthyltransférase	*erm(41)*	*MAB_2297*
					le polymorphisme du gène modifie la cible		*MAB-r5052*
tétracyclines	tigecycline, doxycyline, omadacycline	ribosome (ARN 30S)			dégradation de l'antibiotique par une flavine monooxygénase	*tetX*	*MAB_1496c*
oxazolidinones	linezolide, tedizolide, delpazolide	ribosome (ARN 50S)			export de l'antibiotique par une pompe d'efflux	*MmpL9*	*Rv2339*
rifamicines	rifampicine, rifabutine, rifapentine	ARN polymérase (sous-unité β)	*rpoB*	*MAB_3869c*	inactivation de l'antibiotique par l'ADP-ribosyl transférase	*arr*	*MAB_0591*
fluoroquinolones	moxifloxacine	réplication de l'ADN	*gyrA*	*MAB_0019*			
benzimidazoles	SPR719, SPR720		*gyrB*	*MAB_0006*			
diarylquinoléines	bedaquiline	ATP synthase	*atpE*	*MAB_1448*	export de l'antibiotique par une pompe d'efflux	*MmpL*	*MAB_2301*
phenazines	clofazimine	NADH quinone oxidoreductase (H_2O_2)				*MmpL*	*MAB_1134c*
cyclopeptides	rufomycine	ClpC1 complexe (protéines chaperones)					
	auranofin	thiorédoxine reductase					

I.A.3.e.v Mécanismes de résistance antibiotique

L'apparition de souches de *Mtb* réfractaires à la rifamycine ou à la rifampycine, qui sont suivies par l'OMS montre que cette problématique est préoccupante. Parmi les MNTs, les sous-espèces du complexe MABSC possèdent le plus grand nombre de mécanismes de résistance (Tableau 1).

Historiquement, l'évaluation des anti-tuberculeux efficaces contre *Mtb* a montré que ces derniers n'avaient qu'une action limitée sur *Mabs*. L'isoniazide, l'éthambutol ou la thioacétazone bloquent la formation de l'enveloppe mycobactérienne mais *Mabs* y est peu sensible. A l'inverse, les β-lactamines empêchent la synthèse du peptidoglycane bactérien mais *Mabs* produit une enzyme qui

les dégrade, la β-lactamase. Afin de contourner ce mécanisme, l'utilisation d'inhibiteurs de β-lactamase en association avec les β-lactamines est une solution synergique alternative[138].

Les macrolides, les aminoglycosides et les tétracyclines qui limitent la synthèse des protéines en interagissant avec le ribosome sont sujettes à de nombreux mécanismes d'évitement. Dans un travail de 2011, Bastian a séquencé les gènes *rrl* et *erm(41)* de 90 souches cliniques du complexe MABSC[139]. Le premier code une région peptidyltransférase de l'ARN 23S dont certaines mutations engendrent une résistance à la clarithromycine, et le second pour une méthylase dont la production peut être induite par l'exposition à la clarithromycine afin de modifier l'ARN 23S et d'y résister[107]. Dans le même temps, elle a déterminé la concentration minimale d'inhibition (MIC) par la clarithromycine de ces mêmes souches sur 14 jours. De cette manière, elle constate des délétions dans le gène *erm(41)* pour *Mmas*, accompagnées de MIC_{50} inférieur à 1µg/mL. Pour *Mabs*, la position 28 du gène *erm(41)* présente un polymorphisme T/C associé à une valeur de MIC_{50} respectivement supérieure à 16µg/mL ou inférieure à 1µg/mL. *Mbol* ne possède ni délétion ni polymorphisme et une MIC_{50} supérieure à 16µg/mL. La détermination en amont du profil des gènes *erm(41)* et *rrl* fut encore récemment utilisé pour orienter la médication des patients[140]. D'autre part, Il existe au moins deux acétyltransférases et une phosphotransférase qui modifient les aminoglycosides, diminuant leur affinité potentielle avec la cible ribosomale[141]. Enfin, les tétracyclines seraient modifiées et dégradées par une flavine monooxygénase même si son action seule n'a pas pu être validée chez *Msmeg*[142].

Les rifamicines entravent l'activité de l'ARN polymérase mais leur ADP-ribosylation par une enzyme spécifique de *Mabs* rend ce dernier résistant[143]. Toujours parmi les acides nucléiques, la réplication de l'ADN est ciblée par les fluoroquinolones et benzimidazoles. Aucun mécanisme pérenne de résistance n'a été décrit pour ces deux classes de drogues. C'est également le cas de celles qui déstabilisent le transport des acides mycoliques. Celui-ci doit s'effectuer sous forme de tréhalose monomycolate (TMM) via le transporteur « Myco membrane protein Large » 3 (MmpL3). La spécificité au genre *Mycobacterium* de ce processus biologique encourage les nombreux développements pour bloquer cette voie[144]. Des polymorphismes génétiques engendrent aussi des modifications de la séquence protéique de la cible et une moindre affinité de l'antibiotique. La plupart des mutations identifiées parmi les mycobactéries sont listées dans une revue de Cattoir[145].

Un dernier exemple de mécanisme de résistance communément décrit par Gutierrez pour la bédaquiline et la clofazimine est leur export hors du cytoplasme par deux transporteurs de type MmpL et codés par les gènes *MAB_2301* et *MAB_1134c* [146]. Leur niveau d'expression augmente *via* le

régulateur transcriptionnel TetR qui aurait préalablement fixé les drogues qui ne sont plus en quantité suffisante pour agir sur leurs cibles. Des phénomènes similaires ont été proposés pour les nouveaux dérivés de thioacétazone[147]. La multirésistance antibiotique compte parmi les paramètres qui favorisent le maintien et la virulence des mycobactéries.

I.A.3.e.vi Facteurs de virulence de *Mabs*

Le critère pour estimer la virulence de *Mabs* dépend du modèle biologique choisi. La drosophile, la souris, le poisson-zèbre, l'amibe et la culture primaire de macrophages sont les principaux modèles utilisés[83].

Transfert génétique horizontal

Une partie des caractéristiques responsables de cette virulence résulte du transfert horizontal de gènes à partir de bactéries distantes du genre *Mycobacterium*. C'est le cas de l'acquisition d'une phospholipase C : cette protéine, très similaire à celle rencontrée chez *Pseudomonas aeruginosa*, participe à l'évasion du phagosome. Chez *Mabs* elle provoque aussi la lyse des macrophages et améliore la survie dans l'amibe[148]. En outre, le génome de *Mabs* code deux protéines MgtC, fortement similaires à celles présentes chez *Salmonella enterica serovar typhimurium*, qui facilite la croissance sur un milieu pauvre en magnésium ou dont le pH est faible[149]. L'acquisition de cette protéine est envisagée comme une adaptation à la survie intracellulaire. Enfin, *Mabs* possède un gène codant la peptide methionine sulfoxide réductase (MsrA) qui pourrait améliorer la survie lors d'un stress oxidatif[150].

Systèmes de sécrétion de type VII (ESX)

Les systèmes de sécrétion sont des complexes qui exportent des effecteurs protéiques afin d'avantager la mycobactérie. Parmi eux, *Mabs* possède deux systèmes de sécrétion ESX-3 et ESX-4. Le premier est impliqué dans l'utilisation du fer en association avec la mycobactin[151]. La délétion des gènes qui codent ESX-3 diminue la croissance de *Mabs* dans les macrophages ainsi que leur production de cytokines proinflammatoires TNF-α, IL1-β et IL-6[152]. L'ensemble du 2ème système de sécrétion ESX-4 limite l'acidification du phagosome, et favorise l'échappement vers le cytoplasme grâce aux produits des gènes *esxU* et *esxT* capables de perméabiliser les membranes[153].

Protéines de type MCE

En outre, 7 opérons codant des protéines de type « Mammalian Cell Entry » (MCE) ont été identifiés dans le génome de *Mabs*[154]. L'utilisation de fragments d'ADN contenant les opérons *mce* de

Mtb chez une souche non pathogène d'*Escherichia coli* lui a permis d'infecter des cellules HeLa[155]. Malgré tout, les études portant sur l'expression des gènes *mce* de *Mabs* ne sont pas suffisantes pour soutenir un rôle central dans sa virulence, comparativement à d'autres espèces.

Protéines de transport type MmpL

Parce qu'elles participent notamment à l'élimination de certains antibiotiques ou qu'elles contribuent à la sécrétion de glycolipides pariétaux, les protéines de transport de type MmpL contribuent à la virulence des mycobactéries. De plus, parmi les mycobactéries, le génome de *Mabs* compte le plus grand nombre de gènes codant des MmpL[156]. Parmi les 27 gènes annotés, certains ont des fonctions biologiques connues qui peuvent être corrélées avec la virulence de *Mabs*. Par exemple, la protéine MmpL3 contribue au transport des acides mycoliques sous forme de TMM afin de bâtir une mycomembrane externe fortement hydrophobe et imperméable[157]. La protéine MmpL4 transporte les glycopeptidolipides (GPLs) dont les rôles biologiques, qui seront discutés plus loin, servent non seulement sa dissémination mais aussi son cycle pathogène. La délétion du gène *MmpL8*, associé à l'export d'un nouveau glycolipide appelé Glycosyl Diacetylated Nonadecyl Diol (GDND), diminue la virulence de *Mabs* chez la souris[158]. D'ailleurs, son homologue chez *Mtb* exporte le sulfolipide SL-1 également très impliqué dans la virulence[159]. Enfin, comme indiqué précédemment, les gènes *MAB_2301* et *MAB_1134c* codent les transporteurs MmpL5 et MmpL9 qui éliminent la bedaquiline et la clofazimine[146].

Régulateurs transcriptionnels

Les régulateurs de transcription dont la plupart ont été décrits chez *Mtb,* contrôlent des fonctions biologiques favorables à la virulence de *Mabs*. Sous l'effet d'un pH acide ou de l'infection de macrophages, la protéine VirS favorise l'expression de l'opéron *mymA* dont les gènes codent des protéines impliquées dans la modification des acides gras. Dans le même temps, la morphologie du bacille et de sa paroi est altérée chez le mutant *ΔVirS*. Le génome de *Mabs* contient justement un homologue de VirS dont la fonction reste à élucider[154]. Il présente aussi 5 homologues de facteurs de transcription de type Sig A, C, D, E et H : le premier est essentiel à la croissance de *Msmeg*[160] ; lorsque l'expression du second est interrompue chez *Mtb*, la létalité chez la souris diminue mais pas la charge bactérienne[161] ; dans le cas de sigD, il influence la sécrétion de TNFα par les macrophages[162] ; les deux derniers régulateurs de transcription, sigE et sigH, confèrent un avantage de croissance dans les conditions défavorables d'augmentation de la température et de stress oxydatif ainsi que de diminution du pH[163],[164]. Récemment décrit chez *Mabs,* le facteur de transcription SigH peut porter une

mutation ponctuelle qui favorise la résistance à la tigecycline par l'expression du système de dégradation des tétracyclines Mab TetX. En parallèle, la mutation n'affecte pas l'expression des gènes de la thioredoxin ou des protéines de choc thermiques associées au stress. Cette mutation est donc considérée comme un gain de fonction favorable à l'évolution de *Mabs* vers la virulence[165].

Lsr2 est un autre facteur de transcription, similaire aux histones par son interaction avec l'ADN, qui a été largement étudié. Chez *Msmeg*, il est corrélé à la morphologie des colonies et à leur capacité à former des biofilms qui s'accompagne d'une diminution du méromycolyl diacylgycérol (MMDAG) dans l'enveloppe[166]. Chez *Mtb*, il s'avère essentiel et contribue à la protection de l'ADN du stress oxydatif suite à la production d'espèces réactives de l'oxygène[167]. Le même rôle a été démontré chez *Mabs,* puisque le mutant rugueux *Δlsr2* est plus sensible au peroxyde d'hydrogène comparativement à la souche sauvage. Il présente également une moindre croissance intracellulaire chez l'amibe et le macrophage murin, ainsi qu'une diminution de la mortalité provoquée chez les embryons de poisson zèbre[168]. Ce même travail montre que le niveau d'expression de lsr2 est plus élevé chez le morphotype rugueux que chez la forme lisse, réputée moins virulente lors du cycle infectieux.

I.A.3.e.vii Cycle infectieux de *Mabs*

Les 2 morphotypes rugueux et lisse de *Mabs*

Au début des années 2000, Byrd parvient à isoler une souche clinique de *Mabs* qu'il observe sous 2 formes distinguées par l'aspect morphologique des colonies sur agar. La forme rugueuse présente une texture irrégulière et sèche, voire plurilobée alors que celle de la forme lisse, issue de la 1ère, a un aspect plus homogène, aplani et brillant. A l'époque, cette découverte n'est pas originale puisque d'autres espèces mycobactériennes ont été précédemment décrites sous ces deux aspects. Par contre, le fait que la forme rugueuse persiste dans les poumons de souris immunodéficientes, se multiplie dans les monocytes humains ou forment des cordes au sein de lignées cellulaires de fibroblastes lui est exclusive[169]. En 2006, la même équipe montre que les repiquages successifs de la forme lisse (390S) s'accompagnent de l'apparition spontanée d'une forme nouvellement rugueuse (390V). Sa caractérisation approfondie révèle une morphologie rugueuse identique à son ancêtre (390R), ainsi qu'une capacité à produire des cordes serpentiformes. La persistance dans les poumons de souris et la multiplication dans les monocytes humains sont deux caractères virulents également restaurés. Par contre, seule la forme lisse est motile et produit des biofilms ainsi que des glycopeptidolipides (GPLs) pariétaux[170].

A la même période et au moyen de la souche CIP104536T, Catherinot et collaborateurs obtient un résultat identique puisqu'elle isole la forme rugueuse qui ne produit pas de GPLs et qui se montre plus virulente chez la souris[171]. De plus, les macrophages murins produisent du TNFα à la suite de l'exposition aux deux formes rugueuses 390R et CIP104536R. La forme rugueuse augmente donc la réponse inflammatoire. En parallèle, la forme lisse n'induit pas cet effet mais les GPLs qu'elle produit ne modulent pas la production de TNFα. L'équipe de Byrd poursuit son travail sur la souche 390 en milieu liquide. Par microscopie électronique et coloration de Ziehl-Neelsen, elle demontre que les bacilles de la forme rugueuse s'agglutinent selon un axe longitudinal ce qui aboutit à la formation d'une structure macroscopique, appelée corde, qui ressemble à celle produite par *Mbov BCG*. A l'inverse, les formes lisses 390S et DSMZ44196 ne produisent pas de cordes[172].

L'étude de 76 souches cliniques du complexe MABSC renforce le lien établi entre la morphologie rugueuse et l'absence de GPLs dans l'enveloppe. En effet, parmi ces échantillons, les auteurs individualisent la souche rugueuse Asan 50594 de type II appartenant à la sous-espèce *Mmas*, du fait de l'absence de GPLs. De façon cohérente, son génotype se caractérise par une dizaine de délétions dans le locus de biosynthèse des GPLs[173]. Dans le même temps, Pawlik et collaborateurs constatent des modifications génétiques localisées dans le locus de biosynthèse des GPLs du génome des souches 19977-IP-R (=CIP104536T), 19977-AT-R et CF-R. Chez la 1ère, l'insertion CG dans le gène *mps1* entraine une diminution de la transcription des gènes impliqués dans la biosynthèse des GPLs. D'autre part, le clonage d'une portion s'étendant du gène *mbtH* à *gap* permet de transformer une souche *Msmeg* d'aspect initialement rugueux vers une forme lisse, tout en rétablissant la production d'un niveau élevé de GPLs[174]. Les résultats successifs tendent à confirmer l'influence majeure de la présence de GPLs sur le morphotype observé. Le suivi de souches cliniques de 9 patients en 2015 abonde dans ce sens. En effet, sur l'ensemble des prélèvements de forme rugueuse, toutes possédaient des mutations, insertions ou délétions dans les gènes *mps1*, *mps2*, *MmpL4* et *MmpS4* du locus de biosynthèse des GPLs[175].

Les deux morphologies de *Mabs* ne sont pas uniquement dues à une différence de production des GPLs. En effet, les travaux de Llorens et collaborateurs montrent que le traitement à l'éther de pétrole de trois souches rugueuses (390R, 44196R et BE48R) entraîne la disparition des cordes serpentiformes caractéristiques. Or cette fraction contient notamment des tréhaloses polyphléates (TPPs) en plus grande quantité que le tréhalose dimycolate (TDM), initialement supposé comme responsable de la formation des cordes. De plus, ils démontrent la possibilité de former *in-vitro* des agrégats à partir de billes fluorescentes recouvertes par les TPPs, et que ces derniers n'ont pas d'effets

délétères sur les macrophages[176]. Ces observations suggèrent que la proportion des glycolipides pariétaux conditionne la morphologie de *Mabs.* De même, Halloum et collaborateurs montrent que la délétion d'une déhydratase altère la proportion d'acides mycoliques α/α' en même temps que la capacité à former des cordes, suggérant l'implication des acides mycoliques dans la morphologie[177].

En outre, deux études distinctes démontrent la possibilité d'un passage de la forme rugueuse à la forme lisse par l'insertion de transposon dans les gènes *MAB_3168c*[178] et *MAB_3083c*[179] qui codent respectivement pour une acétyltransférase et pour une RNase. Dans les deux cas, les auteurs indiquent que la composition lipidique demeure identique entre la forme rugueuse « sauvage » et le mutant de transposition lisse, notamment la composition en GPLs dans le cas de *MAB_3083c*. Par contre la motilité augmente et la capacité des bacilles à s'agréger diminue chez les mutants lisses. Les autres phénotypes observés sont probablement liés aux fonctions spécifiques de chaque gène : dans le cas du mutant *MAB_3168c*, il est plus sensible au lysozyme ainsi qu'à l'amikacine, et sa croissance dans les macrophages ou sa capacité à former des biofilms diminuent ; dans le cas du mutant *MAB_3083c*, la sensibilité au lysozyme et la survie intra-macrophagique sont inchangées. Ainsi, la faculté supérieure d'étalement de la forme lisse lui confère une plus grande aptitude à coloniser son environnement. A l'inverse, les agglomérats de bacilles de la forme rugueuse conduisent à des structures macroscopiques qui échappent plus facilement à la phagocytose et se multiplient plus efficacement dans les phagosomes.

Interaction avec le système immunitaire

Compte tenu de leur phénotype, la forme lisse est considérée comme l'entité capable de coloniser le milieu environnant : elle est responsable de la primo-infection. A l'inverse la forme rugueuse qui possède la plus forte virulence est responsable de la réaction inflammatoire (Figure 4).

En 2009, Rhoades et collaborateurs montrent que seuls les variantes rugueuses 390R et CF-R induisent la production de TNFα par les macrophages dérivés de monocytes[180]. Cette dernière peut être bloquée par des anticorps dirigés contre le récepteur TLR-2. Ces études montrent également qu'une température de 37°C est nécessaire pour observer ce phénomène ainsi qu'une différence de production de GPLs entre la forme rugueuse et lisse. De plus, la souche 390S débarrassée de ces GPLs au moyen d'isopropanol parvient à induire la production de TNFα par les macrophages dérivés de monocytes via TLR-2. En recherchant les composés inflammatoires parmi les autres glycolipides pariétaux, elles démontrent que des billes recouvertes de phosphatidyl-inositol dimannosides (PIM_2) provoquent la production de TNFα contrairement à celles recouvertes de GPLs. L'abondance de ces

derniers dans l'enveloppe de la forme lisse limite donc l'effet pro inflammatoire des PIM_2, ce qui ne peut pas être le cas pour la forme rugueuse.

Puisque *Mabs* provoque principalement des maladies pulmonaires, son effet a été évalué sur les cellules alvéolaires épithéliales A549[181]. Ces dernières produisent une plus grande quantité d'ARN_m codant la défensine HβD2 qui est un peptide antimicrobien, en présence des souches 390R et 390V qu'en présence de 390S. La délétion du gène *MmpL4b* chez cette dernière, participant au transport des GPLs, entraine l'augmentation de l'expression de l'ARN_m d'HβD2. L'implication de TLR-2 a également été vérifiée lors de cette réponse, ainsi que sur la production d'IL-8 induite par la forme rugueuse et le mutant *ΔMmpL4b*. Or, la cytokine IL-8 est reconnue pour contribuer au recrutement des neutrophiles sur le site d'infection. Même si cette dernière étude ne caractérise pas les partenaires moléculaires de TLR-2, elle renforce l'idée que la forme lisse est une entité faiblement activatrice du système immunitaire inné. Le phénotype du mutant *ΔMmpL4b* se révèle également moins motile avec une diminution de la capacité à former de biofilms. De plus, il stimule fortement le récepteur TLR-2, ce qui avait été démontré précédemment, mais il est capable de se multiplier rapidement dans les macrophages humains[182].

Une comparaison de la réponse des neutrophiles soumis à *Staphylococcus aureus* (*Sa*) ou à la forme lisse de *Mabs* indique que cette dernière provoque une réponse moins exacerbée que *Sa*. En effet, la libération d'anion superoxide et la régulation des gènes codant les cytokines ou chémokines est moins élevée dans le cas de *Mabs*. D'autre part, parmi les cytokines, seule la sécrétion d'IL-8 se maintient à un niveau équivalent. C'est également le cas des chémokines CCL3, CCL4 et CXCL2 qui favorisent la réaction pro-inflammatoire, en particulier le recrutement des cellules du système immunitaire adaptatif. La proportion de bacilles de *Mabs* tués par les neutrophiles est moindre comparativement à *Sa* ainsi que leur phagocytose par ces mêmes cellules. Enfin, grâce à 2 tests *in-vitro*, les auteurs démontrent que la coexistence des bacilles de *Mabs* et des neutrophiles favorisent la formation de biofilm[183]. D'ailleurs, l'exploration des poumons de patients atteints de mucoviscidose confirme la présence de la forme lisse de *Mabs* sous forme de biofilms[184].

L'utilisation de l'embryon du poisson-zèbre doté d'un système immunitaire inné, montre que la forme rugueuse des trois sous-espèces du complexe MABSC dégrade la survie et contribue à la formation de nombreux abcès au niveau du système nerveux central, comparativement à la forme lisse. L'apparition des abcès est due à la formation de cordes serpentiformes uniquement dans le cas de la forme rugueuse ainsi que du mutant de la forme lisse *ΔMmpL4b*. En effet, les macrophages

phagocytent de façon équivalente les bacilles de la forme lisse et rugueuse. Dans le cas de la forme lisse, les macrophages contribuent préférentiellement à la formation de granulomes et à l'installation d'une infection chronique. A l'inverse, dans le cas de la forme rugueuse, puisque les bacilles induisent davantage l'apoptose, leur libération extracellulaire favorise la croissance ainsi que la formation des cordes qui ne peuvent plus être contrôlées par la phagocytose[185].

En poursuivant l'étude de la physiopathologie de *Mabs* lors de l'infection de l'embryon du poisson zèbre, Bernut et collaborateurs confirment que les macrophages initient l'expression du gène codant le TNFα plus fortement en présence de la forme rugueuse que de la forme lisse[186]. Cette induction se produit dans les cellules présentes au site d'infection et dans le granulome mais pas dans les abcès. En parallèle, ces études démontrent que l'absence du récepteur du TNFα aboutit à la formation plus rapide d'abcès et à une mortalité accrue pour les deux formes. Dans ce cas, les macrophages présentent une charge mycobactérienne plus élevée puisqu'ils ont une moindre capacité à tuer les bacilles *via* la libération d'espèces réactives de l'oxygène. La voie de signalisation du TNF permet aussi le recrutement des neutrophiles avec l'IL-8, et le maintien du granulome, limitant la sévérité de l'infection.

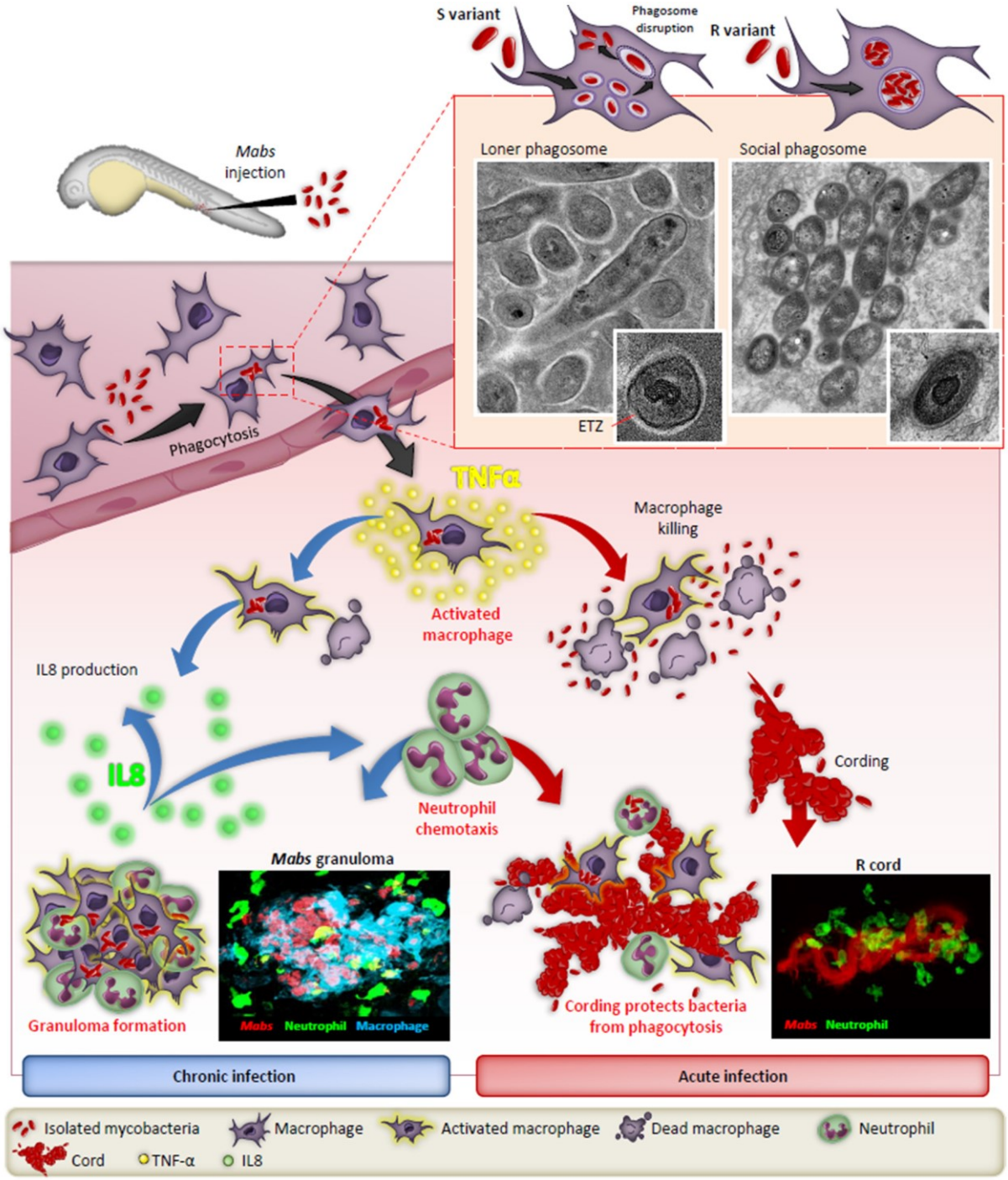


Figure 4 : Interaction des morphotypes lisse et rugueux de Mabs avec les macrophages et au sein de l'embryon de poisson zèbre d'après Bernut[79]. La forme lisse est davantage associée à une infection chronique caractérisée par la présence de granulomes. La forme rugueuse hypervirulente provoque une forte inflammation ainsi que des cordes qui ne peuvent pas être prises en charge par les macrophages.

Les deux formes de *Mabs* sont phagocytées par les macrophages murins et s'y multiplient autant, mais elles génèrent des phagosomes d'aspect différent en microscopie électronique[187]. Alors qu'un seul bacille de la forme lisse par vésicule s'observe majoritairement, ceux de la forme rugueuse

sont plusieurs par phagosome. D'autre part, la membrane du phagosome est intimement accolée aux bacilles de la forme rugueuse mais beaucoup plus éloignée de celui de la forme lisse. Cette dernière empêche l'acidification du phagosome et endommage sa membrane. De cette manière, elle échappe à la dégradation par le phagolysosome. A contrario, la forme rugueuse précipite l'apoptose des macrophages THP-1 s'autorisant ainsi la possibilité de se disséminer ou de former les cordes serpentiformes extracellulaires.

Très récemment, le poisson-zèbre adulte a été utilisé pour tester l'implication de l'immunité adaptative dans la réponse aux infections par *Mabs*[188]. La charge mycobactérienne de la forme lisse est supérieure après 28 jours comparativement à la forme rugueuse. En outre, pour ce dernier, la formation du granulome serait plus rapide. Ils montrent également que la voie de signalisation du TNFα protège de la forme lisse mais favorise le développement de la forme rugueuse. Par contre, les lymphocytes T retrouvés dans les granulomes sont plus impliqués dans le contrôle de la prolifération de la forme rugueuse par rapport à la forme lisse.

Lors de l'infection, la forme lisse de *Mabs* déjoue la plupart des mécanismes de défense immunitaire de l'organisme pour se maintenir sous forme de granulomes disséminées ou de biofilms, et sa persistance multiplie le risque d'apparition d'une forme isogénique rugueuse dont l'hypervirulence débouche sur une réaction inflammatoire aigüe, des abcès nombreux et incontrôlables dont l'issue est létale dans 50% des cas (Figure 4). C'est pourquoi le choix d'un traitement rapide et approprié est primordial.

I.A.3.e.viii Traitement des infections associées à *Mabs*

Conduite thérapeutique

D'après l'Infectious Disease Society of America (IDSA), les multiples résistances développées par *Mabs* au cours de son évolution nécessite l'utilisation d'au moins trois antibiotiques[189]. La durée de traitement qui se divise en deux phases s'étale sur plusieurs mois, mais elle doit être adaptée au statut du patient. En effet, une amélioration après 2 à 4 mois n'est pas synonyme d'une guérison complète et doit faire l'objet de cultures négatives pendant 6 à 12 mois[190]. La figure 5 présente la conduite thérapeutique préconisée par la « British Thoracic Society » selon l'issue du test de susceptibilité aux macrolides et de la fonctionnalité du gène *erm(41)*. Compte tenu de la longueur du traitement et des effets secondaires engendrés, la recherche d'alternatives utilisant des molécules thérapeutiques existantes se poursuit.

Nouveaux traitements

De nombreuses molécules sont en cours d'évaluation contre les infections provoquées par *Mabs* . Par exemple, l'omadacycline est une tetracycline qui montre une action synergique significative avec la clarithromycine chez le poisson zèbre[191]. L'oritavancine est un peptide antibactérien qui possède également une action synergique en particulier avec la cefoxitine ou le meropenem, contribuant à diminuer la charge bactérienne dans les poumons de souris, sans être cytotoxique pour les macrophages THP-1[192]. L'utilisation de molécules qui ciblent le transporteur d'acides mycoliques MmpL3 n'ont jamais été évaluées chez l'homme à ce stade de leur développement, mais n'en demeurent pas moins des candidats prometteurs au traitement des mycobactéries au sens large. Enfin, l'utilisation des phages est une stratégie qui permet de cibler spécifiquement les MNTs ainsi que *Mabs* tout en adaptant la charge virale à la charge mycobactérienne. En 2019 et 2022, Dedrick décrit 2 situations cliniques dans lesquelles cette méthode est utilisée avec succès : la 1ère concerne un adolescent atteint de mucoviscidose et transplanté des deux poumons pour lequel un cocktail de trois phages a été administré, faisant diminuer la proportion de nodules dans l'ensemble du corps[193] ; la seconde concerne l'inhalation de phages chez un patient adulte immunocompétent atteint par *Mabs* et ayant développé des anticorps contre les phages, contribuant à l'amélioration de son état global[194].

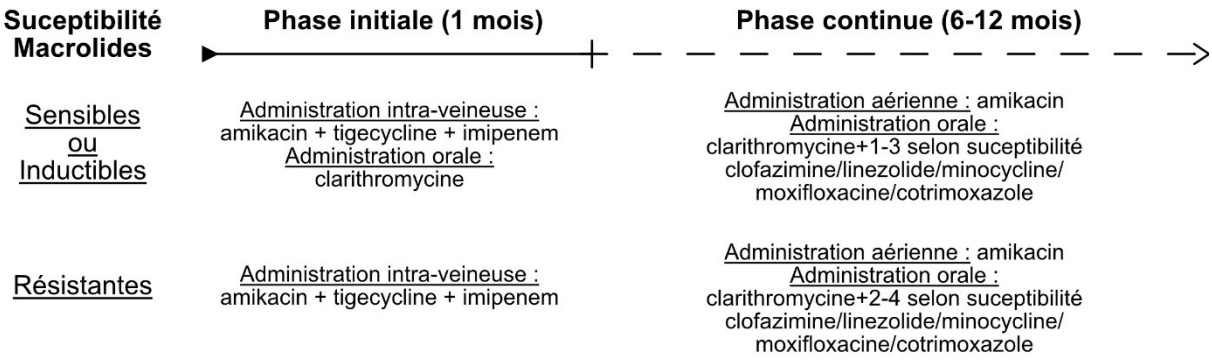


Figure 5 : Conduite thérapeutique préconisée par l'IDSA[186] suite au diagnostic d'une infection pulmonaire associée à l'une des 3 sous-espèces du complexe MABSC et selon la détermination de la suceptibilité à la clarithromycine ou autres antibiotiques.

I.B ORGANISATION DE L'ENVELOPPE MYCOBACTERIENNE

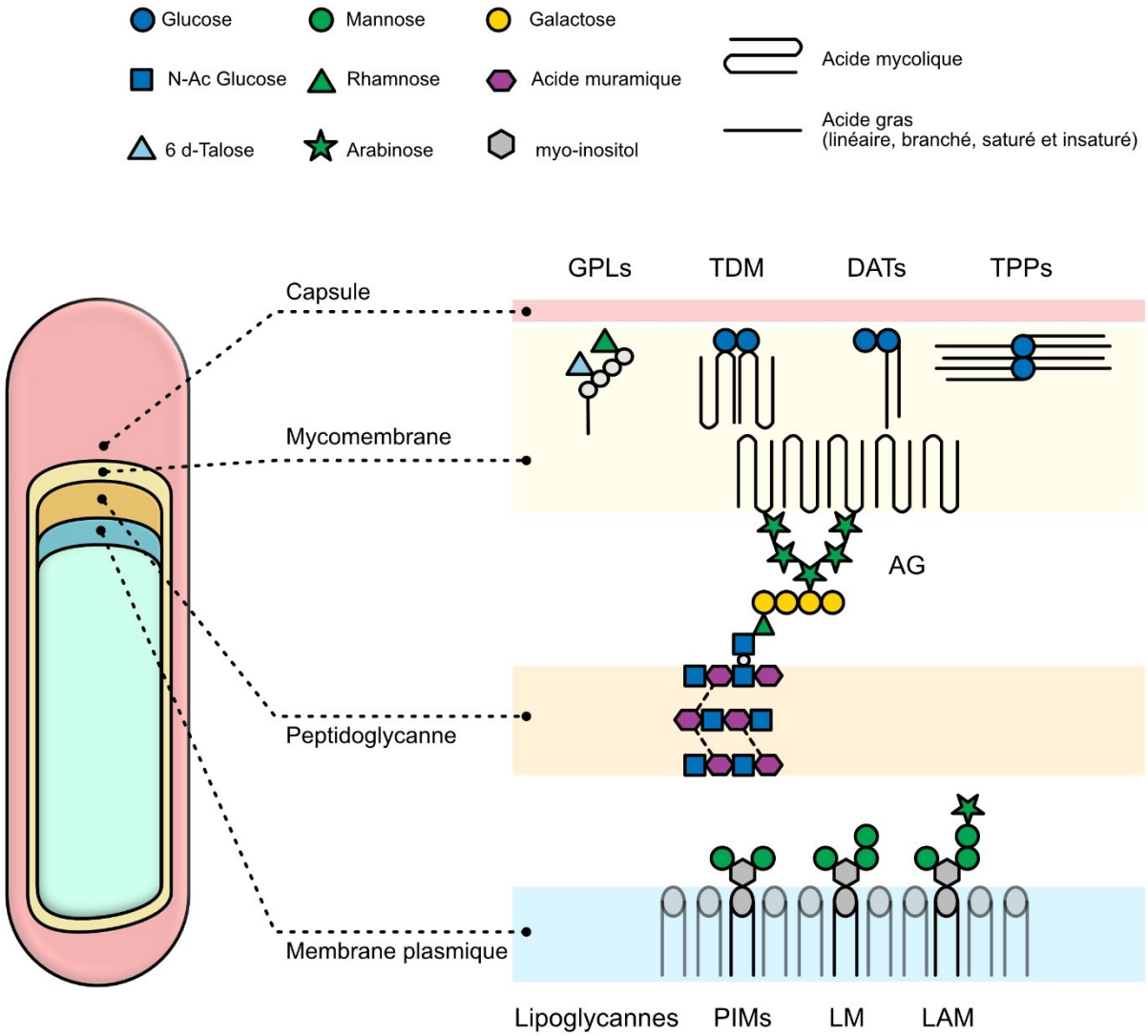


Figure 6 : Schéma de la paroi des mycobactéries, de ses compartiments et des structures glycolipidiques décrites chez *Mabs* dans la litérature[195],[196],[197]. Les différents monosaccharides sont repris de la nomenclature SNFG mais leur forme furane ou pyrane, l'anomérie et la position de leurs liaisons ne sont pas précisées. De même, la structure des acides mycoliques et des lipides n'est pas détaillée. (GPLs = glycopeptidolipides ; TDM = tréhalose dimycolate ; DATs = diacyltréhaloses ; TPPs = tréhalose polyphléates ; AG = arabinogalactane ; PIMs = phosphatidylinositol mannosides ; LM = lipomannane ; LAM = lipoarabinomannane). Les LOSs et les PGLs ne sont pas indiqués.

L'enveloppe des mycobactéries est une structure complexe subdivisée en plusieurs feuillets et constitutée d'un grand nombre de glycolipides [198],[199],[196]. La membrane plasmique est le feuillet le plus interne correspondant à une double couche phospho-lipidique dans laquelle sont insérés les lipoglycannes (Figure 6). Ils regroupent les phosphatidylinositol mannosides (PIMs), le lipomannane (LM) et le lipoarabinomannane (LAM). C'est également dans la membrane plasmique que sont enchâssées les protéines MmpL. Le périplasme est l'espace entre cette membrane interne et le

squelette mycolyl-arabinogalactane-peptidoglycanne. Côté interne, le peptidoglycanne commun à de nombreuses bactéries, entoure et protège la mycobactérie, et le polymère d'arabinogalactane y est lié de façon covalente, maintenant l'intégrité architecturale de l'enveloppe. Des acides mycoliques à longue chaîne sont estérifiés sur l'arabinogalactane formant le feuillet interne de la mycomembrane. Son feuillet externe est constitué de glycolipides dérivés d'acides mycoliques, principalement le TDM dont la partie hydrophile se tourne vers l'extérieur. Les acides mycoliques sont spécifiques de l'espèce et en partie responsables de la très faible perméabilité aux petites molécules hydrophiles. Par exemple, la perméabilité aux céphalosporines de *Mche* est respectivement 3 fois et 10 fois inférieure à celle d'*Escherichia coli* et de *Pseudomonas aeruginosa*. Les autres glycolipides peuvent également être dérivés du tréhalose comme les acyl-tréhaloses, les tréhaloses polyphléates (TPPs)[200] ou les lipooligosaccharides (LOSs)[195]. Certains contiennent principalement des déoxyhexoses comme les glycopeptidolipides (GPLs)[201] ou les phénolglycolipides (PGLs)[202]. Enfin, selon les espèces, on retrouve également une capsule faiblement liée à la paroi de la mycobactérie et entourant celle-ci[203]. Elle contient essentiellement des glycannes de structure similaire au glycogène et aux lypoglycannes ainsi que des protéines qui ne seront pas traités dans les chapitres suivants[204].

I.B.1 LA MEMBRANE PLASMIQUE ET LES LIPOGLYCANNES

I.B.1.a LES PHOSPHATIDYL-INOSITOL MANNOSIDES (PIMs)

Les glycoconjugués pariétaux composés de mannose sont essentiels à la survie de la lmycobctérie. En effet, la délétion du gène codant la phosphomannose isomérase chez la souche *Msmeg* mc^{2}155 interrompt la croissance sur un milieu sans mannose[205]. Ce phénotype s'accompagne d'une diminution de la quantité dans le cytoplasme des méthylmannosides matures, et entraine également une modification de la répartition des PIMs en faveur de ceux ayant une taille intermédiaire. De même, la délétion du gène *ino1* de *Msmeg*, responsable de la conversion du glucose 6-phosphate en inositol, aboutit à une souche auxotrophe pour cette source de carbone[206]. Elle demeure capable de croitre sur un milieu enrichi en inositol mais lorsque ce dernier est absent du milieu de culture, la viabilité chute avec la quantité détectée de protéines. La cardiolipine et la phospho-éthanolamine ne sont pas impactés alors que les GPLs et l'arabinogalactane (AG) diminuent bien qu'ils ne contiennent pas d'inositol. Plus directement, l'inositol n'est plus catabolisé et ainsi renouvelé pour soutenir la biosynthèse des PIMs.

La biosynthèse des PIMs à partir d'un précurseur di-*O*-acyl glycéro-phosphatidyl-*myo*-inositol ancré dans la membrane plasmique est proposée par Ballou et collaborateurs dans les années 1960

[207]. L'étude approfondie de la forme pentamannosylée indique que les deux premiers α-D,mannopyranoses sont reliés aux positions 2 et 6 du *myo*-inositol[208], donnant respectivement le phosphatidyl-inositol monomannoside (PIM_1) et le phosphatidyl-inositol dimannosides (PIM_2). Puisqu'ils sont ancrés dans la membrane plasmique, les PIMs ne font pas l'objet d'un transport par des protéines membranaires de type MmpL. Les produits des gènes *Pim* participent à l'ajout des résidus de mannose et les gènes codant les protéines *Pat* sont responsables du transfert de groupements acyles (Figure 7).

I.B.1.a.i Biosynthèse des PIMs

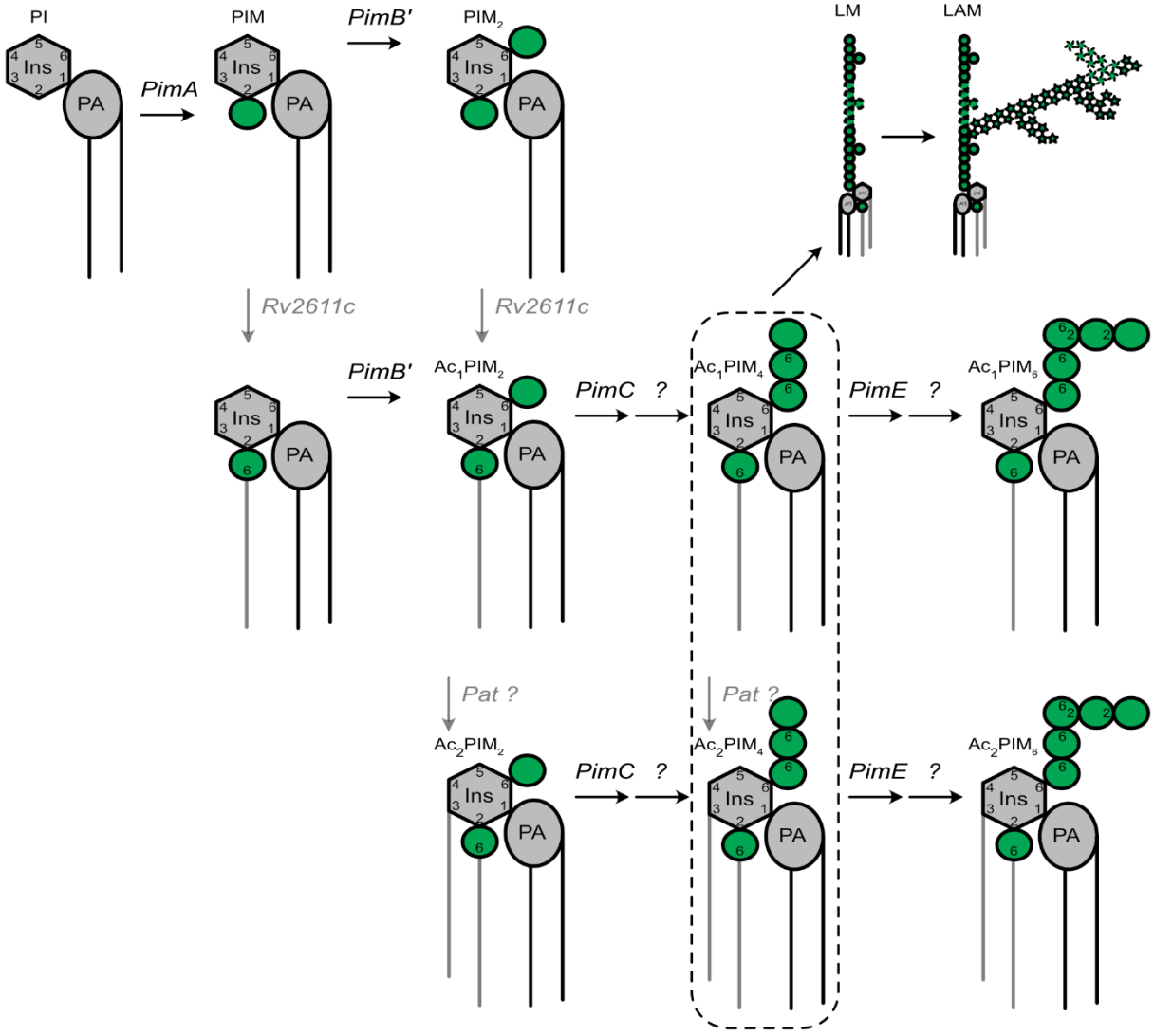


Figure 7 : Schéma général de biosynthèse des PIMs indiquant les enzymes impliquées dans l'ajout des résidus de mannopyranoses. La position porteuse de la liaison glycosidique est indiquée ainsi que l'anomérie lorsqu'elle est dans l'orientation β. Une seule acétyltransférase PatA codée par le gène *Rv2611c* chez *Mtb* est connue. A partir des PIM_4, 2 voies vers la biosynthèse des PIM_6 ou du LM et du LAM sont possibles.

Ajout successif des mannoses

Chez *Msmeg*[209] et *Mtb*[210], le produit du gène *PimA* participe à l'obtention du PIM_1 par l'addition d'un résidu d' α-D,mannopyranose en position 2 du *myo*-inositol. Lorsque l'expression de ce gène est interrompue, la croissance et la viabilité du bacille diminue significativement non seulement en culture mais aussi à l'intérieur des macrophages. Ces observations soulignent le caractère essentiel de ce gène aussi bien dans la physiologie des mycobactéries, que dans la virulence des espèces pathogènes. Dans la 1ère étude, c'est la surexpression chez *Msmeg* ou l'expression hétérologue du gène *PimA* chez *Escherichia coli* qui aboutit à la détection par autoradiographie du PIM_1 marqué grâce au donneur ^{14}C [GDP-Man]. *PimB'* est le gène codant l'enzyme qui relie exclusivement le 2ème résidu d' α-D,mannopyranose à la position 6 du *myo*-inositol du PIM_1[211]. Dans cette étude, les auteurs insistent sur le fait que l'ajout des deux premiers résidus de mannose doit se faire d'une manière ordonnée. En effet, en présence de phosphatidyl *myo*-inositol *in vitro*, la vitesse d'obtention du PIM_1 est plus faible avec PimB' qu'avec PimA. De plus, la quantité de PIM_2 détectable est négligeable lorsque PimB' agit avant PimA alors que la réaction inverse aboutit à des quantités équivalentes de PIM_1 et PIM_2. A nouveau, le gène *PimB'* est essentiel puisque sa délétion conditionnelle entraine un défaut de croissance.

L'une des spécificités structurales du PIM_2 de *Mabs* est la présence d'un groupement méthyl en position 2 sur ce 2ème résidu de mannose[212]. Cette modification se répercute sur le lipomannane (LM) et le lipoarabinomannane (LAM) mais son origine métabolique et son rôle dans la suite de la biosynthèse des PIMs ne sont pas établis.

La 3ème mannosyltransférase est codée par le gène *PimC* qui transfère l'α-D,mannopyranose sur la position 6 du précédent. La mutation de ce gène chez *Mtb* n'empêche pas la production de LM et de LAM et n'entrave pas la croissance du bacille. De plus, la distribution du gène *PimC* parmi des souches cliniques de *Mtb* ne s'observe pas dans la totalité des génomes, suggérant une voie alternative de biosynthèse[213]. L'obtention de la forme tétramannosylée (PIM_4) via une seconde liaison α1-6 requiert l'action d'une mannosyltransférase qui n'a toujours pas été caractérisée.

C'est à partir du PIM_4 que bifurquent les voies de biosynthèse du PIM_6 d'une part et du LM puis du LAM, d'autre part, faisant intervenir deux nouvelles mannosyltransférases. La 1ère, PimE, est codée chez *Msmeg* par le gène *MSMEG_5136* [214]. Sa délétion entraine la disparition de l'Ac_1PIM_6 et de l'Ac_2PIM_6 au profit de l'Ac_1PIM_4 et de l'Ac_2PIM_4 après une analyse simultanée CCM et MALDI-TOF[215]. La 2ème fut proposée chez *Mmar* et appelée PimF[216]. Le gène codant cette enzyme possèdait 78%

d'homologie avec le gène *Rv1500* retrouvé chez *Mtb*. Sa délétion entrainait non seulement l'accumulation d'Ac_4PIM_5 mais également la diminution de la quantité de LAM. Néanmoins, ce travail a été remis en question et le gène maintenant appelé *LosA* est réattribué à la biosynthèse des lipooligosaccharides (LOSs)[217]. Enfin, c'est le polyprénolmannose phosphate qui est le substrat donneur de mannose pour l'obtention du PIM_5 et du PIM_6 alors que le GDP mannose est le substrat donneur lors des réactions catalysées par PimA et PimB'.

Composition de l'ancre lipidique

Puisque les PIMs sont construits à partir d'un acide phosphatidique, les deux premiers acides gras présents sont intégrés dans le di-*O*-acylglycérol, contenant le plus souvent au moins un acide 10 méthyl stéarique ou tuberculostéarique. L'acyltransférase PatA produite chez *Msmeg* à partir du gène *Rv2611c* de *Mtb* est essentielle. En effet, son absence contribue à limiter les capacités de croissance sur le milieu LB. Cette enzyme catalyse l'estérification d'un acide gras sur la position 6 du résidu de mannose ajouté par PimA. A partir des PIM_1 et PIM_2, elle produit ainsi les formes tri-acylées Ac_1PIM_1 et Ac_1PIM_2. Lorsque l'expression du gène *Rv2611c* est interrompue, le niveau général d'acylation des PIMs diminue sans disparaître, supposant l'existence d'une enzyme qui compense son absence[218].

En 1995, Khoo et collaborateurs ont démontré que le résidu d'acide gras dans l'Ac_1PIM_2 est porté par l'un des deux résidus de mannose [219]. Ils ont également déterminé que des acides lysophosphatidiques sont parfois présents dans l'Ac_1PIM_2 et que ce dernier est présent dans le lipommannane (LM) et le lipoarabinomannane (LAM). La présence d'une structure commune dans les trois lipoglycannes renforçait ainsi leurs liens biosynthétiques. La dernière acylation est celle de l'inositol en position 3 qui a d'abord été observée dans les phosphatidyl-inositol dimannosides tétra-acylées (Ac_2PIM_2) et le LM de *Mbov* BCG[220]. Les acides gras palmitique, stéarique et tuberculostéarique sont les quatre lipides estérifiés sur le glycérol, le *myo*-inositol et le mannose. L'analyse par spectrométrie de masse MALDI-TOF[8] et ESI-IT[221] confirme que parmi les formes mannosylées de PIMs, l'acide myristique et l'acide stéarique sont minoritaires alors que l'acide palmitique et l'acide tuberculostéarique sont majoritaires. Néanmoins, l'enzyme responsable de cette modification n'a toujours pas été décrite.

I.B.1.a.ii Rôles biologiques des PIMs

Malgré la présence de grande quantité de PIMs dans l'enveloppe mycobactérienne et le caractère essentielle des 1ères enzymes de leur biosynthèse, le rôle de ces composés dans la physiologie de la mycobactérie n'est pas clair[222]. Avec les phospholipides, ils sont les constituants majeurs de la

membrane plasmique et modulent sa perméabilité. Les formes tri-acylées du PIM_1 et du PIM_2 sont d'ailleurs nécessaires à la survie. Les PIMs peuvent diffuser passivement hors de la membrane plasmique[223] et ils contribuent ainsi à la physiopathologie de l'interaction hôte-pathogène. Bien qu'ils ne soient pas exclusifs aux espèces pathogènes, les PIMs sont des agonistes des récepteurs TLR-2 qui activent l'immunité innée[180]. Les PIMs fortement mannosylés lient préférentiellement le récepteur à mannose. De cette manière, les PIM_5 et PIM_6 sont capables de participer à l'inhibition de la fusion entre phagosome et lysosome[224]. Le récepteur DC-SIGN possède une forte affinité pour le PIM_6 comparable au lipoarabinomannane coiffé de mannose (ManLAM) de *Mtb*. Par contre, l'arrêt de leur production par *Mbov BCG* n'a pas d'effet sur la reconnaissance des bacilles par ce même récepteur et la production de cytokines qui s'ensuit.

I.B.1.b LIPOMANNANE (LM)

Le lipomannane résulte de l'élongation du PIM_4 par des unités (1-6)α-D-mannopyranoses (Figure 8). Les gènes codant les enzymes MptA et MptB respectivement responsables de l'allongement de la partie distale et proximale de la chaine mannane ont été caractérisés *in vitro* ainsi que chez *Corynebacterium glutamicum (Cglu)*. Ils possèdent des orthologues *Rv2174* et *Rv1459c* chez *Mtb* non identifiés chez *Mabs*. Le rôle de MptB pourrait être assuré par une protéine redondante[196]. La chaine mannane obtenue présente des points de branchement d'un seul résidu α-D-mannopyranose qui se font en position 2 chez *Mbov BCG*[220] mais en position 3 chez *Mche*[225]. Cette caractéristique structurale fut confirmée récemment chez *Mabs*. De plus, d'après le spectre de masse de sa forme déacylée, la longueur de la chaine mannane du LM varie de 32 à 44 résidus, ce qui est supérieure à celle de *Msmeg*[226]. Par contre, l'espacement entre les points de branchement n'a pas pu être clairement défini. De même, le rôle de l'enzyme MptC qui permet le branchement en position 2 dans le LM de *Cglu* ainsi que son orthologue potentiel ne sont pas clairement identifié chez *Mabs*.

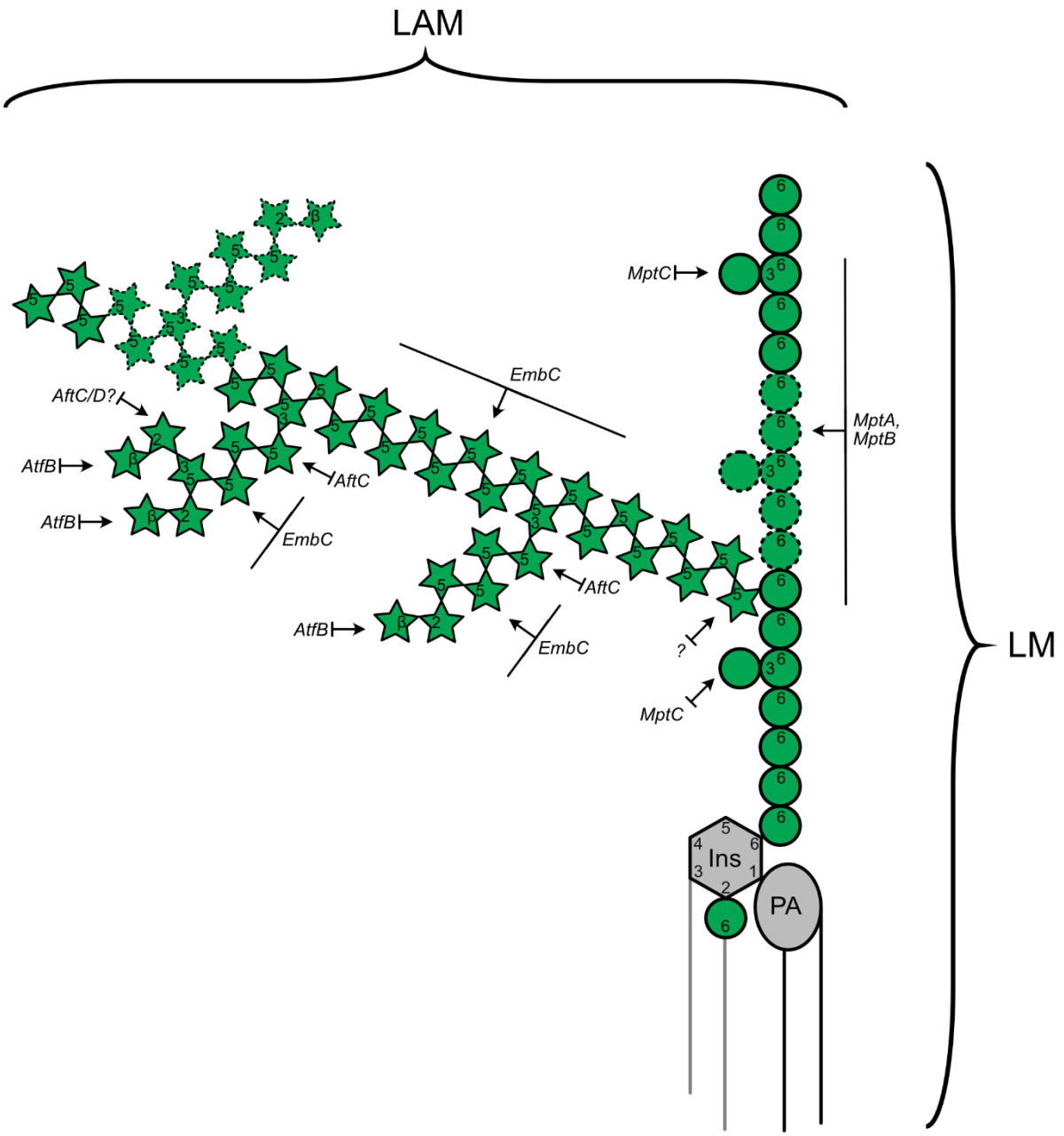


Figure 8 : Schéma général de l'enchainement oligosaccharidique du LM et du LAM indiquant les enzymes impliquées dans l'ajout des résidus de mannopyranoses et d'arabinofuranoses. La position porteuse de la liaison glycosidique est indiquée ainsi que l'anomérie lorsqu'elle est dans l'orientation β. Seule la forme AraLAM propre à *Mabs* est schématisée alors que les coiffes mannose (ManLAM) et phosphatidylinositol (PILAM) ne sont pas représentées.

I.B.1.c LIPOARABINOMANNANE (LAM)

Dans le LAM, les arabinoses sont exclusivement sous forme furanose (Figure 8). L'initiation de la chaine arabinane sur la chaine mannane n'a pas été décrite à ce jour, mais elle se fait dans la chaine linéaire de mannane sur des résidus de mannoses non branchés. Par contre, la seconde arabinofuranosyl-transférase responsable de l'élongation de la chaine linéaire d'arabinane est

l'enzyme EmbC également impliquée dans la formation de l'AG chez *Msmeg*[227]. Elle catalyse la formation d'une liaison glycosidique α(1-5). Deux autres enzymes sont communes avec la biosynthèse de l'AG, AftC et AftB. AftC introduit des points de branchement en position 3 conservant l'anomérie α[228], tandis qu'AftB coiffe les ramifications non réductrices d'un dernier résidu en position 2 avec une anomérie β[229]. La ramification terminale du LAM est constituée de deux structures caractéristiques tetra-arabinofurannosides linéaires ou hexa-arabinofurannosides bi-antennées. Leur répartition chez *Mabs* est égale, mais la proportion de points de branchements succinylés est supérieure à celle des points de branchements acétylés dans les deux structures[212]. En effet, le gène *sucT* responsable de la succinylation a été d'abord identifié chez *Msmeg*[226] puis confirmé chez *Mabs*.

Enfin le LAM peut être coiffé par 3 motifs distincts selon l'espèce :

- ManLAM : L'enzyme CapA transfère un premier résidu d'α-D-mannopyranose en position 5 de l'arabinose terminal dans le cas de *Mtb*[230] ; après cela l'enzyme de branchement de la chaine mannane MptC transfère deux résidus d'α-D-mannopyranose relié en (1-2)[231].
- PILAM : Chez *Msmeg*, un phosphoinositol est transféré à l'extrémité non réductrice sur la position de l'arabinofuranose[232];
- AraLAM : La coiffe terminale demeure le résidu de β-D-arabinofuranose lié en (1-2)[225].

La découverte d'une unité de méthylthioxylofuranose (MTX) qui s'ajoute à chaque molécule de LAM sur une unité d'arabinofuranose terminal chez *Mtb*[233], ou sur la chaine mannane de *Mycobacterium kansasii (Mkan)*[234] est la dernière originalité structurale récemment décrite chez les mycobactéries. L'analyse RMN des produits purifiés met également en évidence deux formes oxydées diastéréoisomériques en plus du MTX.

Enfin, le LAM n'est pas restreint à la membrane plasmique et peut être exporté par le produit du gène *LprG* qui est une lipoprotéine possèdant une forte affinité pour le ManLAM [235 236].

I.B.1.d ROLES BIOLOGIQUES DU LM ET DU LAM

Le LM et le LAM contribuent à maintenir l'intégrité de l'enveloppe et à réguler la perméabilité membranaire. En effet, les mutants de *Msmeg* et de *Mtb* dont les LM et LAM ont des structures non branchées ou raccourcies peuvent présenter un défaut de coloration et une plus grande susceptibilité aux β-lactamines ou aux antibiotiques. A l'inverse, la surexpression des gènes normalement impliqués dans la biosynthèse ne rétablit pas systématiquement le phénotype sauvage[237]. Un défaut de succinylation altère la surface de la paroi de *Msmeg*[226] et limite la survie intracellulaire de *Mabs*[212].

D'autre part, le rôle immunomodulateur du LAM dans les infections par *Mtb* a été proposé via sa propension à lier les « Pattern Recognition Receptor » (MR, DC-SIGN, TLRs). En particulier la liaison du ManLAM au récepteur à mannose des macrophages limite la maturation du phagosome, le niveau d'oxydation, l'inflammation et l'apoptose. Malgré tout, les effets provoqués par le ManLAM avec ou sans la coiffe mannose sur la réponse immunitaire mettent uniquement en évidence l'impact sur l'apoptose des macrophages[238]. Enfin, l'étude approfondie du résidu MTX oxydable sous l'action du couple d'enzyme methionine sulfoxide reductase MsrA/MsrB sous-tend le rôle de protection envers les espèces réactives de l'oxygène[239].

I.B.2 LE PERIPLASME ET LE COMPLEXE MYCOLYL ARABINOGALACTANE PEPTIDOGLYCANNE (mAGP)

Le périplasme est l'espace délimité par la membrane plasmique côté interne et la mycomembrane côté externe. C'est un compartiment bactérien stratégique qui circonscrit le déroulement de nombreux processus biologiques dont la biosynthèse du complexe mAGP.

I.B.2.a SPECIFICITES DU PEPTIDOGLYCANNE DES MYCOBACTERIES

Le peptidoglycanne est comparable à un exosquelette qui préserve l'intégrité de l'enveloppe, en particulier lors de stress osmotique. Cette structure est commune à toutes les bactéries Gram+, Gram- ainsi qu'à la famille des mycobactériacées. Le peptidoglycanne est formé de chaines polysaccharidiques constituées d'alternances d'unités *N*-acétyl glucosamine et d'acides *N*-acétyl muramiques en β(1-4) et reliées entre elles par des liaisons amines *via* des tétrapeptides L-alanyl-D-isoglutaminyl-meso-diaminopimelyl-D-alanine[196]. L'une des originalités propres aux mycobactéries est la possibilité d'oxydation du résidu *N*-acétyl en *N*-glycolyl[240]. La liaison peptidique entre les 2 peptides s'établit soit entre le COOH de la D-alanine et le NH_2 de l'acide diaminopimélique (liaison 3-4) soit entre 2 résidus d'acides diaminopiméliques (liaison 3-3). Dans le cas des mycobactéries le taux de branchements entre les chaines oligosaccharidiques est particulièrement élevé pouvant atteindre jusqu'à 80% en phase stationnaire de croissance[241]. De plus, chez les mycobactéries, 10 à 12% des résidus de *N*-glycolyl acides muramiques sont modifiés en position 6 par un groupement α-L-rhamnopyranosyl(1-3)N-acétyl-α-D-glucopyranosylamine-1-phosphate qui constitue le point d'ancrage de la chaine arabinogalactane (AG)[242].

I.B.2.b L'ARABINOGALACTANE (AG)

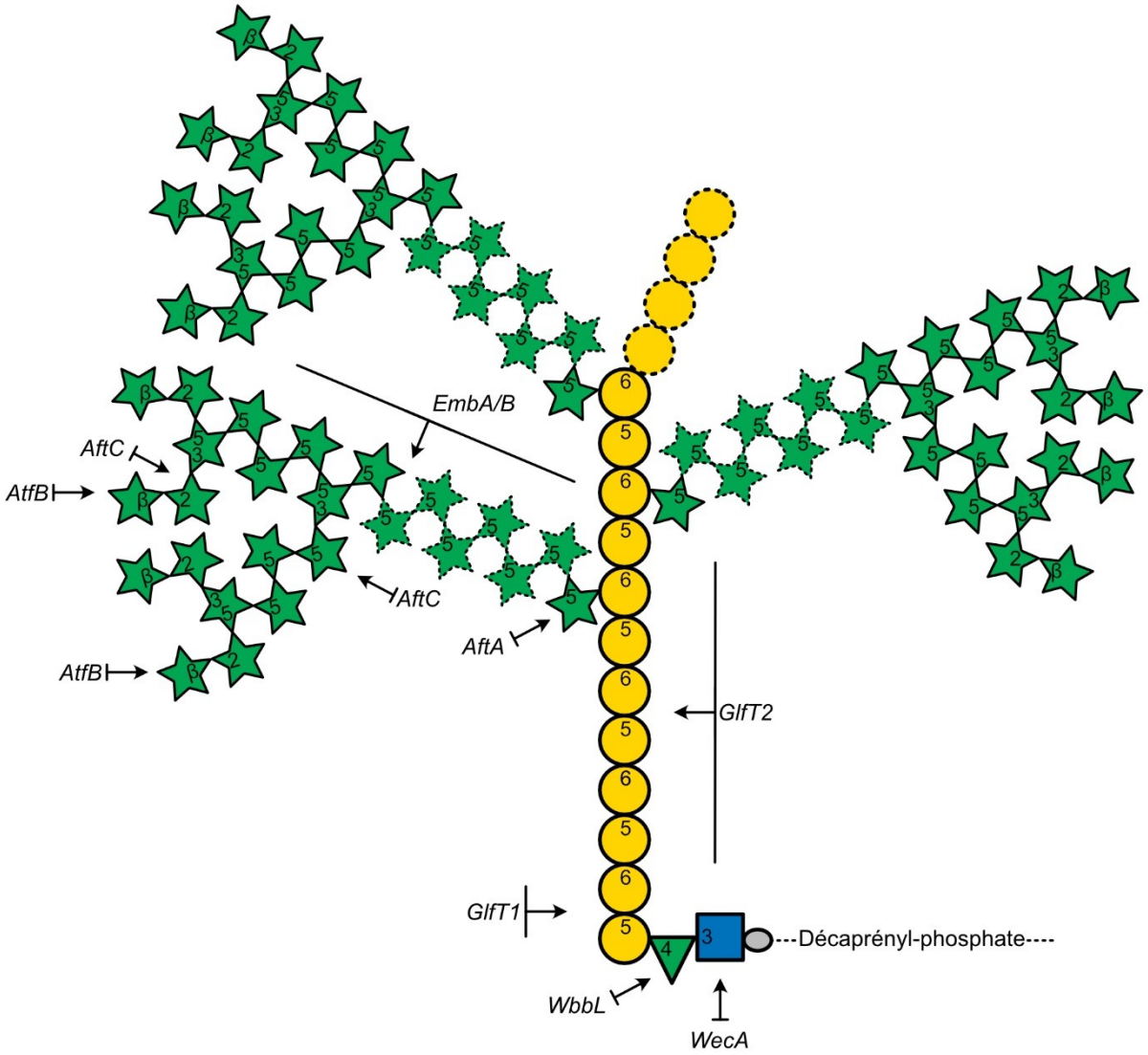


Figure 9 : Schéma général de l'arabinogalactane d'après Jankute[194] indiquant les enzymes impliquées dans sa biosynthèse. Celles de la chaine arabinane sont communes avec le LAM. Tous les résidus de galactofuranoses sont reliés avec une anoméire β et ceux d'arabinofuranoses avec une anomérie α, sauf le résidu terminal côté non réducteur. L'estérification des acides mycoliques sur ces derniers n'est pas illustrée.

I.B.2.b.i La chaine galactane

Egalement décrite chez les corynébactéries, l'arabinogalactane est synthétisée exclusivement à partir de formes furanoses (Figure 9). La chaine galactane provient d'un enchainement d'unités galactofuranoses reliés alternativement en β(1-5) et β(1-6)[243]. A partir du point d'ancrage sur le rhamnose, l'enzyme GlfT1 ajoute les deux premiers galactoses[244] alors que GlfT2 poursuit l'élongation jusqu'à une trentaine de résidus[245].

I.B.2.b.ii La chaine arabinane

L'initiation de la chaine arabinane chez *Mtb* est une fonction essentielle à la survie. Chez *Cglu,* il a été observé que les 8ème, 10ème et 12ème galactofuranoses sont liés non seulement en β(1-6) mais aussi en position 5 au premier α-D-arabinofuranose suite à l'action d'AftA[246]. Les arabinosyltransférases responsables de la synthèse du LAM sont impliquées dans l'élongation de l'arabinane du mAGP[243] : EmbA et EmbB participent à l'élongation en produisant des liaisons glycosidiques α(1-5), AftC ou AftD initie des branchements sur les positions 3 et AftB ajoute les résidus terminaux liés en β(1-2) (Figure 9). L'addition d'un groupement succinyl par le produit du gène *SucT* sur la position 2 des arabinofuranoses correspondant aux points de branchement représente un point commun supplémentaire avec la structure du LAM, notamment chez *Msmeg*[226], contrairement à la possibilité d'y observer un résidu de D-galactosamine[247]. Cette modification pourrait être considérée comme une étape déterminante pour la mycoloylation des unités arabinofuranoses terminales. En effet, la chaine arabinane s'organise en motifs hexa-arabinofuranosides sur lesquels les acides mycoliques sont estérifiés pour donner naissance au feuillet interne de la mycomembrane.

I.B.2.c ESTERIFICATION PAR LES ACIDES MYCOLIQUES

Les acides mycoliques ne sont pas exclusifs des mycobactéries et se retrouvent chez les genres *Corynebacterium*[248], *Rhodococcus* ou *Nocardia*… Malgré tout, la première description des acides mycoliques remonte à une étude d'Anderson en 1938 sur la souche H37 de *Mtb*. C'est ensuite Asselineau qui démontre que leur structure générale aujourd'hui acceptée, s'articule autour d'un α-alkyl β-hydroxy acide gras[249] dont le nombre total de carbones varie de 30 à 90 selon le genre.

I.B.2.c.i Structure des acides mycoliques

Chez les mycobactéries (Figure 10), les acides mycoliques sont formés d'une chaine α alkyle saturée de 24 carbones et d'une chaine méromycolique prolongeant la fonction β-hydroxy potentiellement modifiée par une (forme α') ou deux modifications (forme α) en positions proximale puis distale[250],[251]. Deux types de modifications existent : les insaturations et les fonctions oxygénées. Les insaturations sont constituées d'une double liaison éthylénique ou d'une fonction cyclopropyl, chacune pouvant être en configuration *cis* ou *trans*. Dans ce dernier cas, le carbone vicinal n-1 ou n+1 porte un branchement C-méthyl. La fonction proximale est systématiquement porteuse d'une insaturation. La position distale, en plus de la possibilité d'être insaturée, peut être modifiée par une fonction cétone (keto), ester (wax) ou époxy. Un groupement méthoxy a également été identifiée en position ω de la chaine méromycolique, comme décrit chez *Mycobacterium alvei*[252]. Chez *Msmeg*, les

formes diéthyléniques α' cis, α *cis/trans* et époxy *trans* sont présentes, alors que chez *Mtb* seules des formes cyclopropyles α *cis/cis* ainsi qu'hydroxy, méthoxy et keto ont été identifiées, et la forme wax n'a été détectée que chez *Mav*. Les fonctionnalités vont orienter le repliement des chaines hydrocarbonées. Dans le cas de la fonction cétone, elles adoptent une conformation semblable à un W[253].

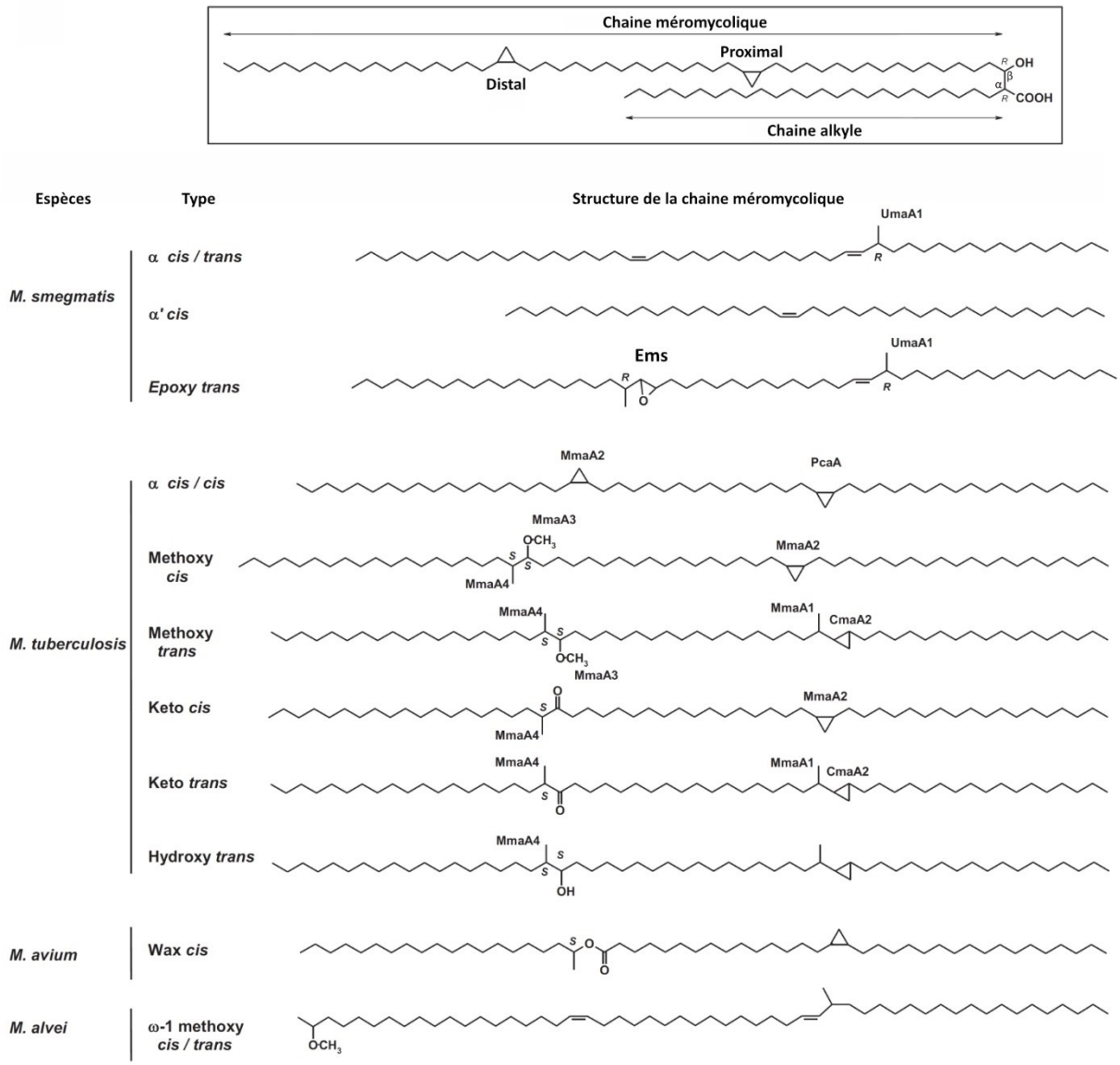


Figure 10 : Structure générale (en haut) et spécificités de la chaine méromycolique (en bas) des acides mycoliques retrouvés chez les principales espèces de mycobactéries d'après Marrakchi[247]. Les enzymes impliquées dans l'obtention des fonctions distale et proximale sont indiquées.

I.B.2.c.ii Biosynthèse des acides mycoliques

Les acides mycoliques sont synthétisés dans le cytoplasme mycobactérien par une réaction de condensation de Claysen catalysée par l'enzyme pks13, d'une chaine méromycolique préalablement modifiée sur un acide gras saturé. Alors que ce dernier est produit par l'acide gras synthase, FAS-I, rencontré chez de nombreux organismes, la chaine méromycolique est issue de l'activité d'une autre acide gras synthase, FAS-II. Ce complexe utilise comme substrats un acyl-CoA produit par FAS-I ainsi qu'un résidu de malonyl pour allonger la chaine carbonée dans un processus itératif de 4 étapes :

- KasA/B condense le résidu de malonyl sur l'acyl-CoA produisant un β céto-acyl ;
- MabA réduit le groupement cétone en fonction hydoxyl ;
- HadA/B élimine l'équivalent d'une molécule d'eau donnant un trans-2-énoyl ;
- InhA réduit la double liaison obtenue suite à la déshydratation.

La très grande diversité structurale des fonctions proximales et distales associées dans la chaine méromycolique ne reflète pas l'éventail des enzymes caractérisées contribuant à sa modification. L'ajout des doubles liaisons est sujet à de nombreuses hypothèses dont aucune ne s'impose aujourd'hui, d'autant que les formes diéthyléniques sont considérées comme précurseurs de toutes les autres formes[254]. Certains travaux s'orientent vers la nécessité d'enzymes codants des désaturases[255].

La synthèse des groupements cyclopropyl et des fonctions oxygénées fait intervenir au moins six enzymes : PcaA[256] ou MmaA2[257] sont responsables de la formation des cis-cyclopropanes, MmaA1 et CmaA2[258] de celle des trans-cyclopropanes et MmaA3[259] ainsi que MmaA4[260] de l'obtention des formes hydroxy, méthoxy et kéto. Ces enzymes assurent parfois des fonctions redondantes[261]. En outre, l'implication de deux déhydratases, HadC et HadD, dans l'élongation de la chaine méromycolique après l'ajout de fonctions oxygénées[262],[262] montre que l'allongement se réalise parfois après la modification de la chaine méromycolique. La différence entre les longueurs de chaine des formes α' et α pourrait se justifier par une activation *via* deux partenaires protéiques alternatifs FadD32 ou FadD32S, caractérisés chez *Msmeg* et *Mabs*[263]. Enfin, la formation des acides mycoliques de forme époxy nécessite l'action d'une « epoxy mycolate synthase » codée par le gène *MSMEG_1350*[264].

I.B.2.c.iii Liaison à la chaine arabinane

Avant d'être transférés sur les arabinofuranoses terminaux, les acides mycoliques sont transportés à la surface bactérienne. Pour cela, ils sont rapidement associés au tréhalose sous la forme de TMM. En effet, ce dernier va être pris en charge par le transporteur MmpL3 ancré dans la membrane plasmique[157]. Dans l'espace périplasmique, deux molécules de TMM peuvent être condensées entre elles pour former le TDM ou utilisées comme donneur d'acide mycolique pour estérifier l'AG. Une comparaison des trois mycolyltransférases de type Ag85 identifiées chez *Mtb* fournit des indices sur ces enzymes potentiellement responsables du transfert[265]. En effet, leur alignement révèle que le domaine d'interaction avec les carbohydrates est le plus variable. De plus, si l'isoforme C présente une plus grande spécificité pour l'accepteur tréhalose, l'isoforme A préfère le TMM. L'expérience de docking et des tests *in-vitro* entre des motifs penta-arabinofuranoside et hexa-arabinofuranoside et Ag85A démontrent que cette dernière est capable de transférer des acides mycoliques sur les motifs terminaux du mAGP. Le tréhalose libéré est ensuite recyclé par le transporteur LpqY-SugABC et dirigé vers la synthèse nouvelle de TMM ou d'autres glycolipides[266].

I.B.2.c.iv Le dimycolyl diarabino glycerol (DMAG)

L'analyse systématique des glycolipides pariétaux de *Mma* a mis en évidence un glycolipide original composé de deux résidus d'arabinose susbtitués par des acides mycoliques de forme kéto C_{84} et C_{86}. Son analyse structurale par RMN, spectrométrie de masse et chromatographie en phase gazeuse a permis de le décrire comme un 5-*O*-mycolyl-β-D-Ara*f*-(1→2)-5-*O*-mycolyl-α-D-Ara*f*-(1→1)-Gro, désigné dimycolyl diarabinoglycérol (DMAG). Les similarités structurales avec les motifs terminaux du mAGP ainsi que sa susceptibilité à l'éthambutol et à l'isoniazide, tous deux connus pour inhiber la biosynthèse de l'arabinogalactane, suggèrent que le DMAG résulte du catabolisme du mAGP[267].

I.B.3 LA MYCOMEMBRANE ET SES GLYCOLIPIDES

A l'image des PIMS décrits précédemment, la mycomembrane contient deux autres grandes catégories de glycolipides, ceux dérivés du tréhalose et les glycopeptidolipides.

I.B.3.a LES GLYCOLIPIDES DERIVES DU TREHALOSE

Le tréhalose est un disaccharide non réducteur α-D-glucopyranosyl(1-1)α-D-glucopyranose décrit pour la première fois par Berthelot[268]. Chez les mycobactéries, trois voies biochimiques Ots-A/Ots-B, Tre-Z/Tre-Y et Tre-S permettent la biosynthèse du tréhalose[269]. Ce dernier n'est pas seulement une molécule de stockage d'énergie carbonée mais participe également à la protection

contre le stress osmotique. De plus, il est la source de plusieurs glycolipides en fonction de la structure fine des lipides et de leur position sur le tréhalose (Tableau 2).

Tableau 2 : Acyl tréhaloses des principales espèces. Seuls les TPPs natifs ont été caractérisés chez *Mabs*.

GLYCOLIPIDES	ACYLATIONS		ESPECES	REFERENCES
	positions	lipides		
DAT	2,3	acide gras linéaire en 2; acides gras mono ou polyméthyl branchés jusqu'à C_{27} en 3,6,2',6'	*Mtb, Mfor*	
TATs	2,3,4 ou 2,3,6		*Mtb, Mper, Mche, Mfor*	
tétraATs	2,3,6,2'		*Mtb*	
PATs	2,3,6,2',6'			
SL-I	2,3,6,6'	acide gras linéaire en 2; phtiocéraniques en 3,6,6'		
TMM	6	acides mycoliques	Toutes	
TDM	6,6'			
TPP A	2,3,4,6,2',3',4',6'	acide gras linéaire en 2; acides phléiques jusqu'à C_{40}	*Msmeg, Mabs, Mche, Mphlei, Mav*	
TPP B	2,3			
TPP C	2,4,6,2',3',4',6'			
TPP D	2,3,4,2',3',4',6'			

I.B.3.a.i Les acyl-tréhaloses

Position et structure des lipides

Des acides gras saturés à chaîne linéaire ainsi que des acides gras ramifiés polyméthylés peuvent substituer plusieurs positions du tréhalose. Dès 1988, Daffé décrit chez *Mtb* un penta-acyltréhaloses dont le tréhalose est substitué par des acides gras polyméthylbranchés de type mycolipenate ou mycolipanolate[270]. Au début des années 1990, Lemassu et collaborateurs purifient, toujours chez *Mtb,* le composé SL-IV qui n'est pas retenu sur colonne échangeuse d'ions contrairement au sulfolipide SL-1 contenant un groupement sulfate. Elle démontre par RMN et MS qu'il s'agit d'un di-*O*-acyl-tréhalose (DAT) substitué en position 2 et 3[271]. L'année suivante, Besra et collaborateurs identifient deux familles par CCM en phase inverse[272]: le DAT-1 est composé d'un tréhalose substitué en position 2 par un acide gras linéaire (C_{16}, C_{18} ou C_{19}), et en position 3 par un acide 2,4 diméthyl docosanoïque ou mycosanoïque, ainsi que les variantes à 25 et 26 carbones ; le DAT-2 diffère par la présence d'un acide 3-hydroxy 2,4,6-triméthyl tétracosanoïque, également nommé mycolipanolique, à la place de l'acide mycosanoïque sur la position 3 du tréhalose.

Chez *Mfor*, trois acyltréhaloses avec une polarité croissante ont été purifiés. Le premier est un 2,3,6 tri-*O*-acyltréhalose (TAT) ; le second est le 2,3,4 tri-*O*-acyltréhalose (TAT) légèrement plus polaire du fait de la présence de la fonction alcool primaire libre en position 6 ; le dernier est le 2,3 di-*O*-acyltréhalose (DAT) encore plus polaire car il présente un substitut acyl en moins. L'analyse RMN des spectres COSY autorise l'emplacement des substitutions sur la même unité glucosidique. Les substituants lipidiques sont des chaines saturées C_{16} ou C_{18} ainsi que des acides gras 2-méthyl α,β insaturés. Le 2,3,4 TAT est retrouvé dans plusieurs souches du complexe *Mfor* dont *Mche* et *Mycobacterium peregrinum*[273] : leur contenu en lipides étant essentiellement représenté par les acides gras $C_{16:0}$, $C_{18:0}$ ou $C_{18:1}$ et 2-Me $C_{18:1}$ ou 2-Me $C_{18:2}$. Les 2,3 DAT ont ensuite été séparés en plusieurs formes qui diffèrent par leurs proportions de ces mêmes lipides sur le tréhalose[274],[275].

Très récemment, une étude exhaustive du contenu en acyl-tréhaloses d'une souche de *Mtb* cultivé sous forme de biofilm, révèle la richesse des positions et des lipides identifiables. En effet, elle montre la présence d'un nouveau tétra-*O*-acyl tréhalose (tétraAT) en plus des DATs, TATs et PATs également détectés. Elle propose via la fragmentation observée en LC-MS que les positions 2' et 6' de la 2ème unité glucosidique soient estérifiés dans le TétraAT et le PAT. Enfin, un nouvel acide gras contenant un groupe cyclopropyl est présent dans les TATs, TetraATs, PATs en plus des acides gras lignocérique en C_{24}, mycosanoïque en C_{27}, mycolipénique, mycolipodiénoïque, ou mycolipanolique[276].

Biosynthèse et transport du di-*O*-acyltréhalose (DAT)

L' enzyme PapA3 transfert les lipides sur les positions 2 et 3 du tréhalose[277]. Les DATs obtenus sont transportés vers la surface de la mycobactérie par le transporteur MmpL10, pour ensuite être converti en PATs par le produit du gène *Chp2* codant une seconde acyltransférase[278]. Les autres acétyltransférases ainsi que leurs spécificités pour les chaines linéaires ou branchées n'ont pas pu être évaluées davantage.

Rôles biologiques

D'après Saavedra et collaborateurs, les 2,3 DATs et les TATs limitent la prolifération des lymphocytes T chez la souris, suggérant un rôle dans leur hypersensibilité dans la tuberculose chronique[279]. En comparant le mutant *Δmsl3* incapable de produire les mycolipénates et mycosanoates à la souche *Mtb* H37Rv de type sauvage, il a été démontré que les DATs et les PATs influencent les propriétés d'hydrophobicité de l'enveloppe ainsi que l'adhésion aux cellules phagocytaires et non phagocytaires[280]. Enfin, le sulfolipide SL-1 est un acyl-tréhalose sulfaté en position 2 spécifique à

Mtb[281]. Récemment, il a été observé qu'il active les neurones nociceptifs de l'homme et de la souris, et provoque la toux du cobaye, contrairement au tréhalose seul, au TMM ou au TDM[282].

I.B.3.a.ii Les mycolyl-tréhaloses

Tréhalose dimycolate (TDM) et monomycolate (TMM) identifiés selon les espèces

Le TDM fut décrit la même année que *Mabs*[283]. Nommé également « cord factor », il est impliqué dans l'agrégation des bacilles sous forme de cordes serpentiformes qui échappent à la phagocytose du système immunitaire, induisant ainsi la virulence[284]. L'estérification des acides mycoliques en position 6 du glucose a été déduite des hydrolyses ménagées du produit purifié[285] puis l'analyse RMN du composé natif chez *Msmeg* a confirmé la position de l'acide mycolique sur le groupement glucosyl[286]. Alors que le TDM porte deux acides mycoliques chacun en position 6, le TMM n'en possède qu'un seul. C'est pourquoi les variations structurales observées de ces deux molécules résident essentiellement dans celles des acides mycoliques précédemment décrits.

En particulier, l'analyse systématique par MALDI-MS du contenu en TDM et en acides mycoliques sous forme d'esters méthyliques (MAMEs) issus de multiples souches de *Mtb*, *Mbovis*, *Mav, Mkan*, *Mycobacterium phlei*, *Mycobacterium flavescens* et *Msmeg* permet de distinguer nettement les espèces[287]. Ainsi, les auteurs prouvent que *Mtb*, *Mbovis* et *Mkan* synthétisent exclusivement des TDM incluant les acides mycoliques α (C_{76} à C_{84}), méthoxy et keto (C_{85} à C_{89}) excepté la souche *Mbovis* Conaught dépourvue de la forme méthoxy. En parallèle, Les TDM de *Mav*, *Mycobacterium phlei* et *Mycobacterium flavescens* ne contiennent pas non plus d'acides mycoliques méthoxy mais des formes α, keto et wax ester de 78 à 87 carbones. De la même manière, l'étude du contenu en acides mycolique du TMM aboutit à des compositions identiques selon les espèces, renforçant son lien métabolique avec le TDM[288]. Dans le cas de *Mabs*, la composition en acides mycoliques ne met en évidence que les formes α C_{79} et α' C_{64} [177,289] laissant supposer qu'elles pourraient se retrouver dans le TDM mais ce dernier n'a jamais été caractérisé pas plus que le TMM.

Biosynthèse et transport du TMM : découverte du glucose monomycolate (GMM)

Le TMM est généralement présent en plus faible abondance car il s'agit d'un intermédiaire de biosynthèse du TDM qui est exporté hors du cystosol par MmpL3 [157]. Ce TMM peut être estérifié par l'antigène Ag85 (i) sur un autre TMM pour produire le TDM final ou (ii) sur la chaîne arabinane du complexe mAGP pour former la mycomembrane[290]. L'obtention du TMM à partir du tréhalose et d'acide mycolique libre ou activé est probablement due au même antigène 85 dont les trois isoformes ont des spécificités particulières pour leur substrat[265].

En présence d'une concentration élevée de glucose de l'ordre de 5%, l'espèce *Mav* produit un nouveau glycolipide plus apolaire que le TDM et le TMM et dont l'ion majoritaire m/z 1452 est compatible avec un hexose estérifié par un acide mycolique C_{85} de forme wax-ester. La détermination de sa composition en monosaccharides révèle la présence de glucose, et les tests in-vitro montre que l'Ag85A de *Mav* assure la médiation d'un échange de TDM vers ce GMM, s'expliquant par une compétition avec le TMM[291]. Pour conclure, ce GMM est détectable en conditions physiologiques mais dans des proportions largement inférieures au TMM et TDM.

Rôles biologiques du TDM

Tout d'abord, le TDM est un composé ubiquitaire des mycobactéries indispensable à la formation de la mycomembrane. De plus, le TDM semble posséder un rôles dans l'immunité innée, comme le suggère la diminution de la production de cytokines IL-1β, TNFα, IL-6 et MIPα en présence de bacilles délipidés[292]. De plus, la survie de bacilles délipidés dans les macrophages murins type C57BL/6 est amoindrie mais se restore avec l'apport de TDM. D'autre part, la délipidation montre que des bacilles dépourvus de TDM sont rapidement transportés dans le lysosome[293]. Parallèlement, des billes recouvertes de TDM ou de l'extrait apolaire montrent un défaut de localisation dans le lysosome par rapport à des billes recouvertes de BSA ou de tréhalose dibéhénate. Le TDM module donc la réponse inflammatoire et empêche que le bacille soit dégradé par le lysosome.

Les travaux d'Ishikawa et collaborateurs suggèrent quant à eux que la sécrétion du TNFα par les macrophages est induite par la reconnaissance du TDM par le récepteur C-Mincle[294]. *In vivo*, le TDM induit bien la formation de granulomes dans les poumons de souris mais pas chez les mutantes n'exprimant pas le récepteur C-Mincle. Ce récepteur est donc un ligand essentiel du TDM pour l'activation des macrophages et la formation du granulôme. Par ailleurs, il a été démontré que la longueur du lipide associé au TDM influence l'affinité[295].

Un modèle biophysique de membrane enrichie en TDM démontre un regain d'intégrité et de fluidité après un cycle de déhydratation/réhydratation comparativement une membrane constituée uniquement de phospholipides[79]. Au moins 25% de TDM en mélange avec les phospholipides est nécessaire pour résister à la dessiccation. Cette caractéristique représente un avantage stratégique pour se maintenir et persister en cas de stress hydrique.

Le rôle du TDM dans la formation des cordes a été l'objet de nombreuses études. Parmi celles-ci, Behling démontre que des billes de polystyrène recouvertes de TDM s'agrègent davantage que des billes nues[284]. En comparant deux souches de *Mtb* virulente et avirulente, Hunter observe que la

première forme une masse plus dense, irrégulière et “multilobée”, qu’elle recouvre l’interface eau-air d’une large pélicule rigide et qu’elle s’arrange sous forme de cordes serpentiformes après coloration de Kinyoun. De plus, la souche virulente est détectable par des anticorps anti-TDM en IHC et sa fraction éther de pétrole contient le TDM, contrairement à la souche avirulente[296]. Enfin, Les travaux plus récents de LLorens montrent que dans le cas de *Mabs*, la contribution du TDM à la formation des cordes pourrait s’avérer moins importante que celle des tréhaloses polyphléates[176].

I.B.3.a.iii Les tréhaloses polyphléates (TPPs)

L'acide phléique est un acide gras polyinsaturé à longue chaîne de C_{30} à C_{40} qui a été décrit au début des années 1970[297]. Dans un travail récent, Burbaud et collaborateurs identifient un locus de biosynthèse largement distribué parmi les espèces de mycobactéries dont les gènes présentent des similarités avec ceux du locus de biosynthèse des PATs ou du SL-I chez *Mtb*. La délétion des gènes MmpL10 et pE entraine l’accumulation d’un 2,3 di-*O*-acyl-tréhalose (TPP B) estérifié par un groupe acyle C_{14} à C_{19} et par un groupe phléyle $C_{36:5}$ à $C_{40:6}$. Par contre chez les souches sauvages de *Msmeg*, *Mabs*, *Mche* et *Mav*, la forme polyphléyle (TPP A) reste exposé en surface[298]. Dans son travail sur l’implication des TPPs dans l’agrégation *Mabs*, Llorens et collaborateurs confirment l'existence de deux nouveaux intermédiaires TPP C et TPP D (Tableau 2), chacun d'entre eux étant dépourvu d'un ou deux acides phléiques aux positions 3, 4 et/ou 6[176]. L'enzyme responsable du tranfert acides phléiques, l’acétyl-transférase pE a été récemment identifiée[299].

I.B.3.a.iv Les lipooligosaccharides (LOSs)

Structures parmi 3 espèces de MNTs et *Mycobacterium tuberculosis* (*Mtb*)

Cette famille de dérivés du tréhalose est présente chez certaines MNTs mais n’a jamais été décrite chez *Mabs* (Tableau 3). En particulier, sept LOSs I-VII pourvus d’un noyau tétrasaccharidique commun β-D-Glc*p*(1-4)β-D-Glc*p*(1-4)α-D-Glc*p*(1-1)α-D-Glc*p* ont été identifiés chez *Mkan.* Ils sont pour certains substitués par du 3-*O*-méthyl rhamnose, du fucose ou du xylose[300,301]. Trois LOSs A, B1 et B2 ont également été identifiés chez *Msmeg*[302,303]. Ils présentent un noyau identique β-D-Glc*p*(1-4)β-D-Glc*p*(1-6)α-D-Glc*p*(1-1)α-D-Glc*p* ainsi que des résidus pyruvyl reliant les positions 4 et 6 d’un ou deux Glc*p* terminaux. La recherche de LOSs amènera à les identifier aussi chez *Mmar* puisque Burguière isole en 2005 4 d’entre eux intégrant le même noyau tétrasaccharidique[217]. Parmi eux, le LOS-I est un pentasaccharide contenant un résidu terminal de 3-*O*-Me-α-L-Rha*p*. Les 3 autres LOSs dont le LOS-II, possèdent un α-D-xylopyranose supplémentaire en position 4, en plus d’autres résidus non identifiés[304]. Il s’agira finalement d’un résidu 3,6-didéoxy,4-C-(D,altro1,3,4,5-tetrahydroxyhexyl)-α-D-

xylo-hexopyranose portant le nom de caryophyllose qui se retrouve unique dans le LOS-II, et qui se répète dans le LOS-III[305]. Enfin, le LOS-IV arbore un N-acyl-4-amino-4,6-didéoxy-α-D-Gal*p* terminal qui se retrouvait également dans l'un des LOSs identifiés par Daffé chez *Mtb*[306]. En 1991, il caractérise 2 LOSs présents chez *Mtb* dont le plus simple est constitué d'un octasaccharide, alors que le second porte un N-acyl-4-amino-4,6-didéoxy-Gal*p* supplémentaire[307]. Cette différence de structure s'accompagne d'une différence de réactivité antigénique.

Tableau 3 : LOSs des principales espèces. En gras, la structure du noyau commun est construite autour du résidu de tréhalose. En italique, l'ajout ou la modification des résidus dans les nouvelles formes de LOSs ou les différentes souches montre leur diversité.

ESPECES	GLYCOLIPIDES	NOYAU OLIGOSACCHARIDIQUE COMMUN *Motif ajouté à l'extrémité non-réductrice*	REFERENCES
Mkan	LOS-I	**3-*O*- Me-α-L-Rhap(1→3)β-D-Glcp(1→3)β-D-Glcp(1→4)α-D-Glcp(1↔1)α-D-Glcp**	
	LOS-II-III	*6-D-Xylp(1→4)* $_n$	
	LOS-IV-VII	*4,6-didéoxy-2-O-Me-3-C-Me-4-(2'-methoxypropionamido)-α-D-Glcp(1→3)α-L-Fucp(1→4)*	
Msmeg	LOS-B1	**4,6-(1-carboxyethylidene)-β-D-Glcp(1→4)β-D-Glcp(1→6)α-D-Glcp(1↔1)α-D-Glcp**	
	LOS-B2	*3-O-Me-6-D-Glcp(1→3)*	
	LOS-A	*4,6-(1-carboxyethylidene)-3-O-Me-6-D-Glcp*	
	mc² 22	**β-D-Glcp(1→3)4,6-(1-méthoxycarbonylethylidene)-β-D-Glcp(1→4)β-D-Glcp(1→6)α-D-Glcp(1↔1)α-D-Glcp**	
	mc² 11	*4,6-(1-méthoxycarbonylethylidene)-3-O-Me-6-D-Glcp*	
Mmar	LOS-I	**3-*O*- Me-α-L-Rhap(1→3)β-D-Glcp(1→3)β-D-Glcp(1→4)α-D-Glcp(1↔1)α-D-Glcp**	
	LOS-II*	*6-D-Xylp(1→4)*	
	LOS-II	*(3-OH)-α-D-Carp(1→4)*	
	LOS-IV	*(3-OH)-α-D-Carp(1→C)*	
	LOS-V	*Z-4N-4,6-didéoxy-α-D-Galp(1→C)*	
Mtb	LOS-I	**2-*O*- Me-α-L-Fucp(1→3)β-D-Glcp(1→3)2-*O*- Me-α-L-Rhap(1→3)2-*O*- Me-α-L-Rhap(1→3)β-D-Glcp(1→3)4-*O*- Me-α-L-Rhap(1→3)6-*O*- Me-α-D-Glcp(1↔1)α-D-Glcp**	
	LOS-II	*N-acyl-4-amino-4,6-didéoxy-Galp(1→?)*	

Tous les LOSs sont tri-acylés mais la nature et la position des lipides sur le résidu de tréhalose varie selon les espèces. Substitué en positions 3, 4 et 6 les LOSs de *Mkan* et *Mtb* contiennent respectivement des acides 2,4-diméthyltétracosanoïque ou 2,4-diméthylhexadecanoïque, 2,4,6,8-tétraméthyloctadecanoïque ainsi que 2-méthyl-3-hydroxyeicosanoïque. Inversement les LOSs de *Msmeg* et *Mmar* sont respectivement estérifiés sur les positions 3,4,2' ou 2,3,2'.

Rôles biologiques des LOSs de *Mycobacterium marinum (Mmar)*

Le rôle des LOSs dans l'immunomodulation a surtout été décrit chez *Mma*. Un défaut dans leur biosynthèse limite la capacité à infecter les macrophages. Le LOS-IV inhibe la production de TNFα par les macrophages THP-1[305] et active les macrophages en favorisant l'expression des antigènes de surface ICAM-1 et CD-40[306]. D'ailleurs, un mutant qui ne produit pas de LOS-IV est hypervirulent chez l'embryon de poisson zèbre[308]. De plus, il a été démontré que les souches déficientes dans la

production de LOS étaient phagocytées plus efficacement que celles n'accumulant que les intermédiaires précoces, alors que les souches sauvages et mutantes dépourvus de LOS-IV étaient les moins phagocytées. Ce dernier LOS possède donc un rôle central dans la pathogénicité de *Mma*[309]. Dans le cas de *Mtb*, les LOSs furent évalués pour le sérodiagnostic[310] ou la détermination de la morphologie[311] avec moins de succès que les glycopeptidolipides.

I.B.3.b LES GLYCOPEPTIDOLIPIDES (GPLs) PORTEURS DE DEOXYHEXOSES

Au sein de l'enveloppe mycobactérienne, les GPLs représentent une véritable frontière moléculaire et taxonomique distinguant les MNTs de *Mtb* qui en est totalement dépourvue[4]. En 1968, Lanéelle et Asselineau confirment l'essentiel de la composition et de la structure du mycoside C chez *Mycobacterium* 1217 qui est une espèce non photochromogène apparenté aux mycobactéries atypiques[312]. Ce composé possède un acide tétracosanoïque β-hydroxylé de la forme D qui est estérifié sur la fonction amine d'un tétrapeptide phénylalanine-*allo*-thréonine-alanine-alaninol, lui même substitué par un résidu de 2,3,4 tri-*O*-méthyl ou 3,4 di-*O*-méthyl rhamnose sur l'alaninol, et d'un 6-déoxy talose sur l'*allo*-thréonine.

I.B.3.b.i *Mycobacterium xenopi* (*Mxen*) : nouveau noyau peptidique

Au début des années 1990, l'espèce *Mxen* fait l'objet d'une attention particulière puisqu'elle produit un GPL dont la structure originale révèle néanmoins l'association similaire de trois familles de composés. Le lipide est un acide laurique relié par une liaison amide à un noyau peptidique. La séquence de ce tétrapeptide carboxyméthylé du côté C terminal est Ser-Ser-Phe-*allo*-Thr [313]. Alors que les sérines sont de la série L, la phénylalanine et l'*allo*-thréonine sont de la série D [314]. La première sérine est glycosylée par un résidu de 3-*O*-Me-α-L-6dTal*p* mais pas la seconde qui a la possibilité d'être O-méthylée[315]. Le résidu d'*allo*-thréonine porte le motif tétrasaccharidique 2,3,4-*O*-Me_3-α-L-Rha*p*(1-3)2-*O*-lauryl-α-L-Rha*p*(1-3)-α-L-Rha*p*(1-3)2-*O*-lauryl-4-*O*-acetyl-α-L-Fuc*p*. D'autres souches produisent des GPL de structure plus simple telle que le GPL *Mxen* ATCC19250 substitué par le disaccharide α-L-Rha*p*(1-3)2-*O*-lauryl-α-L-Rha*p* dont le rhamnose terminal est possiblement acylé par un acide gras à 8 ou 10 carbones. De plus, la présence d'un groupement 2-*O*-acétyl sur le 3-*O*-méthyl-α-L-6dTal*p* constitue une seconde variation structurale.

I.B.3.b.ii *Mycobacterium avium* (*Mav*) : une grande diversité de structures

C'est en 1974 que Draper décrit un objet fibrillaire, observable au microscope électronique, dont il extrait des 6-déoxy hexoses non caractérisés, des quantités équimolaires de thréonine, d'alanine et d'alaninol ainsi que des traces de phénylalanine, de valine et de leucine[316]. La présence de

l'un de ces 2 derniers à la place de la phénylalanine dans le noyau tripeptidique est confirmée chez la souche *Mycobacterium avium intracellulare* 700737 grâce à la détermination de la composition en acides aminés par LC-MS[317].

Tableau 4 : GPLs de *Mycobacterium avium (Mav)*. Liste des structures identifiées parmi les sérovars de *Mav* à partir du nsGPL d'après Chatterjee[201].

Serovar	Haptenic oligosaccharide extending from α-L-Rha*p*-(1→2)-L-dTal[1]
1	Non-extended core; α-L-Rha*p*-(1→2)-L-dTal
	Group 1 : R → α-L-Fuc*p*-(1 → 3)-core
2	4-*O*-Ac-2,3-di-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
4	4-*O*-Me-α-L-Rha*p*-(1 → 4)-2-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
14	4-formamido-4,6-dideoxy-2-*O*-Me-3-*C*-Me-α-L-Man*p*-(1 → 3)-2-*O*-Me-α-D-Rha*p*-(1 → 3)-core 2-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
20	2-*O*-Me-α-D-Rha*p*-(1 → 3)-2-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
3	2,3-di-*O*-Me-α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*A-(1 → 4)-2,3-di-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
9	4-*O*-Ac-2,3-di-*O*-Me-α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*A-(1 → 4)-2,3-di-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
25	2-*O*-Me-α-D-Fuc*p*NAc-(1 → 4)-β-D-Glc*p*A-(1 → 4)-2-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
26	2,4-di-*O*-Me-α-L-Fuc*p*-(1 → 4)-β-D-Glc*p*A-(1 → 4)-2-*O*-Me-α-L-Fuc*p*-(1 → 3)-core
	Group 2 : R → -α-L-Rha*p*-(1 → 3)-core
12	4-(2-OH)propanamido-4,6-dideoxy-3-*O*-Me-β-D-Glc*p*-(1 → 3)-4-*O*-Me-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-core
17	3-(3-OH-2-*O*-Me)butanamido-3,6-dideoxy-β-D-Glc*p*-(1 → 3)-4-*O*-Me-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 3)-core
19	3,4-di-*O*-Me-β-D-Glc*p*A-(1→3)-3-*C*-Me-2,4-di-*O*-Me-α-L-Rha*p*-(1 → 3)-α-LL-Rha*p*-(1 → 3)-core
	Group 3 : R → -α-D-Glc*p*-(1 → 3)-core
8	4,6-*O*-(1-carboxyethylidene)-3-*O*-Me-β-D-Glc*p*-(1 → 3)-core
21	4,6-*O*-(1-carboxyethylidene)-β-D-Glc*p*-(1 → 3)-core

Finalement, la diversité des structures identifiées de GPLs est la plus vaste chez *Mav* qui compte jusqu'à 26 sérovars[201]. Dans sa forme la plus simple, appelée GPL apolaire ou non spécifique (nsGPL), le noyau tripeptidique Phe-*allo*-Thr-Ala et le lipide sont conservés de même que le 6dTal relié à l'*allo*-thréonine et le 3,4-*O*-Me_2-α-L-Rha*p* relié à l'alaninol. Sur la position 2 du 6dTal, le produit du gène *rtfA* ajoute un résidu α-L-Rha*p* formant ainsi le GPL du sérovar 1[318]. La répartition des sérovars dans les groupes 1, 2 ou 3 se base sur l'ajout respectif d'un résidu α-L-Fuc*p*, α-L-Rha*p* ou α-D-Glc*p* en position 3 [201] (Tableau 4).

Groupe 1 : type fucose

Le produit du gène *gtfD* transfère α-L-Fuc*p* sur la position 3 du rhamnose donnant naissance au GPL du sérovar 2[319]. Ce résidu de fucose est O-méthylé en position 2 et 3, et *O*-acylé en position 4[201]. D'autre part, la présence d'un unique groupement acétyl en position 4 du 6dTal potentialise l'interaction du GPL du sérovar 2 avec le récepteur TLR-2. Cette interaction induit la production de TNFα alors que la forme complétement déacétylé n'est pas active[320].

Le gène *hlpA* code une 1-4 rhamnosyltransférase permettant la synthèse du GPL du sérovar 4 et potentiellement du sérovar 20 à partir du sérovar 2 [321]. Pour le premier, le rhamnose porte alors un

groupement O-méthyl en position 4 alors qu'il est en position 2 dans le cas du GPL du sérovar 20. Les deux isomères de position induisent des réponses pro-inflammatoires plus forte pour le GPL du sérovar 4 que du sérovar 20[322].

Au sein des derniers sérovars 3, 9, 25 et 26 du groupe 1, ce rhamnose est remplacé par le disaccharide α-L-Fuc*p*1(1-4)β-D-Glc*p* sur la position 4 du fucose. C'est le nombre et la position des groupements *O*-méthyl et/ou *O*-acyl qui différencient alors chacun des 4 sérovars[201].

Groupe 2 : type rhamnose

Les sérovars 7, 12, 17 et 19 possèdent tous le motif disaccharidique α-L-Rha*p*(1-3)-α-L-Rha*p* relié au premier rhamnose en position 3. Pour les sérovars 12 et 17, le dernier rhamnose porte une fonction O-méthyl en position 4 alors que pour le sérovar 19, 2 groupements supplémentaires O-méthyl et C-méthyl sont respectivement ajoutés aux positions 2 et 3. A l'extrémité non réductrice, les sérovars 12 et 17 se terminent par un β-D-didéoxyGlc*p* relié à la position 3 du rhamnose terminal. Une modification de type amido sur la position 4 ou 3 distingue les sérovars 12 et 17.

Dans le sérovar 7 décrit plus tard par Fujiwara, la condensation d'un acide 2-hydroxy propanoïque sur la fonction 4 amino du β-D-didéoxyGlc*p* aboutit également à la présence de ce groupement amido [323].

Groupe 3 : type glucose

Les GPLs des sérovars 8 et 21 sont issus de l'action d'une β(1-3)glucosyltransférase codée par le gène *gtfTB* sur le rhamnose du GPL du sérovar 1. Ensuite, le résidu de glucose est modifié par une pyruvyltransférase reliant les positions 4 et 6 et une méthyltransférase ajoutant le groupement O-méthyl en position 3. L'absence de cette dernière modification caractérise le GPL du sérovar 21[324]. D'après Barrow, une fraction contenant les GPLs du sérovar 8 induit une sécrétion de PGE2 et de TNFα comparable au GPL du sérovar 4[322].

Cette pléiade de structures de GPLs détermine ainsi la classification des sérovars de *Mav*. La sulfatation de la position 4 du 6dTal dans le GPL du sérovar 1 fut décrite par Khoo en 1999[325], ce type de modification étant également retrouvée chez l'espèce *Mfor*.

I.B.3.b.iii *Mycobacterium peregrinum* (*Mper*) : absence du 6dTal et présence de sulfate

Parmi les GPL-I à GPL-IV isolés de *Mfor biovar peregrinum*, l'*allo*-thréonine n'est pas glycosylée par le 6dTal mais par un 3-*O*-Me-α-L-Rhap. La présence du disaccharide 3,4-*O*-Me_2-α-L-Rha*p*(1-2)3,4-*O*-Me_2- α-L-Rha*p* sur le résidu d'alaninol complète le motif glycannique du GPL-I (Tableau 5). Dans le

cas du GPL-II, le disaccharide est le 3-*O*-Me-α-L-Rha*p*(1-2)3,4-*O*-Me_2-α-L-Rha*p*[326]. Le GPL-III quant à lui ne possède plus de groupement méthyl sur le rhamnose terminal qui est remplacé à la position 2 par un groupement sulfate dans le GPL-IV[327]. L'utilisation des GPLs identifiés sur des préparations mitochondriales a mis en évidence une déstabilisation marquée de la respiration et de la phosphorylation oxydative par les GPL-III et IV, plus polaires. En parallèle, la perméabilité membranaire passive est fortement augmentée par les GPLs les plus apolaires[328]. En outre, leur réactivité antigénique a été évalué sur plusieurs séra de lapin exposés à différentes espèces de mycobactéries, démontrant une séroréactivité élevée envers les mycobactéries productrices de GPLs et un signal moindre voire nul pour des espèces dépourvues de GPLs[329].

I.B.3.b.iv *Mycobacterium habana* (*Mhab*) : présence d'acides uroniques

Chez cette espèce, un mélange de GPLs apolaires (aGPLs) et très polaires (pGPLs) caractérisés par la présence d'acide glucuronique (GlA) a été observé[330] (Tableau 5). Le lipopeptide est identique à celui retrouvé dans les autres espèces. Chez les aGPLs, l'α-L-6dTal*p* relié à l'*allo*-thréonine porte un groupement *O*-méthyl en position 3 alors que l'α-L-Rha*p* lié à l'alaninol en possède un ou deux en position 3 et 4. Chez les pGPLs, c'est la nature de l'oligosaccharide ancré sur la position 2 du 6dTal qui engendre les formes I, II et III. Dans le cas de la forme I, il s'agit du α-L-Fuc*p*(1-3)6-*O*-Me-β-D-Glc*p*(1-3)4-*O*-Me-α-L-Rha*p*(1-3)-α-L-Rha*p*. Du côté non réducteur, le pGPL-II possède un 2,4-*O*-Me_2-GlA*p* et le pGPL-III un 4-*O*-Me-GlA*p* qui est relié en β(1-4) au α-L-Fuc*p* lui-même O-méthylé en position 3. Compte tenu d'une composition similaire en acides gras et en acides mycoliques, la caractérisation de ces pGPLs parmi les souches de *Mhab* a abouti à sa distinction de l'espèce *Mycobacterium simiae* qui présente un 3,4-*O*-Me_2-GlA*p* comme épitope terminal[331]. Néanmoins, une micro-hétérogénéité existe dans le nombre et la position des groupements méthyls chez ce dernier puisque l'acide glucuronique peut être 2,3,4 tri-*O*-méthylé et le fucose 2,3 di-*O*-méthylé[332].

Tableau 5 : GPLs d'autres espèces mycobactériennes. Liste des structures identifiées parmi 3 espèces *Mhab*, *Mper* et *Mxen* d'après Chatterjee[201]. Seule cette dernière espèce ne présente pas le tripeptide Phe-Ala-Thr.

Species/Strain	GPL	Oligosaccharides/Monosaccharide	
		1. C-mycoside GPLs: Fattyacyl-NH-D-Phe-D-(R_1-*O*-) *allo*Thr-D-Ala-L-alaninol-*O*-R_2	
		R_1	R_2
M. habana TMC5135	pGPL-I	α-L-Fuc*p*-(1 → 3)-6-*O*-Me-α-D-Glc*p*-(1 → 3)-4-*O*-Me-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 2)-3-*O*-Me-α-L-dTal*p*	3,4-di-*O*-Me-α-L-Rha*p*
	pGPL-II[a]	2,4-di-*O*-Me-β-D-Glc*p*A-(1 → 4)-3-*O*-Me-α-L-Fuc*p*-(1 → 3)-6-*O*-Me-β-D-Glc*p*-(1 → 3)-4-*O*-Me-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 2)-3-*O*-Me-α-L-dTal*p*	3,4-di-*O*-Me-α-L-Rha*p*
	pGPL-III	4-*O*-Me-β-D-Glc*p*A-(1 → 4)-3-*O*-Me-α-L-Fuc*p*-(1 → 3)-6-*O*-Me-β-D-Glc*p*-(1 → 3)-4-*O*-Me-α-L-Rha*p*-(1 → 3)-α-L-Rha*p*-(1 → 2)-3-*O*-Me-α-L-dTal*p*	3,4-di-*O*-Me-α-L-Rha*p*
		2. C-mycoside-like GPLs: Fattyacyl-NH-D-Phe-D-(R_1-*O*-) *allo*Thr-D-Ala-L-alaninol-*O*-R_2	
		R_1	R_2
M. fortuitum[b] biovar *peregrinum*	GPL I	3-*O*-Me-α-L-Rha*p*	3,4-di-*O*-Me-α-L-Rha*p*-(1 → 2)-3,4-di-*O*-Me-α-L-Rha*p*
	GPL II	3-*O*-Me-α-L-Rha*p*	3-*O*-Me-α-L-Rha*p*-(1 → 2)-3,4-di-*O*-Me-α-L-Rha*p*
	GPL III	3-*O*-Me-α-L-Rha*p*	α-L-Rha*p*-(1 → 2)-3,4-di-*O*-Me-α-L-Rha*p*
	GPL IV	3-*O*-Me-α-L-Rha*p*	2-*O*-sulfate-3,4-di-*O*-Me-α-L-Rha*p*
		3. Non C-mycoside GPLs: Fattyacyl-NH-(R_1-*O*-)-L-Ser-(±*O*-Me)L-Ser- L-Phe-(R_2-*O*-)-D-*allo*Thr-O-Me	
		R_1	R_2
M. xenopi ATCC 19250	GPL-I	3-*O*-Me-α-L-dTal*p*	α-L-Rha*p*-(1 → 3)-2-*O*-lauryl-α-L-Rha*p*
	GPL-II	2-*O*-Ac-3-*O*-Me-α-L-dTal*p*	α-L-Rha*p*-(1 → 3)-2-*O*-lauryl-α-L-Rha*p*
	GPL-III	3-*O*-Me-α-L-dTal*p*	4-*O*-octanoyl/decanoyl-α-L-Rha*p*-(1 → 3)-2-*O*-lauryl-α-L-Rha*p*
	GPL-IV	2-*O*-Ac-3-*O*-Me-α-L-dTal*p*	4-*O*-octanoyl/decanoyl-α-L-Rha*p*-(1 → 3)-2-*O*-lauryl-α-L-Rha*p*
NCTC10042[c]	GPL-X-IIb[d]	3-*O*-Me-α-L-dTal*p*	α-L-Rha*p*-(1 → 3)-2-*O*-lauryl-α-L-Rha*p*
CIPT 14035004	GPL X-I	3-*O*-Me-α-L-dTal*p*	2,3,4-tri-*O*-Me-α-L-dRha*p*-(1 → 3)-2-*O*-lauryl-α-L-dRha*p*-(1 → 3)-α-L-dRha*p*-(1 → 3)-2,4-di-*O*-acetyl/lauryl-6-deoxy-α-L-Glc*p*

I.B.3.b.v *Mycobacterium smegmatis* (*Msmeg*) : le GPL modèle

Cette espèce non pathogène, à croissance rapide, largement utilisée en laboratoire comme modèle ouvrira la voie à la compréhension de la biosynthèse des GPLs au travers de la caractérisation de ses mutants. En 1983, Daffé et collaborateurs utilisent la spectrométrie de masse pour détecter un ion à *m/z* 1283 correspondant à un adduit sodium et dont la fragmentation confirme la ressemblance avec le mycoside C d'autres espèces de mycobactéries précédemment analysé[333]. Le lipide hydroxylé ou méthoxylé en position 3 possède une double liaison. Il est estérifié côté N-terminal sur le tripeptide Phe-*allo*-Thr-Ala qui se termine par l'alaninol dans la souche mc^2155 [334]. C'est en comparant cette souche sauvage avec un mutant obtenu par transposition que Billman et collaborateurs identifient la région codant les peptides synthases impliquées dans la production du noyau peptidique du nsGPL[334]. L'analyse de la composition en monosaccharides indique une proportion 1/1/2 entre le 2,3,4-*O*-Me_3-Rha, le 3,4-*O*-Me_2-Rha et le 6dTal. L'analyse par spectrométrie de masse permet d'identifier deux formes de GPL diglycosylés dans lequel le 6dTal est porté par l'*allo*-thréonine et le Rha plus ou moins méthylé par l'alaninol[335].

Parmi les quatre glycosyltransférases potentiellement impliquées dans l'ajout des deux déoxyhexoses, le produit du gène *gtf1* transfère le 6dTal sur l'*allo*-thréonine alors que celui du gène

gtf2 catalyse la glycosylation du Rha sur l'alaninol [335]. Grâce à la complémentation, les orthologues respectifs de *gtf1* et *gtf2* ont été identifiés comme étant *GtfA* et *GtfB* chez Mav. Enfin, l'inactivation du gène *gtf4* n'a pas d'impact sur le profil des GPLs alors que celle du gène *gtf3* entraine la disparition de deux GPLs plus polaires et faiblement exprimés. La détermination de la composition en monosaccharides associé à l'analyse par spectrométrie de masse montre que ce GPL triglycosylé porte un deuxième rhamnose 3-*O*-méthylé. Ce composé ainsi que le 3,4-*O*-Me_2-Rha sont plus intensément exprimés lorsque *Msmeg* est cultivé dans un milieu dépourvu de glucose [336] et préférentiellement en phase exponentielle de croissance[337]. Récemment, la production de la protéine gtf3 recombinante a été réalisée chez *Escherichia coli* sans qu'une structure tridimensionnelle résolue ne soit proposée[338].

Plus récemment, l'enzyme qui transfère le groupement *O*-méthyl sur la position 3 du premier rhamnose a été identifiée comme le produit du gène *mtf1* [339]. Les produits de *rmt2* et *rmt4* ajoutent quant à eux les groupements *O*-méthyls en position 2 et 4. De plus, les auteurs indiquent que le gène *rmt4* possède 2 orthologues *MtfB* et *MtfC* chez *Mav* de même que *mtf1* (*rmt3*) est orthologue de *MtfD*[340]. Le fait qu'il n'y a pas d'équivalent de *rmt2* chez *Mav* justifie qu'aucune 2-*O*-méthylation n'ait été décrite sur le rhamnose dans cette espèce. La dernière méthyltransférase identifiée dans la région du génôme responsable de la biosynthèse des GPLs est codée par le gène *Mtf2* dont la mutation n'a pas d'effet sur le niveau de méthylation des déoxyhexoses[341]. Par contre, elle provoque l'accumulation de deux GPLs qui disparaissent au profit de deux autres plus apolaires lorsque le gène *Mtf2* est surexprimé. Les auteurs en déduisent que le produit du gène *Mtf2* modifie la fonction β-hydroxyle déjà décrite du lipide. En 2012, Vats et collaborateurs proposent de positionner la fonction hydroxyle en 5 et non en 3 [342].

Toujours chez la souche mc^2155, l'acétylation des GPLs est étudié chez un mutant d'insertion dans le gène *atf1* qui fut identifié par sa morphologie rugueuse et sa moindre capacité à former des biofilms. En effet, l'hydrolyse alcaline douce des GPLs fortement polaires produits par ce mutant ne modifie ni leur R_f en chromatographie couche mince ni le rapport m/z des ions détectés en MALDI-TOF. A l'inverse, les GPLs de la souche sauvage ont une polarité plus élevée après la saponification et le rapport m/z des ions détectés diminue globalement de 84 Da[343]. Le produit du gène *atf1* participe donc à l'ajout de deux groupements acétyls. Un autre gène *MSMEG1604* pourrait également être impliqué dans l'acétylation puisque son inactivation aboutit à l'accumulation d'un GPL non acétylé et triglycosylé portant le 3-*O*-Me-α-L-Rha*p*1(1-3)3,4-*O*-Me_2-α-L-Rha*p* sur l'alaninol[344].

En plus de l'acide acétique, l'acide succinique est le deuxième acide carboxylique court susceptible d'estérifier les positions hydroxyles libres des déoxyhexoses. Recherchant la capacité de plusieurs fractions lipidiques à inhiber la phagocytose de bacilles entiers par des macrophages, Villeneuve décrit que les GPLs sont ceux qui interfèrent le plus avec ce phénomène, en particulier les formes porteuses d'un groupement succinyl sur la position 2 du rhamnose terminal aussi bien sur le nsGPL que sur la forme triglycosylée[345]. En effet, le GPL succinylé interagirait davantage que le nsGPL avec les récepteurs CR3 et MR pour favoriser la phagocytose des bacilles de *Msmeg*[346]. Néanmoins, le gène codant l'enzyme responsable de cette modification n'a pas été identifié dans le locus de biosynthèse des GPLs.

I.B.3.b.vi *Mycobacterium abscessus* (*Mabs*) : specificités structurales

La première étude de la structure des GPLs de *Mabs* a été réalisée en 1994 parallèlement à celle de l'espèce *Mche* à laquelle elle est apparentée[347]. Dans ce travail, les auteurs montrent que les deux espèces ont jusqu'à cinq GPLs composés de D-phénylalanine, D-*allo*-thréonine, D-alanine, alaninol pour la partie peptidique, de 6 déoxy-talose (6dTal) et de rhamnose (Rha) pour la partie glycannique et un lipide de 26 ou 28 carbones. La présence de trois déoxyhexoses tous en configuration α ainsi que l'existence de trois groupements méthoxy et de deux fonctions acétyles a également été démontrée.

Les analyses démontrent également que *Mabs* (souche CIP104536) et *Mche* (souche CIP104535 produisent une quantité relative plus importante de GPL triglycosylé que *Msmeg* [348]. Dans le même travail, Ripoll et collaborateurs soulignent de nombreuses similitudes dans le locus de biosynthèse des GPLs entre les trois espèces (Figure 11). Cette observation soutient la similarité structurale du nsGPL entre les espèces étudiées mais n'explique pas la proportion différenciée entre GPLs diglycosylé et triglycosylé. Seul le gène *atf* unique chez *Msmeg* est dupliqué chez *Mabs* et *Mche* : le gène *atf1* est orthologue du gène *atf,* et le gène *atf2* s'intercale entre les gènes *rmlB* et *rmt2*. L'hypothèse que les produits des deux gènes aboutissent à deux enzymes ayant une spécificité d'action selon la position d'acétylation en 3 ou 4 du 6dTal n'a pas été démontrée par la caractérisation respective des 3 et 4 mono-*O*-acétyl 6-déoxy talose.

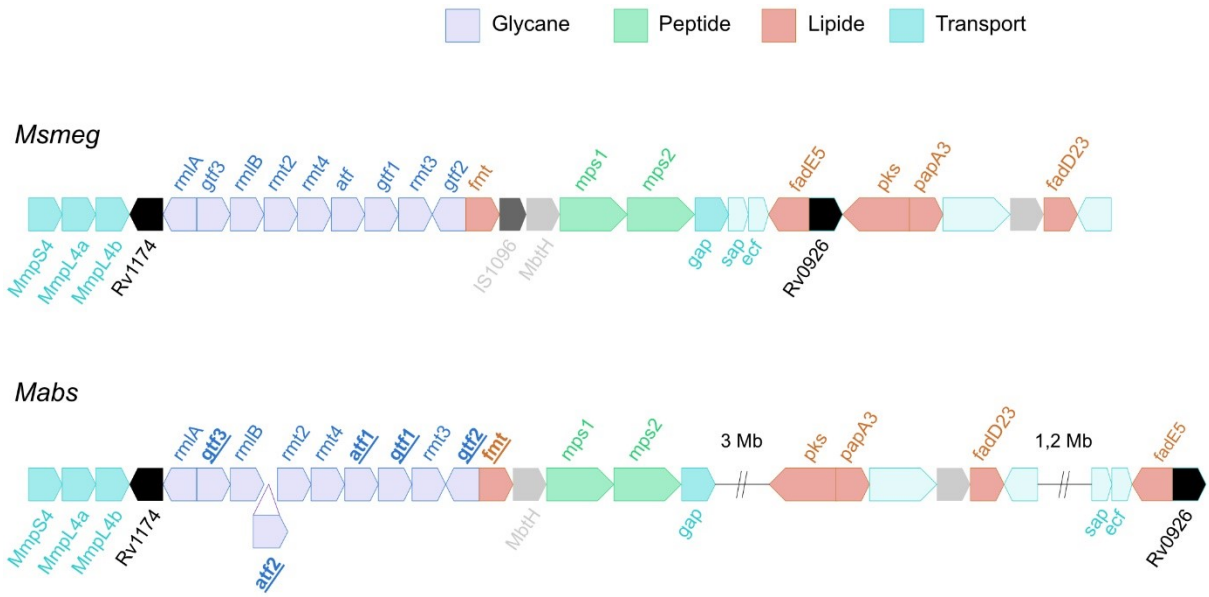


Figure 11 : Comparaison des gènes annotés au locus de biosynthèse des GPLs chez *Msmeg* et *Mabs* [348]. Les gènes *mps1* et *mps2* impliqués dans la synthèse du peptide et ceux impliqués dans le transport vers la surface de l'enveloppe sont identiques. Pour le glycane, le gène *atf* de *Msmeg* est dupliqué chez *Mabs*. Pour le lipide, seul le gène *fmt* demeure dans le locus de *Mabs*. Les rôles des gènes soulignés dans la biosynthèse et les phénotypes de *Mabs* sont présentés dans le chapitre VI.E

Récemment, le consensus établi autour du noyau peptidique a été remis en question puisque d'autres amino alcools branchés ont été mis en évidence. En effet, la culture de *Mabs* et *Mmas* prélevés de patients atteints de mucoviscidose ou sur un milieu synthétique qui reproduit ces prélèvements induit non seulement une proportion plus élevée de forme triglycosylée par rapport à la forme diglycosylée mais aussi la synthèse de GPL contenant un résidu de leucinol à la place de l'alaninol [349]. L'utilisation préférentielle par *Mabs* des acides aminés branchés se confirme puisque la croissance d'un mutant n'exprimant pas le gène *BCKADH*, codant la déhydrogenase des kéto acides branchés, diminue sur les milieux enrichis en valine, leucine ou isoleucine. Dans ce même travail, Wiersma et collaborateurs proposent que la fonction hydroxy ou methoxy du lipide *N*-acylé est placé en α du carbonyle.

Excepté sur ce dernier point, les observations convergent vers une structure tripartite des GPLs du complexe *Mabs* et de ces sous-espèces (Figure 12). Le noyau peptidique est constitué des trois acides aminés D-phénylalanine, D-*allo*-thréonine et D-alanine elle-même reliée à une molécule d'amino-alcool branché. Le second composant est un acide gras de 26 à 33 carbones possiblement insaturé et portant une fonction hydroxyle susceptible d'être méthylée. Enfin, le motif glycannique se décompose en 2 parties : un 6-déoxy-α-L-talopyranose relié à la fonction hydroxyle de la thréonine, et un α-L-rhamnopyranose greffé à l'amino-alcool branché via une liaison glycosidique. Le premier est acétylé en position 3 et 4 et le second méthylé en position 3 et 4 également. Le GPL diglycosylé (GPL-

2a) obtenu est dit apolaire ou non spécifique puisqu'il est commun aux espèces de mycobactéries précédemment citées notamment *Mav*, *Msmeg* et *Mabs*. La position 2 du rhamnose peut rester libre, être méthylé ou glycosylé par un second résidu α-L-rhamnopyranose pour former un GPL triglycosylé (GPL-3). Les GPL-2a et GPL-3 sont majoritairement présents dans l'enveloppe de la forme lisse de *Mabs* et facilement détectables dans l'extrait lipidique total par spectrométrie de masse [59].

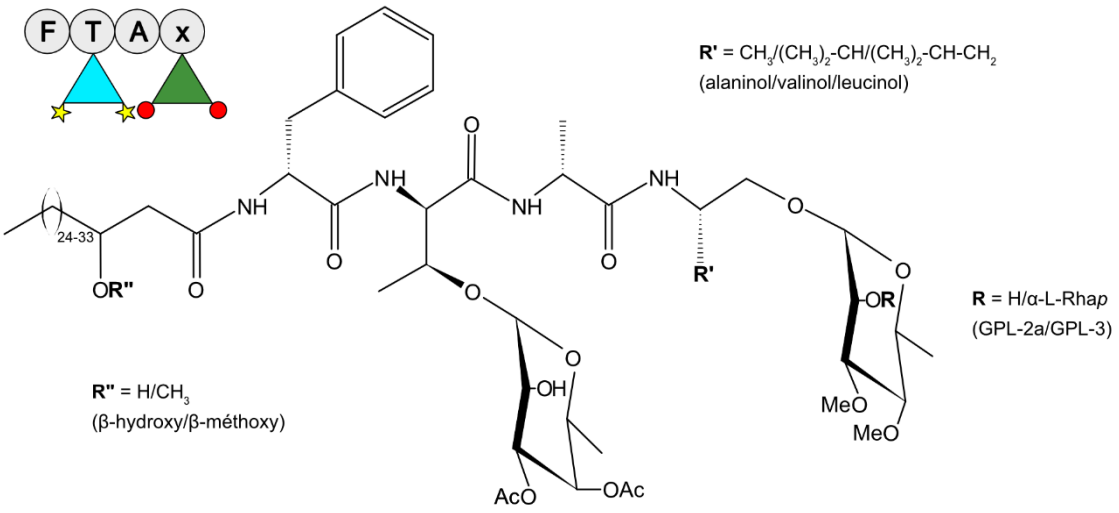


Figure 12 : Schéma général de la structure chimique du GPL de *Mabs*. Concernant la partie glycannique (R), un second α-L-Rha*p* peut être porté par la position 2 du premier donnant le GPL-3. Trois amino-alcools sont possibles du côté C-terminal (R'). Enfin, côté N-terminal, le lipide est hydroxylé ou méthoxylé (R'').

I.B.3.b.vii Transport des GPLs

Au sein des trois clusters de gènes qui participent à la biosynthèse des GPLs de *Mabs*, certains sont destinés au transport et à l'export vers la surface de l'enveloppe où ils se retrouvent massivement. Lorsque le gène *MmpL4b* subi une délétion chez la souche *Mabs* 390S, cela aboutit à la disparition des GPLs [182] et à une virulence exacerbée chez le poisson zèbre[185]. D'autre part, la mutation ponctuelle du gène *MmpL4a* dans le domaine transmembranaire TM10 s'accompagne également d'une perte des GPLs chez *Mbol*[350]. De même, la mutation de ce gène dans la souche *Mabs* CIP104536 inhibe la production des GPLs[351]. En manipulant le gène *MmpS4*, les auteurs obtiennent également un mutant avec une production intermédiaire de GPLs entre le niveau élevé de la forme lisse et l'absence de GPLs chez la forme rugueuse. Précédemment, la comparaison d'un extrait de lipides totaux avec celui de lipides faiblement liés issus de souches sauvage et mutante de *Msmeg*, a confirmé que le gène *MmpS4* ne participe pas directement à la biosynthèse des GPLs mais à leur export[352]. Une méthodologie identique avait mis en évidence un mutant de transposition dans le gène *Gap* qui possède un rôle

similaire[353]. C'est pourquoi l'ensemble des produits issus de ces gènes pourrait s'associer en complexe afin d'assurer le transport des GPLs.

I.B.3.b.viii Rôles biologiques des GPLs de *Mabs*

Le rôle des GPLs est fortement associé au phénotype de la morphologie lisse de *Mabs* qui a été précédemment développée en masquant l'effet proinflammatoire des PIM_2 reconnus par les récepteurs TLR-2 des cellules de l'immunité innée[180] et du tissu pumonaire[181]. Au niveau moléculaire, les GPLs acétylés et succinylés de *Msmeg* sont capables d'interférer avec la phagocytose d'autres espèces productrices de GPLs[346,345], participant donc à moduler ce processus par l'intermédiaire des récepteurs qui l'initient et des autres glycolipides présents.

Le rôle biologique des groupements acétyls a été évalué par Whang et collaborateurs en 2017 en démontrant que l'apoptose induite par la staurosporine, l'oligomycine et même par la forme rugueuse de *Mabs* sur des lignées de macrophages RAW 264.7 et BMDM était diminuée par les GPLs les plus polaires purifiés de la forme lisse de *Mabs*[354]. L'hydrolyse alcaline douce entraîne la perte de cette propriété qui est donc attribuée aux seuls groupements acétyles saponifiables. La diminution de l'apoptose est accompagnée d'une diminution de la production d'espèces réactives de l'oxygène par la mitochondrie. En effet, les GPLs ont été localisés dans la mitochondrie et interagissent avec la cyclophiline D préférentiellement lorsqu'ils sont di-*O*-acétylés.

Enfin, même si la présence des GPLs est le plus souvent corrélée à la capacité de former des biofilms, il a récemment été démontré que même la forme rugueuse était capable d'en produire[355]. Nous n'aborderons pas les multiples rôles attribués aux GPLs des sérovars de *Mav* compte tenu de la plus grande variabilité de la structure glycannique qui n'est pas représenté chez *Mabs*. En outre, certaines espèces produisent des peptidolipides qui ne sont pas glycosylés.

I.B.3.c LES LIPOPEPTIDES

I.B.3.c.i La fortuitine de *Mycobacterium fortuitum* (*Mfor*)

La fortuitine est le premier lipopeptide décrit dès 1963 par Vilkas chez *Mfor*. Il contient un acide gras docosanoïque ou éicosanoïque du côté *N*-terminal et des acides aminés de la forme L[356]. Plus tard en 1965, des expériences de fragmentation en spectrométrie de masse proposent la séquence Val-Leu-Val-Val-Leu-Thr-Thr-Ala-Pro comme noyau peptidique. Les résidus de leucine sont *N*-méthylés, ceux de thréonine sont acétylés et la proline est carboxyméthylé[357]. L'enchainement des acides aminés et leur modification seront confirmées ensuite par l'analyse RMN[358].

I.B.3.c.ii Plusieurs lipopeptides dans le complexe *Mycobacterium avium* (MAC)

Chez la souche MAC702 de l'espèce *Mycobacterium avium*, une recherche d'intermédiaires de la biosynthèse des GPLs s'est faite *via* un marquage avec de la phénylalanine ^{14}C puis de l'isoleucine ^{14}C. Malgré l'absence de dérivés de GPLs, elle a permis aux auteurs de purifier et de caractériser le lipopeptide I, dépourvu de monosaccharides mais composé d'acides aminés hydrophobes. La séquence Phe-(*N*-Me)Ile-Ile-Phe-Ala-Ile-Ala-Phe est estérifiée du côté *N*-terminal par un acide gras à 17 ou 19 carbones. Néanmoins, à l'image de la fortuitine, aucun gène participant à la biosynthèse de ce composé n'a été décrit dans ce travail[359].

Chez *Mycobacterium johnei* - aujourd'hui connu sous le nom de *Mycobacterium avium subs. paratuberculosis* (*Map*) - un lipopeptide est décrit par Lanéelle et collaborateurs en 1965. Le noyau peptidique est une séquence D-Phe-(*N*-Me)L-Val-L-Ile-L-Phe-L-Ala, méthyl-estérifié coté C-terminal et auquel s'ajoute un lipide de C_{17} à C_{22} côté N-terminal[360]. Au début des années 2000, ce même lipopentapeptide, renommé para-LP-01, décrite chez la souche K-10 de *Map* fait l'objet d'un regain d'intérêt car son pouvoir immunogénique laissait entrevoir des possibilités de détection spécifique de cette souche au sein du complexe MAC, à partir du peptide de synthèse. En parallèle, une étude *in silico* du génome de *Map* révèle trois clusters de gènes susceptibles de coder des peptides synthases. L'un de ces clusters consiste en cinq modules dont le premier est capable d'incorporer spécifiquement la forme L de la phénylalanine. De plus, ce module présente 75% de similarité avec le module d'incorporation du premier résidu de D-phénylalanine dans le nsGPL. De même le second cluster code le gène de la méthyltransférase capable de modifier le résidu de valine présente en seconde position[361]. Par contre, aucune délétion de ces clusters n'a été réalisée pour confirmer leur implication dans la biosynthèse du lipopentapeptide.

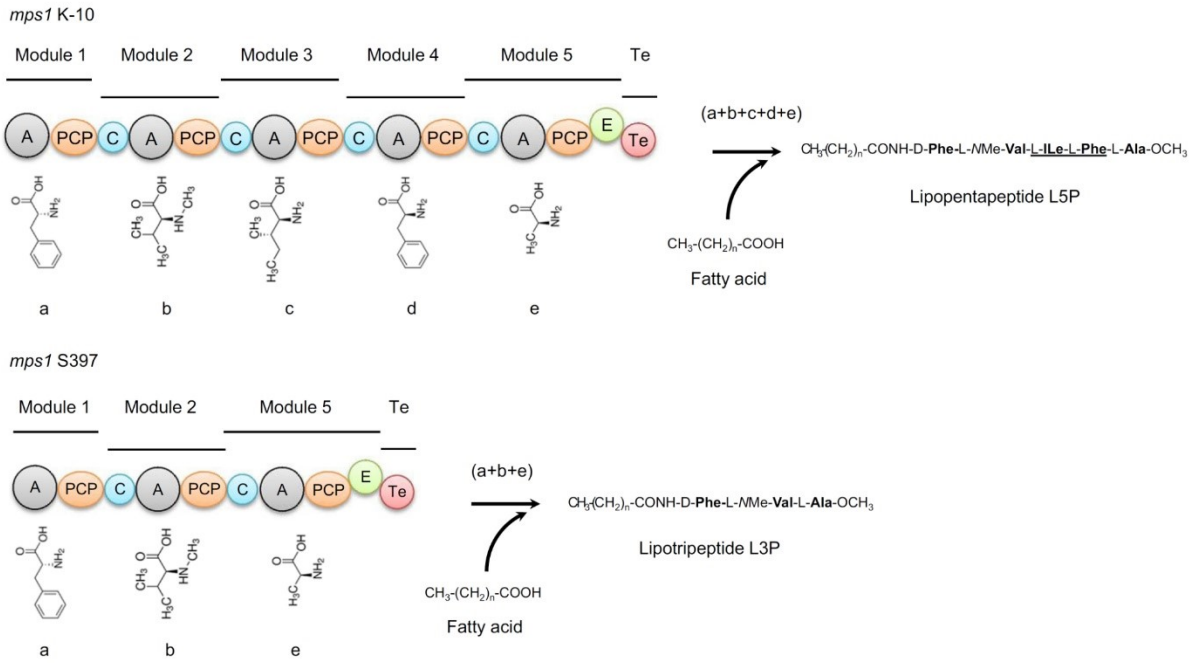


Figure 13 : Comparaison des peptides synthases présentes chez la souche K-10 et S397 de *Map* aboutissant au L5P et L3P d'après Bannantine[354].

En 2008, Biet et collaborateurs montrent que de nombreuses souches de *Map* identifiées chez les bovins, les caprins et les ovins produisent le lipopentapeptide mais pas les GPLs[362] . La comparaison des locus respectifs de biosynthèse avec *Msmeg* et la souche 104 de *Mycobacterium avium subs. hominissuis* (*Mah*) permet d'identifier des gènes codant des glycosyltransférases putatives dans les trois génomes. Néanmoins, seule la syntènie pour les peptides synthases non ribosomales *mps1* et *mps2,* pour les gènes *MmpL4a* et *MmpL4b* ainsi que pour les gènes associés à la synthèse du lipide est préservée. En parallèle, le lipopentapeptide synthètique est détecté spécifiquement dans les séra d'animaux infectés par *Map* mais pas ceux infectés par *M. bovis*, *Mah* ou *Mycobacterium avium intracellulare*. Cette réponse humorale est associée à un titre élevé d'IgG1, IgG2 et IgM. La comparaison de souches bovine et ovine de *Map* montre que la souche MAP-K10 produit le lipopentapeptide de séquence connue tandis que la souche S397 produit un lipotripeptide (L3P) de séquence D-Phe-(*N*-Me)L-Val-L-Ala. Le gène *mps1* de la souche S397 code effectivement trois modules au lieu de 5 (Figure 13). D'autre part, le L3P exprimé à la surface de l'enveloppe active la prolifération des lymphocytes TCD8 et le L5P celle des lymphocytes B chez la vache [363].

S'il permet de bien discriminer les espèces mycobactériennes, la sensibilité de détection des animaux infectés par *Map* grâce l'utilisation du lipopentapeptide synthètique reste inférieure à celle des tests déjà disponibles dans le diagnostic vétérinaire[364].

I.B.3.d LES PHENOLGLYCOLIPIDES (PGLs)

Les phénolglycolipides sont les seuls glycolipides mycobactériens contenant un noyau aromatique de type phénol substitué en position para par une chaine lipidique. De plus, le groupement phénolique est substitué par un oligosaccharide spécifique de l'espèce mycobactérienne (Tableau 6). L'acide gras est quant à lui majoritairement constitué par un motif phtiocérol qui porte deux acides mycocérosiques.

Tableau 6 : PGLs des principales espèces. En gras, la structure du noyau commun est relié au phénol. En italique, l'ajout ou la modification des résidus dans les nouvelles formes de PGLs ou les différentes souches montre leur diversité

ESPECES	GLYCOLIPIDES	NOYAU OLIGOSACCHARIDIQUE COMMUN *Motif ajouté à l'extrémité non-réductrice*	REFERENCES
Mkan	PGL-KI	**2,6-didéoxy-4-*O*- Me-α-L-Ara*p* (1-3)2-*O*- Me-4-*O*- acétyl-α-L-Fuc*p* (1-3) 2-*O*- Me-α-L-Rha*p* (1-3)2,4-*O*- Me_2-α-L-Rha*p***	
	PGL-KII	*2,4-O-Me $_2$ -α-D-Manp*	
	PGL-KIV	*4-O-Me-α-D-Manp*	
	PGL-K8	**2-*O*- Me-α-L-Fuc*p* (1-3)2-*O*- Me-α-L-Rha*p* (1-3)2-*O*- Me-α-L-Rha*p***	
	PGL-K7	**2-*O*- Me-α-L-Fuc*p* (1-3)2-*O*- Me-α-L-Rha*p* (1-3)2,4-*O*- Me_2-α-L-Rha*p***	
	PGL-K6	*2,6-didéoxy-4-O-Me-α-L-Arap(1-3)*	
	PGL-K5	**2,6-didéoxy-4-*O*- Me-α-L-Ara*p* (1-3)2-*O*- Me-4-*O*- propionyl-α-L-Fuc*p* (1-3)2-*O*- Me-α-L-Rha*p* (1-3)2,4-*O*- Me_2-α-L-Rha*p***	
Mbov	PGL-B1	**2-*O*- Me-α-L-Rhap**	
	PGL-B3	*α-L-Rhap(1-3)*	
	PGL-B2	**α-L-Rhap**	
Mtb	PGL-I	**2-*O*- Me-α-L-Rhap**	
	PGL-III	*2,3,4-O-Me $_3$ -α-L-Fucp(1-3)α-L-Rhap*	
	PGL-II	**α-L-Rhap**	
	PGL-IV	*2,3,4-O-Me $_3$ -α-L-Fucp(1-3)α-L-Rhap*	
	souches O/K	*2,4-O-Me $_2$ -α-L-Fucp(1-3)α-L-Rhap*	
Mmar		**3-*O*- Me-α-L-Rhap**	
Mul		**3-*O*- Me-α-L-Rhap**	
Mlep	PGL-II	**2,3-*O*- Me_2-α-L-Rha*p* 1(1-2)3-*O*- Me-α-L-Rhap**	
	PGL-III	*3-O-Me-β-D-Glcp(1-4)*	
	PGL-I	*3,6-O-Me2-β-D-Glcp*	

Les phénolglycolipides ont été décrits dès les années 1960 en tant que mycoside A chez *Mkan*[365], mycoside B chez *Mbov*[366] ou mycoside G chez *Mmar*[367]. La morphologie observée des colonies mycobactériennes trouvait alors une partie de son explication dans le contenu en mycosides selon l'espèce. En 1961, MacLennan et collaborateurs montrent que le mycoside A et le mycoside B contiennent du fucose et du rhamnose[368]. En effet, alors que le mycoside A libère du 2-*O*-Me-Fuc, du

2-*O*-Me-Rha et du 2,4-*O*-Me_2-Rha le mycoside B ne contiendrait qu'un unique 2-*O*-Me-Rha, supposée de forme D et lié en β au groupement phénol. En 1970, l'analyse de la partie lipidique du mycoside A et du mycoside B révèle simplement une différence de deux groupements méthylènes dans la longueur de la chaine phtiocérol[369].

I.B.3.d.i Le mycoside A de *Mycobacterium kansasii* (*Mkan*)

En accord avec les travaux précédents, Fournié et collaborateurs identifient dans le PGL K-I de *Mkan* un nouveau monosaccharide, le 2,6-didéoxy-4-*O*-Me-α-L-Ara*p* lié en position 3 au noyau 2-*O*-Me-4-*O*-acétyl-α-L-Fuc*p*(1-3)2-*O*-Me-α-L-Rha*p*(1-3)2,4-*O*-Me_2-α-L-Rha*p*[370,371]. Ce résidu est remplacé par un groupement 2,4-*O*-Me_2-α-D-Man*p* dans le PGL K-II ou un résidu 4-*O*-Me-α-D-Man*p* dans le PGL K-IV[372]. Finalement, de nouveaux PGL notés 5 et 6 ont été respectivement caractérisées à partir du PGL I suite au remplacement du groupement 4-*O*-acétyl par un résidu propionyl ou à sa disparition (Tableau 6). D'autre part, les PGL 7 et 8 sont des variantes trisaccharidiques qui se distinguent par la présence ou l'absence du groupement 4-*O*-méthyl sur le résidu de rhamnose proximal[373].

I.B.3.d.ii Le mycoside B de *Mycobacterium bovis* (*Mbov*)

Concernant le mycoside B de *Mbov*, sa réévaluation décrit qu'un unique résidu 2-*O*-Me-α-L-Rha*p* est relié au groupement phénol[374,375]. Néanmoins, Vercellone et Puzo ont développé un protocole de purification chromatographique plus élaborée du mycoside B qui a révélé trois nouvelles structures B-1, B-2 et B-3 apparentées à ce dernier[376]. Le PGL-B1 possède un phénolphtiodolone au lieu du phénolphticoérol. Dans le PGL B-2, le rhamnose n'est pas méthoxylé et le noyau phénolphtiodolone demeure. Enfin, le rhamnose du PGL B-3 est lié en position 3 à un α-L-Rha*p* formant la seule variante disaccharidique connue (Tableau 6). D'autre part, le phénolphtiocérol se retrouve à la place du phénolphtiodolone.

La comparaison de la structure du mycoside B met en évidence un composé identique retrouvé chez *Mtb* (souche *canettii*), le PGL-I, ainsi que le PGL-II qui est la forme non méthoxylée en position 2 du rhamnose[374]. L'année précédente, les mêmes auteurs avaient étudié chez la même espèce le PGL-III (PGL-tb1) majoritaire et dérivé de cette forme monosaccharidique puisque sur la position 3 du rhamnose était reliée au motif 2,3,4-*O*-Me_3-α-L-Fuc*p*(1-3)α-L-Rha*p*[377,378]. En outre, l'analyse de la partie lipidique montra la présence de phénol phtiocérol A associé principalement à l'acide mycocérosique C_{29}. De même, le PGL-IV fut ensuite défini comme la forme trisaccharidique non méthylée en position 2 sur le rhamnose proximal[379]. En 1994, dans le but d'améliorer le diagnostic sérique basé sur la reconnaissance des PGLs, Watanabe et collaborateurs déterminent la structure du PGL-III (PGL-tb1)

majoritaire parmi deux autres souches O et K de *Mtb*[380] et identifie un autre PGL caractérisé par la présence d'un 2,4-*O*-Me_2-α-L-Fuc*p* distal et d'un motif phénolphtitriol au lieu du phtiocérol. Au début des années 2000, les 3 gènes *Rv2962c*, *Rv2958c* et *Rv2957c* codant les glycosyltransférases responsables de l'ajout du rhamnose sur le phénol, du second rhamnose et du fucose distal sont caractérisées dans la biosynthèse des PGLs[381].

I.B.3.d.iii Le mycoside G de *Mycobacterium marinum* (*Mmar*)

S'agissant du mycoside G de *Mmar*, l'analyse de l'aglycone montre des similarités structurales avec les mycosides A et B convergeant vers le phénolphtiocérol[382]. Trois ans plus tard, la même équipe identifie dans les souches MNC170 et MNC840 un PGL constitué d'un 3-*O*-Me-α-L-Rha*p*, d'un phénolphtiocérol à 39-37 carbones et d'acides mycocérosiques à 26-27 carbones [383]. Démontrant leur proximité taxonomique, le PGL identifié chez deux souches de *Mul* possède une partie glycanique identique à celui de *Mmar* alors que les lipides estérifiés sur le phtiocérol ne sont pas des acides mycocérosiques[384].

I.B.3.d.iv Les PGL de *Mycobacterium leprae* (*Mlep*)

Les difficultés liées à la culture de *Mlep* ont retardé d'une vingtaine d'années la découverte du premier PGL spécifique à cette espèce par sa partie glycanique. En effet, la comparaison avec le mycoside A de *Mkan* aboutit à des spectres infra-rouges très similaires mais l'analyse RMN ^{1}H révèle la présence de monosaccharides caractérisés par des signaux anomériques distincts. La détermination de la composition en monosaccharides démontrent que les 2,3-*O*-Me_2-Rha, 3-*O*-Me-Rha et 3,6-*O*-Me_2-Glc remplacent les 2,4-*O*-Me_2-Rha, 2-*O*-Me-Rha et 2-*O*-Me-Fuc du PGL de *Mkan*[385]. L'approfondissement de l'étude structurale a permis aux mêmes auteurs de déterminer non seulement la configuration des sucres mais aussi la position des liaisons entre les monosaccharides. Ils ont ainsi défini le PGL-I avec le motif trisaccharidique 3,6-*O*-Me_2-β-D-Glc*p*(1-4)2,3-*O*-Me_2-α-L-Rha*p*1(1-2)3-*O*-Me-α-L-Rha*p*[386] ainsi que le PGL-III qui diffère du PGL-I par l'absence du groupement méthoxy en position 3 du glucose[387]. Plus tard, Daffé et collaborateurs identifient une forme disaccharidique PGL-II dépourvue de glucose terminal[388].

II PRESENTATION DES OBJECTIFS

Dans le chapitre I, nous avons montré que *Mabs* appartient aux MNTs dont l'incidence est de plus en plus forte sur notre santé parallèlement à *Mtb*. L'apparition des clones dominants DCC1,2 et 3 sur plusieurs continents plaide en faveur d'une possibilité de transmission entre les patients, préférentiellement ceux atteints de mucoviscidose. Malheureusement, la très grande résistance antibiotique de *Mabs* complique les traitements lorsque l'infection est avérée, diminuant les chances de survie en deçà de 50%. A l'image d'autres MNTs, *Mabs* possède la particularité d'exister sous deux morphotypes, lisse et rugueux, dont les profils glycolipidiques diffèrent. Dans le cas de la forme lisse, une moindre virulence est associée à la production de grande quantité de GPLs présents à la surface de l'enveloppe. En effet, les GPLs de *Mabs* limitent non seulement la réponse proinflammatoire mais aussi l'apoptose des macrophages. Quant à ceux de *Msmeg*, ils interfèrent avec la phagocytose par les macrophages. Si la structure générale des GPLs de *Mabs* est très similaire à celle de la mycobactérie modèle *Msmeg*, leur biosynthèse, leur répartition ainsi que leur modification détaillée n'ont jamais fait l'objet d'une étude globalisée.

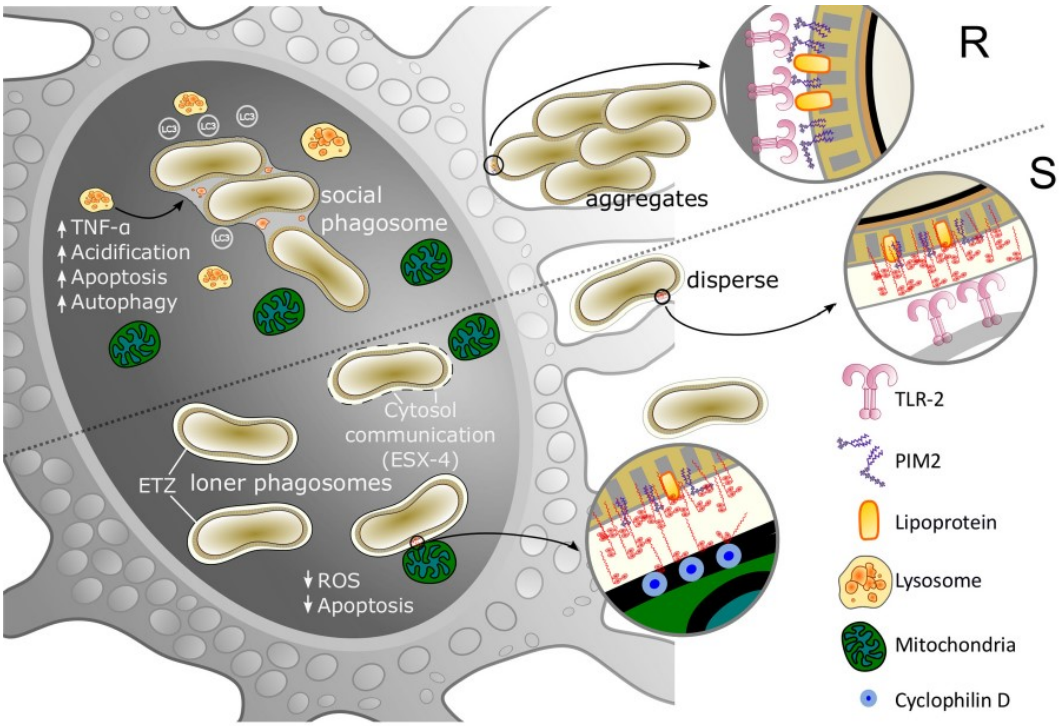


Figure 14 : Devenir de *Mabs* R et S après phagocytose d'après Gutierrez. Les 2 morphotypes rugueux (en haut) et lisse (en bas) de *Mabs* se comportent différemment dans les macrophages. Les PIMs de la membrane plasmique sont reconnus par les récepteurs TLR-2 induisant la production de TNFα et l'apoptose. A l'inverse, la présence de GPLs chez la forme lisse limite ce phénomène mais leur interaction avec la cyclophiline D de la mitochondrie empêche l'apoptose.

Dans le cas de la forme rugueuse, l'absence des GPLs provoque la production de TNFα induite via la reconnaissance des PIMs par TLR-2. Le même rôle a été décrit pour le TDM d'autres espèces de mycobactéries *via* le récepteur C-Mincle. Le GDND est un nouveau glycolipide décrit récemment chez *Mabs.* Son absence chez le mutant Δ*MmpL8* est corrélée avec une diminution de la production d'IL-1β par les macrophages THP-1 et du nombre de granulomes chez le poisson zèbre. Si le rôle immunomodulateur des glycolipides différemment exprimés par les 2 morphotypes justifie partiellement le cycle infectieux de *Mabs* (Figure 14), la possibilité de découvrir de nouveaux composés interagissant avec les cellules du système immunitaire s'explique par le nombre important de gènes codant les MmpL dont les substrats ne sont pas connus.

Mon travail de thèse s'inscrivait dans un projet financé par l'ANR nommé SunLive en collaboration avec l'équipe du Dr Laurent KREMER de l'IRIM à Montpellier. Les objectifs de ce projet étaient 1) d'établir le glycolipidome de *Mabs* avec une attention particulière sur la détermination structurale des glycolipides majeurs de l'enveloppe ; 2) de décrire le rôle de ces glycolipides complexes dans la physiopathologie mycobactérienne et l'immuno-modulation en étudiant leurs propriétés biologiques dans le macrophage, grâce à l'étude de mutants dans leur biosynthèse et/ou grâce à l'utilisation des glycolipides purifiés couplés à des billes ; 3) de corréler le rôle de ces glycolipides et le phénotype des mutants correspondants, dans la morphologie de *Mabs* et sa virulence chez différents modèles animaux.

Dans ce contexte, l'objectif général de mon projet de thèse a été d'établir une liste exhaustive des glycolipides retrouvés dans les deux morphotypes avec un intérêt particulier pour les GPLs. Mes travaux ont d'abord consisté à déterminer les étapes clés de la biosynthèse des GPLs et leur implication *in vitro* dans le cycle infectieux de *Mabs*. Je me suis en particulier focalisé sur trois familles de gènes nommément *fmt*, *gft1-3* et *aft1-2*. Après avoir évalué l'impact de la délétion du gène *fmt* sur la modification du lipide *N*-acylé, j'ai contribué à l'étude des gènes *gtf1*, *gtf2* et *gtf3* codants trois glycosyltransférases également annotées chez *Msmeg*. Finalement, mon attention s'est porté sur les deux gènes *atf1* et *atf2* codants des acétyltransférases et dupliqués par rapport au génome de *Msmeg*. Pour cela, j'ai purifié et identifié l'ensemble des GPLs à partir de la souche sauvage et de tous les mutants de délétion de ces gènes. En parallèle, l'hydrophobicité de l'enveloppe, l'aspect des colonies ou l'interaction avec les macrophages *in vitro* ont été confrontés. En complément, je me suis également attaché à décrire les autres glycolipides libres communs aux deux morphotypes. Puisque les GPLs sont présents en grande quantité dans la forme lisse, j'ai utilisé la forme rugueuse pour identifier essentiellement les mycolyltréhaloses, les DATs et les PIMs. Dans ce contexte, j'ai également

entrepris l'identification d'un glycolipide jusqu'alors inconnu dont la présence avait été détectée par des travaux préliminaires. En effet, ces derniers avaient montré que l'expression du gène *MAB_4690c* était différenciée entre trois souches rugueuses 19977-IP-R, 19977-AT-R et CF-R[174]. De plus, sa délétion entrainait une moindre mortalité chez l'embryon de poisson zèbre et la disparition d'un lipide en CCM (Figure 15A & 15B). Le gène *MAB_4690c* code une peptide synthase et appartient à un locus dont l'organisation rappelle celui de la biosynthèse des GPLs (Figure 15C). En effet, les gènes *MAB_4703c* et *MAB_4704c* codant des MmpL dont le substrat n'est pas connu se trouvent à proximité ainsi que 4 glycosyltransférases. C'est pourquoi, j'ai voulu établir un lien entre l'existence de ce locus et la présence de nouveaux GPLs chez la forme rugueuse de *Mabs*.

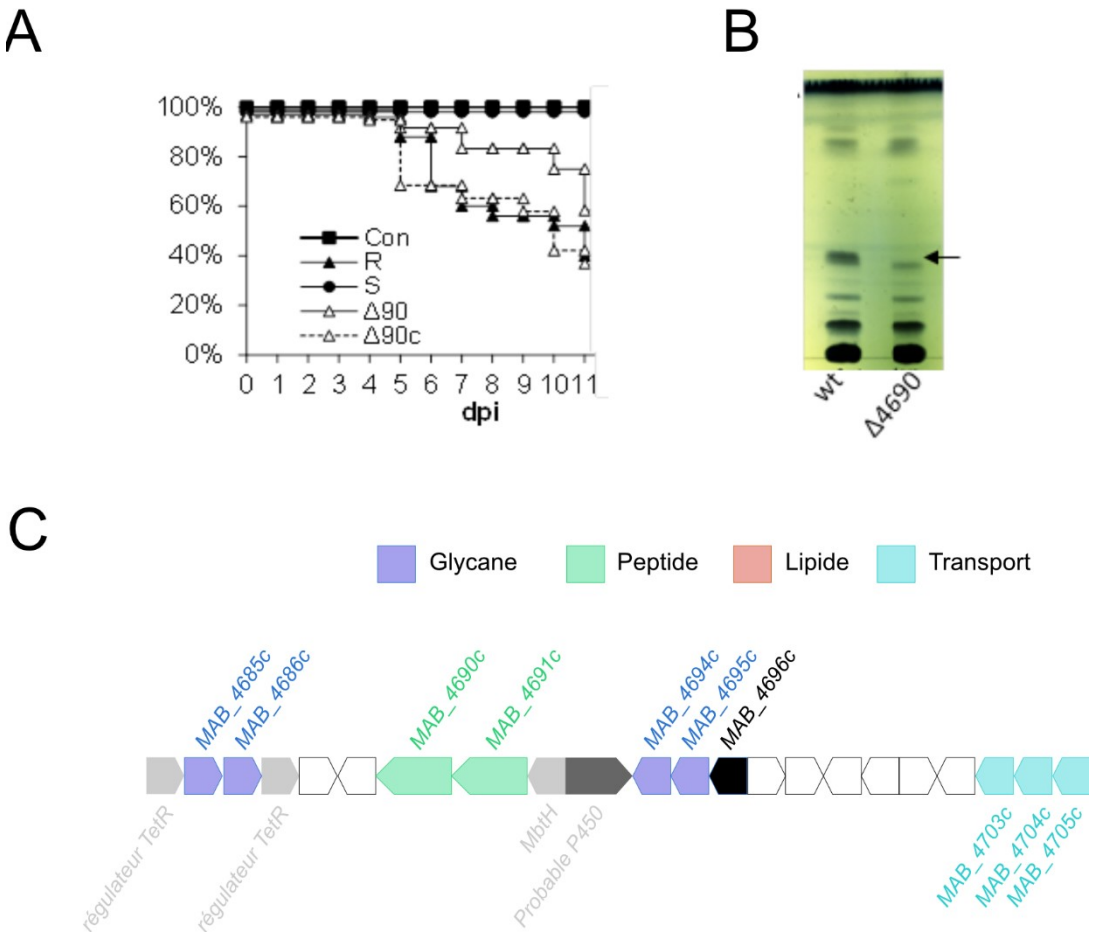


Figure 15 : Le locus autour du gène *MAB_4690c*. Rôles présumés du gène *MAB_4690c* dans la virulence de *Mabs* R chez l'embryon de poisson zèbre (A), dans le profil observé de lipides polaires (B) et dans l'organisation d'un locus semblable à celui impliqué dans la biosynthèse de GPLs. Seuls les gènes hypothétiquement responsables de la synthèse du noyau peptidique, de sa glycosylation et de son transport sont annotés. Excepté les gènes *TetR*, *MbtH* et probable *P450* les autres sont décrits dans le chapitre IV.E.7

III RESULTATS

Après avoir expliqué les méthodes d'extraction et de purification des glycolipides libres, nous détaillerons l'identification du TDM et d'un dérivé di-*O*-acylé du TMM contenu dans la fraction la plus apolaire. Dans un second temps, nous poursuivrons sur l'étude de la fraction polaire à travers la caractérisation de 3 dérivés di-*O*-acylés du tréhalose et des PIMs. Les résultats de ces analyses structurales sont uniquement présentés dans ce mémoire mais feront peut-être l'objet d'une revue sur les spécificités des glycolipides de *Mabs*. Nous reviendrons ensuite sur les résultats publiés s'agissant de la biosynthèse des GPLs par la forme lisse de *Mabs* et de leur impact sur l'interaction avec les macrophages ou sur la morphologie des colonies. Enfin, nous terminerons sur la description structurale précise d'un nouveau glycopeptidolipide nommé GPL4690 potentiellement impliqué dans la virulence de la forme rugueuse de *Mabs,* et pour lequel un 4ème article est en cours de rédaction.

III.A EXTRACTION DES GLYCOLIPIDES

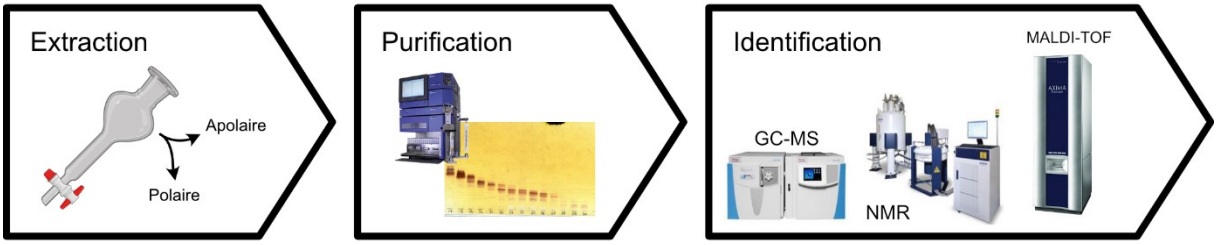


Figure 16 : Méthodologie générale d'analyse des glycolipides. L'ordre d'élution des glycolipides étudiés est indiqué selon le gradient dans le Tableau 26 du chapitre V.C « Matériels et méthodes ».

Les lipides pariétaux ont été séquentiellement extraits par une procédure précédemment mise au point[389]. Elle consiste à préparer d'abord un extrait contenant les lipides les plus apolaires par 3 extractions successives dans un mélange non miscible de méthanol/0,3% NaCl et d'éther de pétrole. Après centrifugation, les fractions éthérées sont regroupées, évaporées et solubilisées dans le dichlorométhane. Le culot est chauffé dans le méthanol, puis extrait à nouveau par deux solutions de chloroforme/méthanol/NaCl 0,3% et centrifugé. Les deux surnageants sont additionnés d'un même volume chloroforme/eau à 0,3% de NaCl et démixés. La phase chloroformique enrichie en lipides polaires, est recueillie, évaporée et reprise dans le chloroforme/méthanol (2:1). Afin de quantifier relativement les glycolipides des souches étudiées, une quantité identique de mycobactéries lyophilisées est utilisée.

III.B CONTROLE DES SOUCHES ET DES PURIFICATIONS PAR CCM

Après chaque extraction, les fractions apolaires et polaires ont été soumises à une séparation monodimensionnelle par chromatographie sur couche mince à basse température. Les systèmes de solvants ont permis de balayer une gamme de polarité suffisamment large pour résoudre les familles de composés recherchés (Tableau 7). Les glycolipides apolaires, les GPLs et les DATs ont pu être distingués grâce au mélange A et B. Les mélanges C,D,E se sont avérés nécessaires pour résoudre les GPLs déacétylés et les PIMs. Plusieurs techniques de révélation ont été utilisées afin de détecter spécifiquement les molécules selon leurs groupements fonctionnels. L'utilisation de l'orcinol-sulfurique met en évidence les glycolipides en produisant des nuances variées de couleur : violette pour les hexoses, orange pour les déoxyhexoses et bleu pour les pentoses. La révélation à l'acide phosphomolybdique renseigne sur le contenu global en lipides. Enfin, l'imprégnation avec le di-iode révéle les glycolipides de façon réversible afin de les purifier par CCM préparative après évaporation des vapeurs d'iode.

Tableau 7 : Mélange de solvants utilisés pour la séparation CCM des différentes familles de glycolipides étudiés ainsi que des esters méthyliques d'acides mycoliques (MAMEs).

Mélange	Chloroforme/acide acétique/méthanol/eau	Glycolipides
A	90/0/10/1	TDM, Ac_2TMM, GPLs, DATs
B	95/0/5/0	GPLs
C	65/0/25/4	DATs et PIMs, dGPLs
D	30/0/8/1	TMM et PIMs, dGPLs
E	40/25/3/6	PIMs
F	Hexane/ethylacétate (95/5)	MAMEs

III.C PURIFICATION DES EXTRAITS

La purification des extraits de lipides apolaires (Figure 17) s'est faite essentiellement grâce à la chromatographie d'adsorption sur gel de silice au moyen d'un instrument de chromatographie flash. L'utilisation de colonnes de granulométrie 15µm donne accès à des purifications rapides et très

résolutives. Les gradients de méthanol dans le chloroforme ont été optimisées pour chaque composé recherché afin d'obtenir les fractions les plus pures. Cette procédure simple s'est révélée suffisante dans la purification du TDM et d'un dérivé di-*O*-acylé du TMM. Néanmoins, ce dernier n'a pas pu être isolé et caractérisé de façon certaine. De même, malgré un essai de purification utilisant l'éther de pétrole PE 60-80 selon la méthodologie décrite par Llorens[176], les TPP n'ont pas été étudiés.

Dans le cas des lipides polaires (Figure 17B & C), l'analyse des GPLs s'est faite également grâce à la chromatographie d'adsorption sur gel de silice avec des proportions de méthanol n'excédant pas 10%. Dans le cas de la souche sauvage, Le GPL-2a est élué avant le GPL-3 plus polaire. Toutes les autres formes de GPL caractérisées présentaient des R_f s'étalant du GPL-2a jusqu'au 1er DAT. Le regroupement des fractions enrichies en GPL a été suivi d'une deuxième séparation par CCM préparative, en particulier pour la préparation des tests *in vitro*. Les DATs sont élués immédiatement après les GPLs lors de la chromatographie d'adsorption à partir de 10% de méthanol. Malgré la charge négative portée par les PIMs, le recours à une séparation par échange d'anions n'a pas été nécessaire pour préparer des quantités satisfaisantes d'AcPIM_2. L'utilisation de proportions plus importantes de méthanol allant jusqu'à 50% a produit des fractions enrichies exploitables. Dans le cas du GPL4690, un second mode de séparation par phase inverse utilisant un gradient de méthanol dans l'eau en présence de 0,1% de TFA a été appliqué.

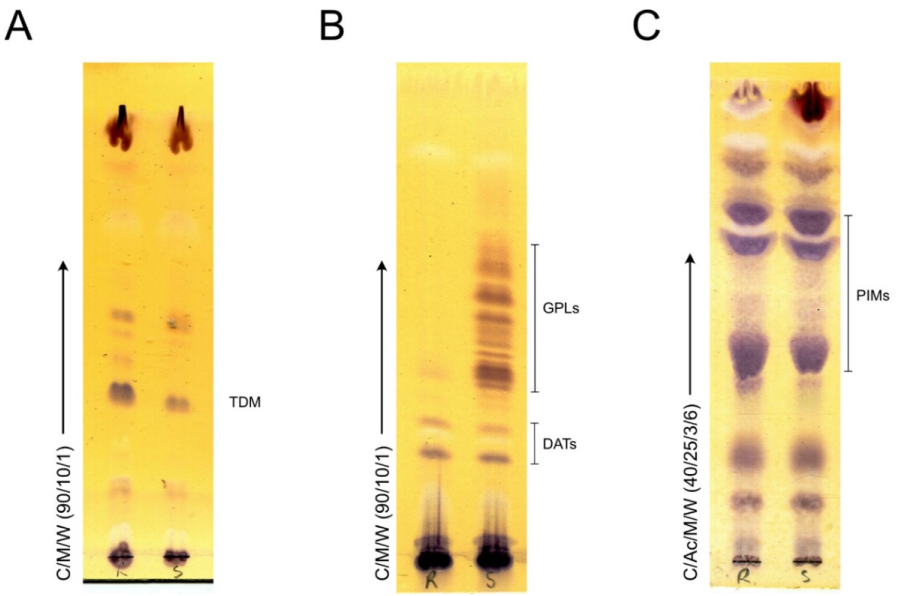


Figure 17 : Profil CCM de *Mabs*. Séparation des lipides apolaires (A) et polaires (B & C) des 2 formes rugueuse (à gauche) et lisse (à droite). Les familles de composés étudiés sont repérées sur la droite de chaque CCM. Dans le cas des lipides apolaires, Le front de migration contient les triglycérides.

III.D ETUDE DE LA FRACTION APOLAIRE

Pour éviter les contaminations par les GPLs partiellement extraits dans la fraction éther de pétrole à partir de la forme lisse, la plupart des lipides apolaires ont été obtenus à partir d'une souche rugueuse après avoir été confirmés par CCM dans les deux morphotypes. D'autre part, les extraits contenaient une grande quantité de lipides neutres sous forme de triglycérides qui ont été éliminés lors des purifications par le chloroforme pur.

III.D.1 PURIFICATION DE DEUX DERIVES DU TREHALOSE ET D'ACIDES MYCOLIQUES

A partir de la fraction apolaire de *Mabs* R cultivé sur milieu 7H10 OADC, j'ai réalisé une purification sur colonne de silice au moyen d'un gradient de méthanol dans le chloroforme. Plusieurs fractions éluées à 10% de méthanol ont révélés des bandes de couleur violette après séparation CCM, et révélation avec l'orcinol-sulfurique (Figure 18A). Les fractions 34 et 37 contenaient majoritairement du glucose identifié sous forme de dérivés itols-acétates par GC-MS (Figure 18B & C). Les fractions 36 à 40 ont été regroupées et ont montré un comportement chromatographique comparable à celui du TDM de *Mmar* qui était disponible au laboratoire (Figure 19A). De plus, la recherche de tréhalose sous forme perméthylé par l'analyse GC-MS a indiqué qu'il était présent en plus forte quantité dans la fraction contenant le TDM natif que dans la fraction 34. Les fractions 33 et 34 sont plus apolaires que le TDM d'après leur R_f plus élevé et ont donc été rassemblées.

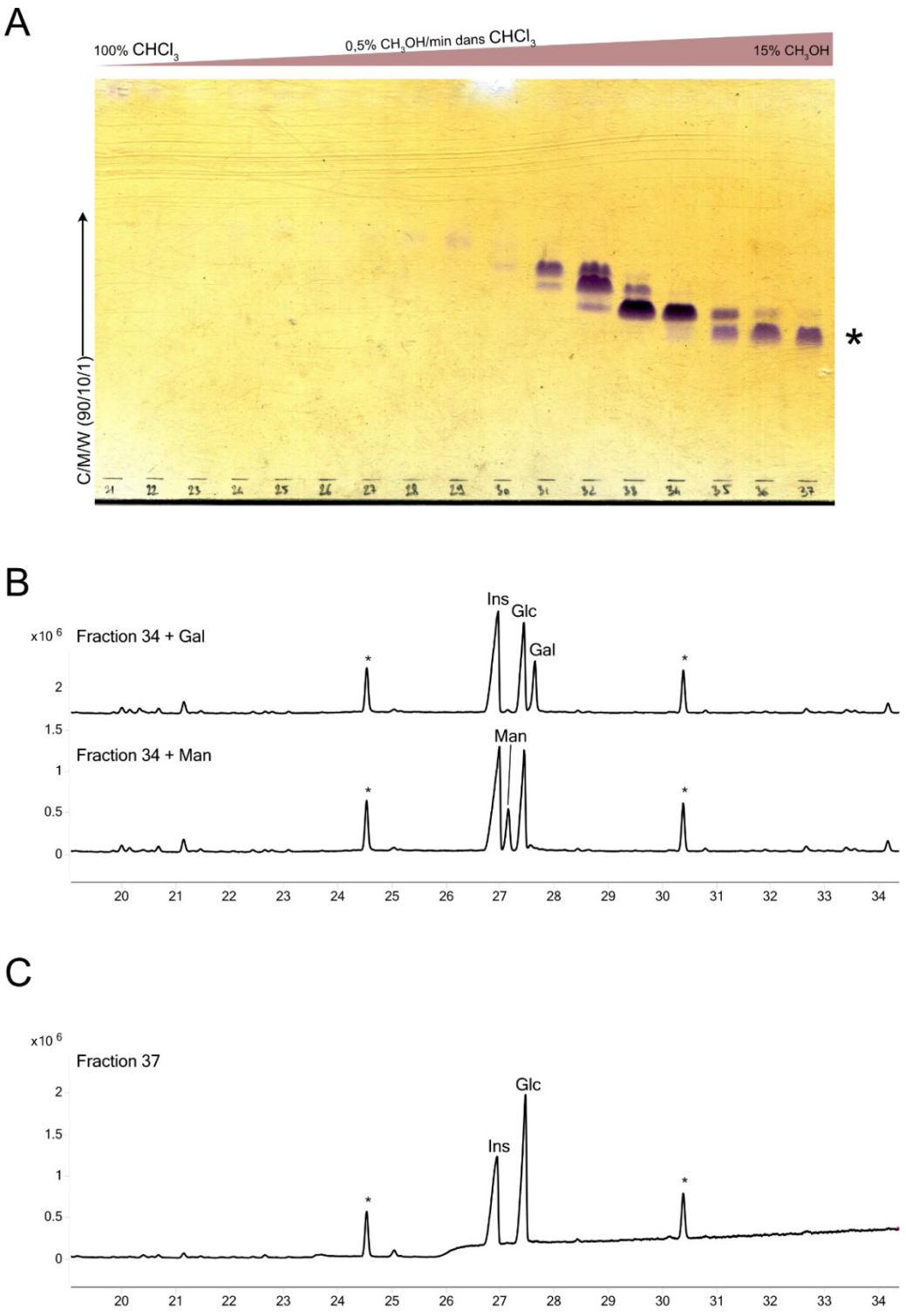


Figure 18 : Caractérisation des lipides apolaires de *Mabs*. Purification de l'extrait apolaire (A) et détermination de la composition en monosaccharide du dérivé di-*O*-acylé du TMM de la fraction 34 (B) et du TDM présent dans la fraction 37 (C). Les chromatogrammes de la fraction 34 ont été additionnés de galactose ou de mannose exogènes afin d'exclure la présence de ces 2 monosaccharides.

III.D.1.a Analyse MALDI-TOF du TDM natif

Le spectre MALDI-TOF de la fraction correspondant au TDM natif révèle un ion majeur de rapport *m/z* 2216 (Figure 19B). D'après cette masse, les deux acides mycoliques de type α' possèdent 64 carbones. Un autre massif se distingue autour de l'ion *m/z* 2425, ce dernier correspondant au remplacement par un acide mycolique de type α comptant 79 carbones. L'ion de rapport *m/z* 2633 équivalent au TDM avec 2 acides mycoliques de type α à 79 carbones est très peu intense. Le spectre de fragmentation de l'ion majoritaire *m/z* 2216 s'explique par la rupture de la liaison glycosidique entre les deux unités glycosyl qui a produit deux fragments *m/z* 1110 et 1128 décalés d'un atome d'oxygène. Le fragment *m/z* 1640 résulte de la perte de la chaine méromycolique (Figure 19C).

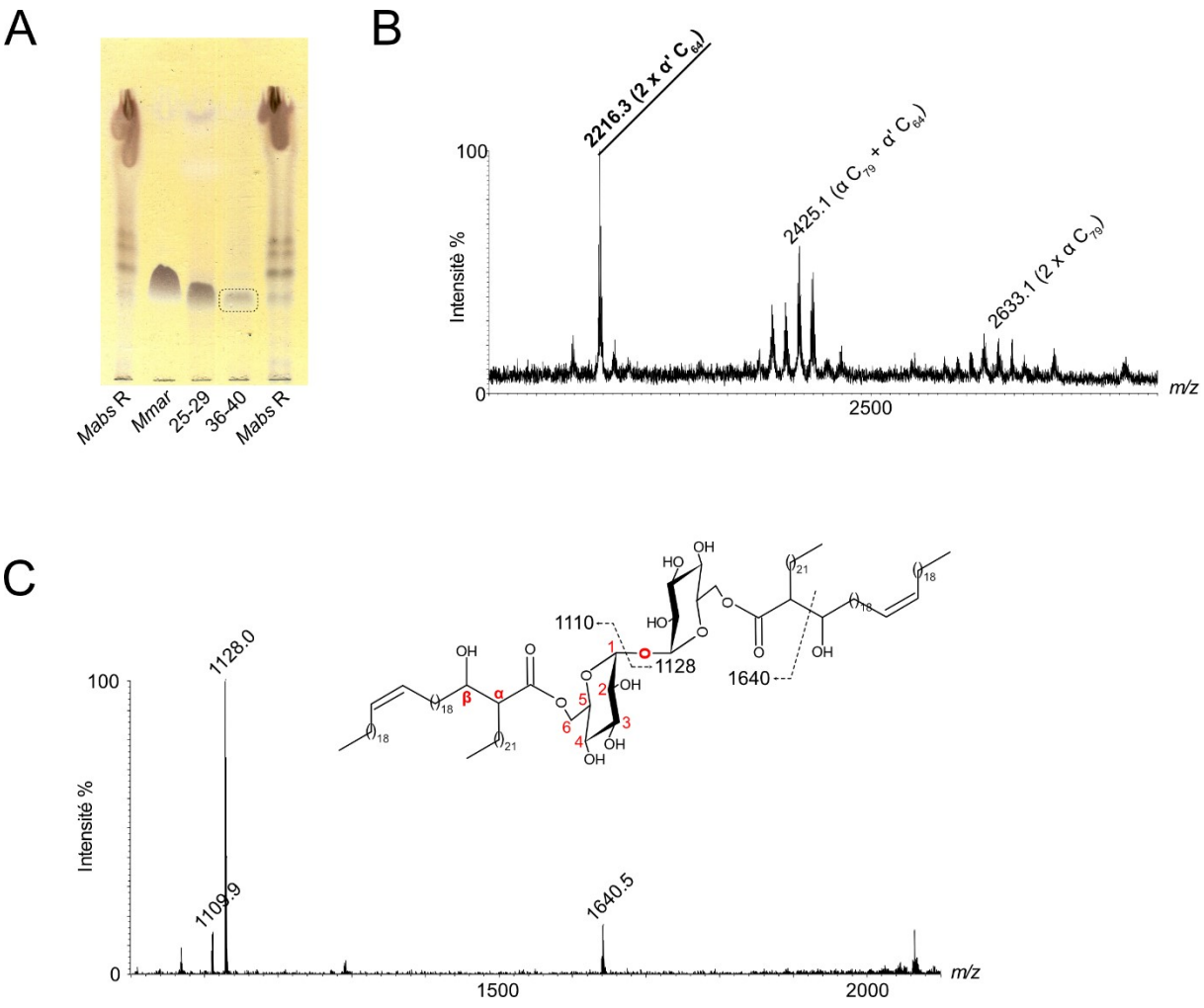


Figure 19 : Analyse MALDI-TOF du TDM de *Mabs*. Comparaison de la migration du TDM provenant de *Mabs* avec celui de *Mmar* (A). Spectre MALDI MS (B) et MS^2 (C) du TDM de *Mabs*. L'ion parent *m/z* 2216 correspond à la forme la plus abondante et contenant 2 acides mycoliques α' comptant chacun 64 carbones.

III.D.1.b Analyse du TDM natif par RMN

Le spectre $^{1}H/^{13}C$ HSQC du TDM natif présente sept signaux attribuables à des fonctions méthines entre 50 et 100 ppm (Figure 20A) qui sont corrélés entre eux dans le spectre $^{1}H/^{1}H$ COSY (Figure 20B), et qui correspondent au système de spin d'un résidu hexose. La constante de couplage scalaire J^3 entre les protons des positions 1 et 2 est égale à 3,4 Hz. D'autre part, les constantes J^3 entre les positions 2 et 3, 3 et 4 puis 4 et 5 varient de 9,3 Hz à 9,5 Hz. La succession des constantes de couplage confirme que l'hexose détecté est donc une unité α-glucosyl (Tableau 8).

Tableau 8 : Signaux RMN du tréhalose du TDM. Déplacements chimiques ^{1}H et ^{13}C de l'unité α-glucopyranose du TDM mesurées dans le solvant $CDCl_3/CD_3OD$ (2:1) à 293K après calibration sur le méthanol.

	Glc -1	*Glc -2*	*Glc -3*	*Glc -4*	*Glc -5*	*Glc -6/6'*
δ ^{1}H (ppm)	5,03	3,46	3,72	3,28	4,03	4,14/4,46
Multiplicité	D	Dd	T (Dd)	T (Dd)	M	Dd
J^3 (Hz)	3,4	3,4/9,4	9,3/9,3	9,5/9,5		6/11,7
δ ^{13}C (ppm)	94,5	72	73,4	71,2	70,5	64

Les deux signaux $^{1}H/^{13}C$ observés pour la position 6 et 6' à 4,14/64 et 4,46/64 ppm possèdent un déblindage plus élevé que la position 5 à 4,03/70,5 ppm ainsi qu'une constante de couplage de 11,7 Hz entre eux, ce qui est due à sa possible acylation. Compte tenu de la symétrie centrale de la molécule de TDM, les signaux des deux unités α-glucosyl sont parfaitement superposés (Figure 20A & B).

De plus, deux signaux à 2,4/53,1 ppm et à 3,64/72,9 ppm sont corrélés entre eux ainsi qu'à des groupements CH_2 sur le spectre $^{1}H/^{1}H$ COSY. Il s'agit des positions α-alkyle et β-hydroxy associés à la présence des acides mycoliques. Enfin, les signaux à 5,31/130,3 ppm et 5,1/125,5 ppm sont attribuables à des fonctions éthyléniques (Figure 20A), et une faible corrélation entre les signaux -0,36 ppm et 0,57 ppm sur le spectre $^{1}H/^{1}H$ COSY est attribuable à une fonction cyclopropyl.

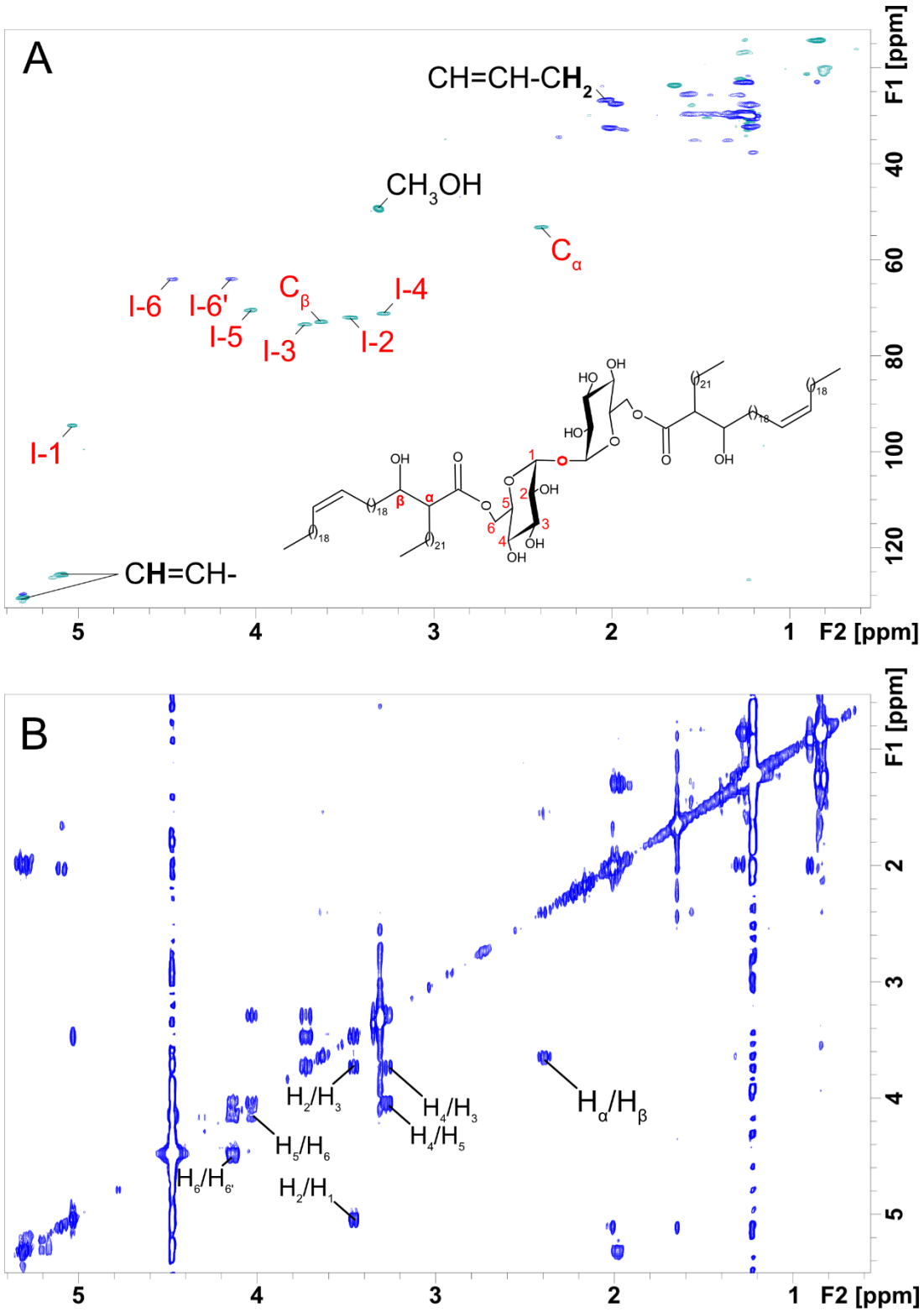


Figure 20 : Analyse RMN du TDM de *Mabs*. Spectres $^{1}H/^{13}C$ HSQC (A) et $^{1}H/^{1}H$ COSY (B) du TDM de *Mabs* mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Les positions 1 à 6 de chaque unité α glucopyranosyl sont superposées ainsi que les signaux des positions α et β des 2 acides mycoliques du fait de la symétrie centrale. Le spectre est calibré sur le méthanol et la zone plus blindée n'est pas annotée. Les signaux en bleu sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2 et ceux en vert à des CH.

III.D.1.c Analyse RMN du dérivé di-*O*-acylé natif du TMM

De façon surprenante, le spectre $^{1}H/^{13}C$ HSQC de l'Ac_2TMM met en évidence au moins quatre signaux anomériques $^{1}H/^{13}C$ (Figure 21A). A partir du signal de la position I-1 à 5,06/94,8 ppm, une corrélation $^{1}H/^{1}H$ COSY est détectée avec le signal à 3,48/71,8 ppm correspondant à la position I-2. Il s'ensuit les corrélations avec les positions I-3 à 3,76/73,5 ppm, I-4 à 3,44/70,3 ppm et I-5 à 3,57/72,8 ppm. La succession des quatre constantes de couplage J^3 implique qu'il s'agit bien d'une unité α-glucosyl (Tableau 9). Les deux protons de la position I-6 correspondent aux signaux ^{1}H 3,65 et 3,72 ppm alignés sur le signal ^{13}C 61,3 ppm.

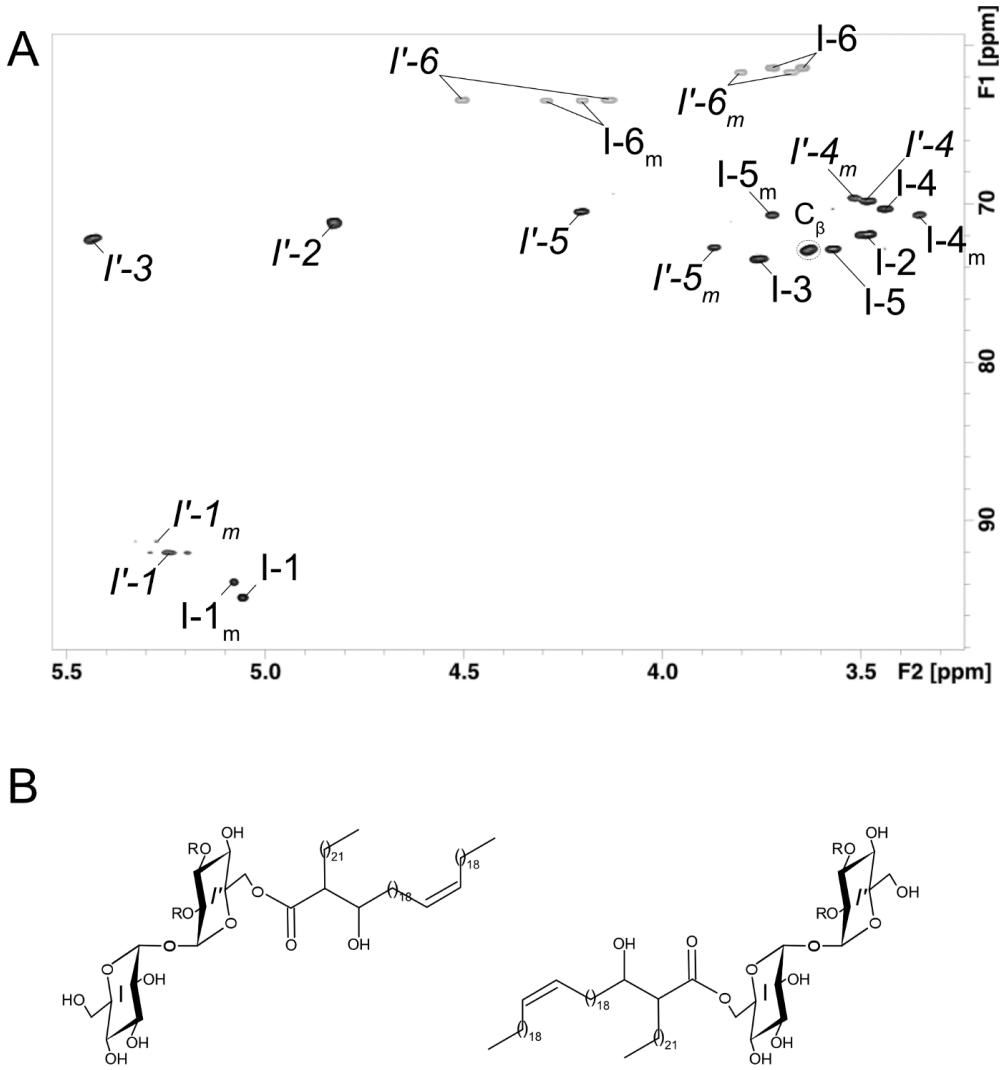


Figure 21 : Analyse RMN de l'Ac_2TMM. Spectre $^{1}H/^{13}C$ HSQC du dérivé di-*O*-acylé du TMM (A) mesurés dans le $CDCl_3/CD_3OD$(2:1) à 293K. Le grand nombre de signaux observables a été attribué au fait que la fraction correspond à un mélange de 2 isomères de position 6 ou 6' pour l'acylation de l'acide mycolique (B). Le composé majoritaire est celui dont l'un des 2 α-glucopyranoses n'est pas acylé. Les signaux en vert sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2.

De l'autre côté de la liaison glycosidique, une corrélation $^{1}H/^{1}H$ NOESY est observée avec le signal *I'*-1 à 5,24/92 ppm (Figure 22A). Ce dernier possède une corrélation avec *I'*-2 à 4,82/71,15 ppm lui-même corrélée à *I'*-3 à 5,43/72,1 ppm. Le déblindage du signal ^{1}H de ces deux positions indique qu'elles sont acylées (Figure 22B). *I'*-4 à 3,48/69,8 ppm est corrélée avec *I'*-5 à 4,2/70,4 ppm. Enfin, ce signal est corrélé avec ceux des deux protons *I'*-6 à 4,15 et 4,50/63,4 ppm dont le carbone est également acylé.

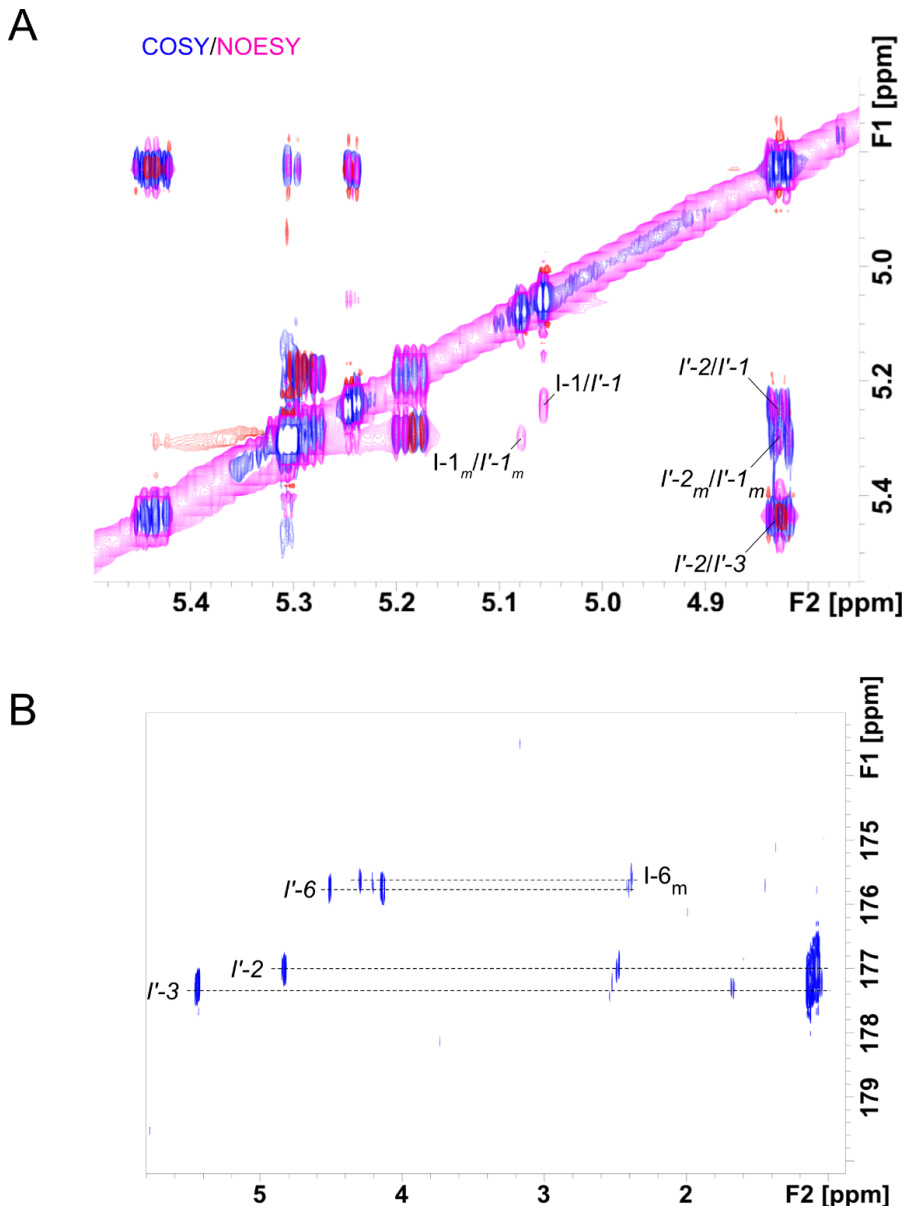


Figure 22 : Acylation de l'Ac_2TMM. Superposition des spectres $^{1}H/^{1}H$ COSY et $^{1}H/^{1}H$ NOESY (A) et $^{1}H/^{13}C$ HMBC (B) du dérivé di-*O*-acylé du TMM, mesurés dans le $CDCl_3/CD_3OD$(2:1) à 293K. Le spectre $^{1}H/^{1}H$ NOESY montre non seulement la liaison entre les positions anomères mais aussi les corrélations avec les positions acylées *I'-2* et *I'-3*. La corrélation des protons *I'-2, I'-3, I'-6* et $I\text{-}6_m$ avec les groupements carbonyles distinguent quatre signaux carbonyles liés à un acide mycolique et à un acide gras court polyméthylbranché.

Les deux autres signaux anomériques à 5,08/93,9 ppm et 5,3/91,3 ppm sont aussi corrélés sur le spectre $^{1}H/^{1}H$ NOESY (Figure 22A). Ils correspondent à une 2ème molécule qui est acylée en position I-6 comme le montre le plus fort déblindage de ses signaux à 4,2 et 4,29/63,5 ppm ainsi que celui I-5 à 3,72/70,6 ppm. A l'inverse, les signaux *I'-5* et *I'-6* sont davantage blindés (Tableau 9). Il s'agit donc d'un isomère de position présent dans une proportion de 0,5 équivalent au signal à 5,06/94,8 ppm. Ce dernier ne peut pas être séparé par chromatographie d'adsorption (Figure 21B).

Tableau 9 : Signaux RMN du tréhalose de l'Ac_2TMM. Déplacements chimiques ^{1}H et ^{13}C du dérivé di-*O*-acylé du TMM, mesurées dans le solvant $CDCl_3/CD_3OD$ (2:1) à 293K après calibration sur le méthanol. La détermination des constantes de couplage J^3 confirme la nature α-glucopyranose de l'hexose I.

			Glc -1	Glc -2	Glc -3	Glc -4	Glc -5	Glc -6/6'
Majoritaire	I	δ ^{1}H (ppm)	5,06	3,48	3,76	3,44	3,57	3,65/3,72
		Multiplicité	**Dd**	Dd	T (Dd)	T (Dd)	M	2,7/12,1 et 3,6/11,9
		J^3 (Hz)	**1/3,5**	3,5/9,7	9,3/9,3	9,5/9,5		
		δ ^{13}C (ppm)	94,8	71,8	73,5	70,3	72,8	61,3
	I'	δ ^{1}H (ppm)	5,24	4,82	5,43	3,48	4,2	4,15/4,50
		Multiplicité	**D**	**Dd**	M	M	M	1,6/12
		J^3 (Hz)	**3,5**	**3,5/10**				
		δ ^{13}C (ppm)	92	71,15	72,1	69,8	70,4	63,4
minoritaire	I	δ ^{1}H (ppm)	5,08	3,5	3,76	3,35	3,72	4,20/4,29
		Multiplicité	**Dd**	Dd	T (Dd)	T (Dd)	M	**Ddd**
		J^3 (Hz)	**1,7/3,6**	3,5/9,6	9,4/9,4	9,5/9,5		**2,5/5,2/8**
		δ ^{13}C (ppm)	93,9	72	73,5	70,7	70,6	63,5
	I'	δ ^{1}H (ppm)	5,3	4,83	5,44	3,52	3,87	3,68/3,80
		Multiplicité	**D**	**Dd**	M	M	Ddd	Dd
		J^3 (Hz)	**3,5**	**3,5/10**			2,5/5,3/7,8	3,5/12,3
		δ ^{13}C (ppm)	91,3	71,13	72,2	69,7	72,8	61,6

Parmi les signaux associés aux acides mycoliques, ceux des liaisons éthyléniques *cis* et *trans* sont présents à 5,3/130,1 ppm, ceux des signaux α-alkyle à 2,39/53,22 ppm et β-hydroxy à 3,63/72,9 ppm sont corrélés entre eux sur le spectre $^{1}H/^{1}H$ COSY. Le spectre $^{1}H/^{13}C$ HMBC confirme la corrélation du groupement α-alkyle et des protons méthyléniques I-6 et I'-6 avec une fonction carbonyle (Figure 22B). De plus, la corrélation des protons avec deux groupements carbonyles dont le déplacement chimique ^{13}C est mesuré à 177,01 et 177,27 ppm confirme l'acylation des positions I'-2 et I'-3. De l'autre côté, ces deux groupements carbonyles sont corrélés avec deux signaux de CH autour de 2,5/38 ppm puis deux signaux méthyls autour de 1,09 et 17,3 ppm, impliquant l'existence de chaine d'acides gras polyméthylbranchés (Figure 22B). A l'inverse, ses caractéristiques de la chaine lipidique n'ont pas été observées dans le TDM natif.

III.D.1.d Analyse MALDI-TOF du dérivé di-*O*-acylé natif du TMM

Le spectre de masse de l'Ac2TMM montre deux ions intenses *m/z* 1570 et 1612 et un massif plus faible autour des ions *m/z* 1779 et 1821 (Figure 23A). De plus, un ion intense de rapport *m/z* 980 a été détecté lorsque la fraction a subi une méthanolyse prolongée de 24h à 100°C. Cet ion est attribuable à un ester d'acide mycolique de type α' possédant 64 carbones. Un 2ème ion moins intense de rapport *m/z* 1188 correspond à l'ester d'acide mycolique de type α avec 79 carbones (Figure 23B). La présence majoritaire de ces deux acides mycoliques explique le décalage général de 208 u.m.a. constaté sur le spectre de la fraction native.

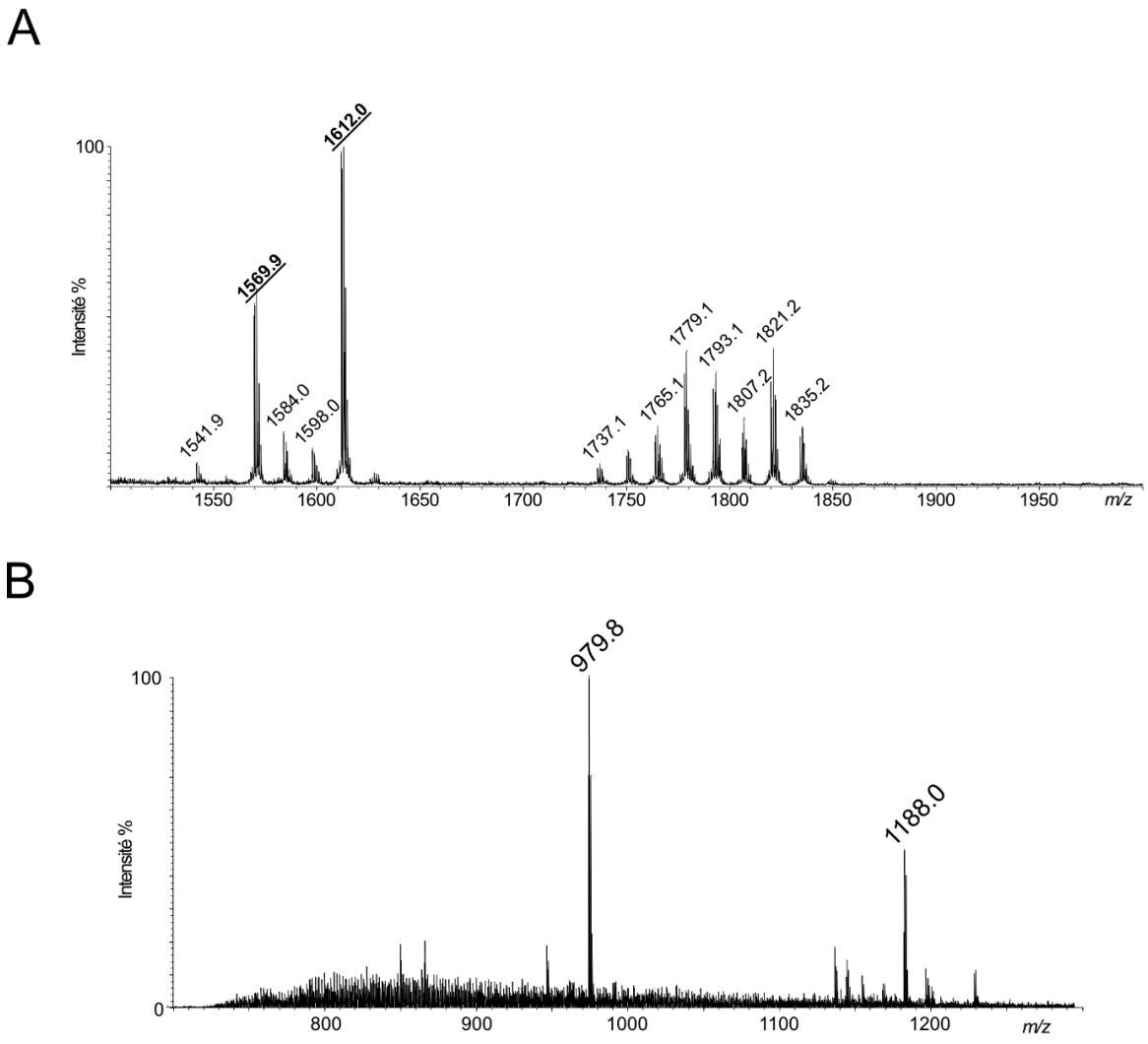


Figure 23 : Analyse MALDI-TOF de l'Ac_2TMM. Spectres de la fraction native (A) et méthanolysée (B) du dérivé di-*O*-acylé du TMM. L'ion *m/z* 1569.9 contient un acide mycolique α C_{64} alors que l'ion *m/z* 1779.1 a la forme α' C_{79}. Ces 2 acides mycoliques sont bien présents sous forme des esters méthyliques à *m/z* 980 et 1188.

La fragmentation des deux ions *m/z* 1570 et 1612 produit un ion fragment inférieur de 162 u.m.a. à l'ion parent résultant de la perte d'un hexose (Figure 24A). Ce résultat est cohérent avec la RMN puisqu'un glucose non acylé avait été identifié. Si l'on considère le spectre MS^2 de l'ion à *m/z* 1570, d'autres ions fragments apparaissent à *m/z* 1413 et 1251. Pour le premier, la différence de masse de 157 u.m.a. est compatible avec le départ d'un acide gras C_9, et la perte de l'hexose produit ensuite le second. Sur le spectre MS^2 de l'ion à *m/z* 1612, davantage d'ions fragments sont présents à *m/z* 1455, 1451 et 1413. Le second correspond à la perte du glucose alors que les deux autres résultent du départ d'un acide gras à 9 ou 12 carbones, comme le confirme les fragments à *m/z* 1292 et 1250 du spectre MS^3 du parent à *m/z* 1450 (Figure 24B). Enfin, les ions fragments à *m/z* 1110 et 1128 sont communs avec les fragments produits par l'ion parent à *m/z* 2216 dans le cas du TDM. Il s'explique par la rupture de la liaison glycosidique, et la perte de la chaine méromycolique produit les ions M-575 à *m/z* 995 et 1037 (Figure 24C).

L'ensemble des résultats obtenus sur la fraction correspondant au dérivé di-*O*-acylé du TMM démontre non seulement qu'il contient les mêmes acides mycoliques que le TDM en position 6, mais aussi que les positions 2 et 3 sont acylées par des acides gras courts et polyméthylbranchés aboutissant à un mélange d'isomères.

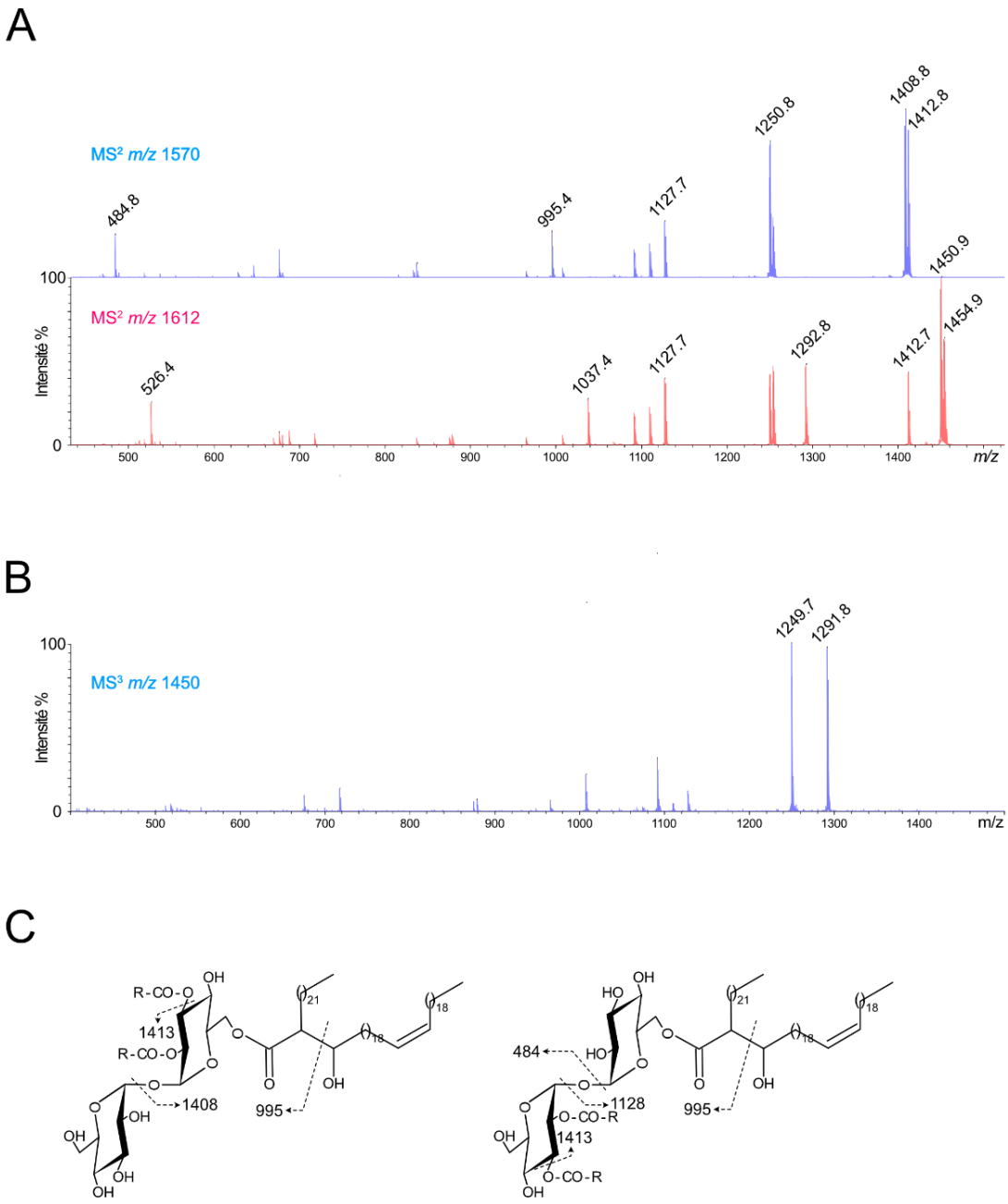


Figure 24 : Fragmentation MALDI-TOF de l'Ac_2TMM. Les spectres MS^2 des ions à *m/z* 1570 et 1612 (A) et MS^3 du parent à *m/z* 1612>1450 (B) confirment le schéma de fragmentation proposé pour l'ion à *m/z* 1570 correspondant aux deux isomères avec un acide mycolique α C_{64} et 2 groupements acyles R à 9 carbones.

III.D.1.e Étude des acides mycoliques de *Mabs*

A partir des culots de souches rugueuse et lisse de *Mabs*, j'ai produit les esters méthyliques d'acides gras afin de les séparer ensuite par CCM. J'ai également utilisé la souche de *Mkan* afin d'obtenir les esters méthyliques d'acides mycoliques (MAMEs) propres à cette espèce notamment les formes oxygénées. La révélation à l'acide phosphomolybdique des profils de *Mabs* R et *Mabs* S ne met

pas en évidence de différences significatives du contenu en esters méthyliques d'acides gras (Figure 25A). D'une part, deux bandes apparaissent avec un R_f de 0,82 et de 0,84 correspondant aux esters méthyliques des acides gras saturés ou insaturés. D'autre part, deux autres bandes avec un R_f de 0,31 et 0,34 sont révélées de façon aussi intense pour *Mabs* R et *Mabs* S. La distance de migration de ces deux dernières bandes s'apparente davantage de celle des MAMEs α ou méthoxy retrouvés chez *Mkan* alors que les MAMEs de type kéto ont un R_f plus faible. Enfin, la préparation des esters méthyliques à partir du TDM précédemment purifié a bien produit les MAMEs retrouvés dans le culot des souches *Mabs* R et *Mabs* S (Fig. 25A).

III.D.1.e.i Caractérisation des MAMEs purifiés par RMN

L'ensemble des MAMEs obtenus par transestérification produit un signal très intense dont le déplacement chimique ^{1}H à 3,78 ppm est attribué au groupement carboxyméthyl terminal (Figure 25B).

Mkan

Les MAMEs de type kéto produisent en RMN ^{1}H deux signaux caractéristiques : le premier est un doublet à 1,04 ppm qui correspond au groupement méthyl branché sur le carbone en α de la fonction cétone ; le second est un quadruplet à 2,5 ppm attribué au proton porté par le carbone en α de la fonction cétone (Figure 25B, spectre bleu).

Il en est de même pour les MAMEs de type méthoxy qui produisent un signal intense à 3,34 ppm correspondant aux protons du deuxième groupement méthoxy ainsi qu'un signal à 0,85 ppm, très proche de celui du groupement ω-méthyl, qui est attribué au groupement méthyl branché sur le carbone en α de la fonction méthoxy (Figure 25B, spectre rouge).

Concernant le dernier type α de MAMEs obtenu à partir de la souche de *Mkan*, ils ne possèdent pas les signaux caractéristiques décrits précédemment (Figure 25B, spectre vert). En outre, aucun signal n'est détecté entre 5 et 5,4 ppm ce qui suppose l'absence de fonction éthylénique. Par contre, trois signaux à –0,33 ppm, 0,55 ppm et 0,64 ppm ont été intégrés afin de confirmer la proportion de protons 1/1/2 associée à la présence de la fonction méthylène et des deux groupes méthines d'un cyclopropyle. De plus, un doublet peu intense à 0,99 ppm est attribué à une fonction méthyle reliée à un groupement CH, supposant qu'une partie des cylopropyles est en configuration *trans* alors que la majorité est en configuration *cis*.

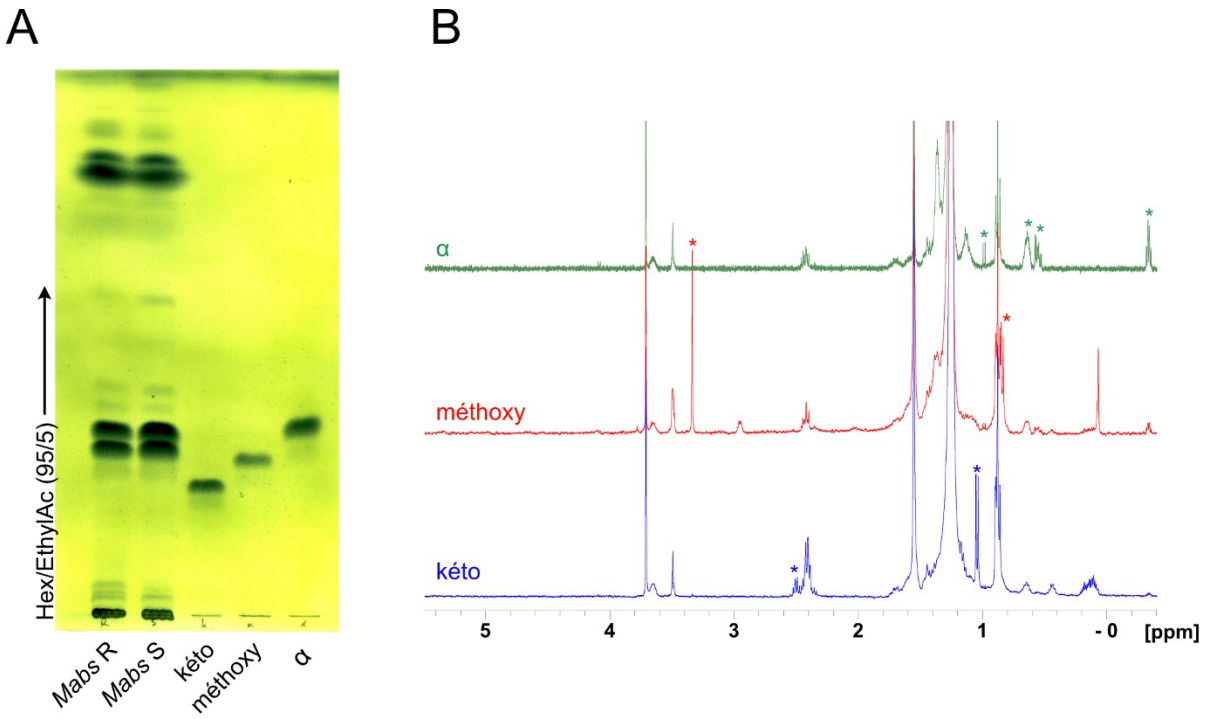


Figure 25 : Contenu en acides mycoliques de *Mabs*. Comparaison CCM (A) des MAMEs des deux morphotypes de *Mabs* avec ceux purifiés chez *Mkan*. Les migrations correspondent à des formes kéto méthoxy et α. Les spectres RMN ^{1}H des 3 MAMEs kéto, méthoxy et α (B), purifiés chez *Mkan*, illustrent les signaux caractéristiques de chacune des formes, marqués par une astérisque.

Mabs

La comparaison des spectres ^{1}H des MAMEs purifiés à partir des esters méthyliques issus des souches rugueuse et lisse ne révèle pas de différences significatives dans les déplacements chimiques des signaux détectés (Figure 26). C'est pourquoi, j'ai réalisé l'attribution des signaux à partir des MAMEs de la forme S qui étaient les plus abondants (Figure 27).De la zone la plus déblindée vers la plus blindée, j'ai identifié :

- les signaux très intenses des groupements éthyléniques cis et trans respectivement à 5,34 et 5,30 ppm sous forme de triplet puisqu'ils sont couplés aux groupements méthylènes à $^{1}H/^{13}C$ 2,01/27,3 et 1,95/32,6 ppm sur le spectre $^{1}H/^{1}H$ COSY ;
- le signal du groupement éthylènique trans à 5,22 ppm qui est couplé à un groupement méthine à $^{1}H/^{13}C$ 2,02/36,5 ppm lui-même couplé à un groupement méthyle $^{1}H/^{13}C$ à 0,93/21,1 ppm ;
- les signaux couplés entre eux sur le spectre $^{1}H/^{1}H$ COSY des positions β-hydroxy et α-alkyl à $^{1}H/^{13}C$ 3,65/72,4 et 2,43/51,4 ppm ;

- les signaux d'intensité plus faible à – 0,34 ppm, 0,55 ppm et 0,64 ppm ainsi que le doublet détecté à 1H 0,98 ppm démontrent la possibilité qu'une petite proportion de fonctions cyclopropyle en configuration trans soit présente.

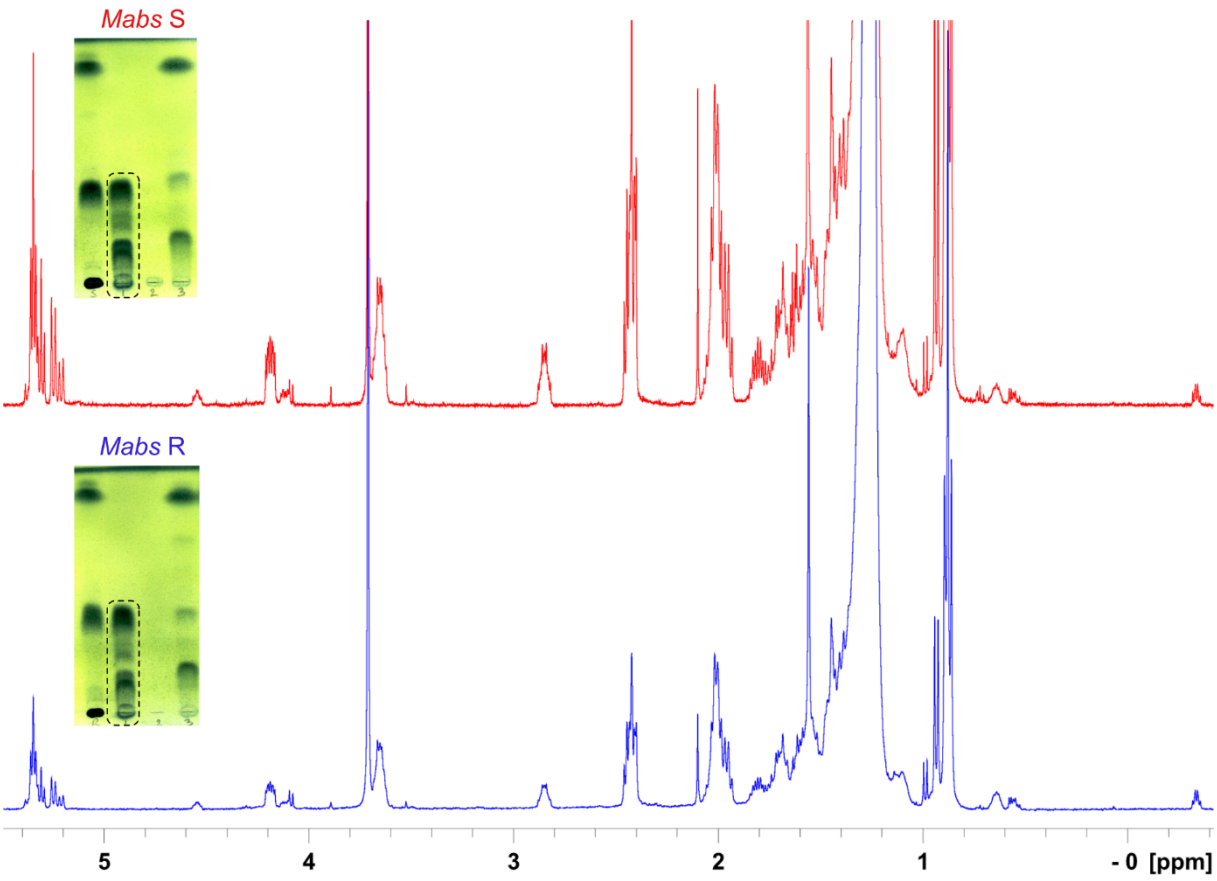


Figure 26 : Comparaison des acides mycoliques de *Mabs* R et S. Spectres RMN ^{1}H comparés des MAMEs purifiés à partir de la souche rugueuse (bleu) et lisse (rouge) de *Mabs*, mesurés à 293K dans le $CDCl_3$ et calibré sur ce dernier. Pour chaque spectre la fraction purifiée en CCM préparative est encadrée (Hexane/éthylacétate ; 95:5 ; x3). Les signaux caractéristiques des formes méthoxy et kéto ne sont pas présents, ceux de la fonction cyclopropyle à droite sont moins intenses que ceux des fonctions éthyléniques à gauche. L'ensemble du spectre est annoté sur la figure 27 en accord avec les expériences $^{1}H/^{1}H$ COSY et HSQC.

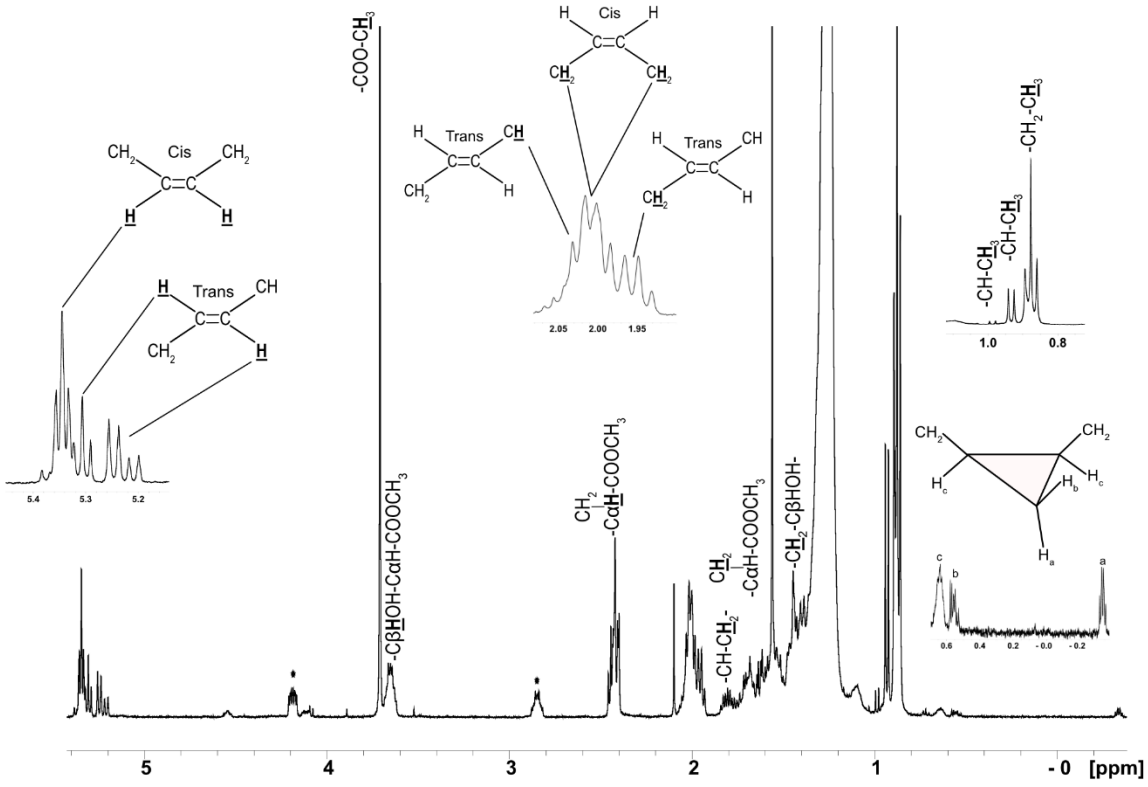


Figure 27 : Caractérisation RMN des acides mycoliques de *Mabs*. Attribution des signaux détectés sur le spectre ^{1}H des MAMEs purifiés à partir de la souche lisse de *Mabs*. Les acides mycoliques de formes α et α' ont majoritairement des configurations cis ou trans éthyléniques et une faible proportion de fonction cyclopropyle. Les positions α-alkyle et β-hydroxy corrèlent bien entre elles sur le spectre $^{1}H/^{1}H$ COSY ainsi qu'avec les CH_2 vicinaux respectifs. Parmi les groupements méthyles, le plus intense correspond à celui des chaines lipidiques et les 2 autres aux fonctions trans éthyléniques ou trans cyclopropyles. Les signaux marqués d'un astérisque corrèlent entre eux mais n'ont pas pu être attribués.

Les résultats obtenus montrent que les acides mycoliques des deux morphotypes sont identiques. De plus, la comparaison avec les acides mycoliques de *Mkan* démontrent que seules des formes non oxygénées α et α' contenant principalement des fonctions éthyléniques sont présentes.

Tableau 10 : Comparaison des acides mycoliques de *Mabs* avec d'autres espèces.

ESPECES	Mtb	Mbov	Mav	Mkan	Msmeg	Mabs
α'					X	X
α	X	X	X	X	X	X
époxy					X	
kéto	X	X	X	X		
méthoxy	X	X		X		
wax ester			X			

Ces résultats sont en accord avec les acides mycoliques supposés dans le TDM et le dérivé di-*O*-acylé du TMM. Ceci a été confirmé pour le TDM grâce à la CCM (Figure 28A) Néanmoins, les analyses par spectrométrie de masse ont montré des ions *m/z* 996 intenses, et décalés de 16 Da avec la masse présumée d'un adduit sodium de l'ester d'acide mycolique α C_{64} à *m/z* 980 (Figure 28B). La présence d'époxyde ou un adduit potassium expliquerait ce résultat, mais il n'a pas été confirmé.

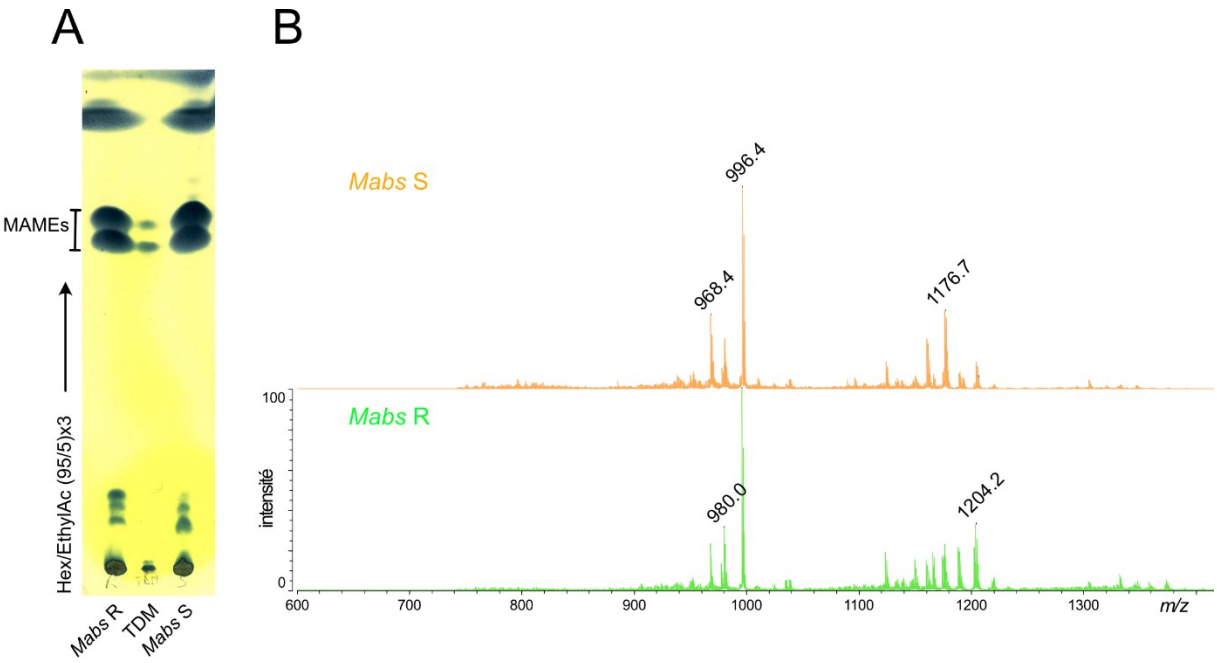


Figure 28 : Acides mycoliques totaux et issus du TDM. Comparaison CCM des MAMEs libérés par méthanolyse du TDM avec ceux obtenus à partir des culots lyophilisés des 2 formes de *Mabs* (A). Spectres MALDI-TOF des MAMEs obtenus à partir des culots lyophilisés des 2 formes de *Mabs* (B). L'ion le plus intense *m/z* 996 correspondrait à un adduit K+ de l'acide mycolique α C_{64}.

III.D.1.f Conclusion

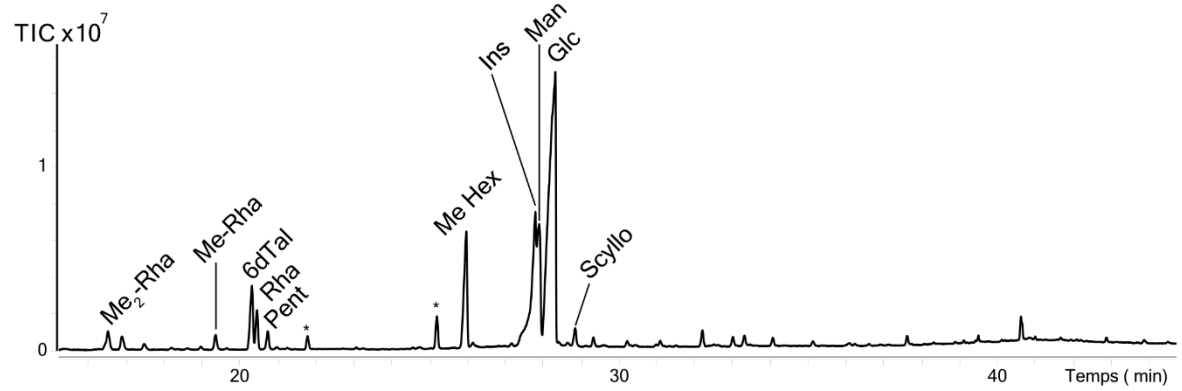
L'analyse de la fraction apolaire a mis en évidence une famille de glycolipides appartenant au TDM déjà identifié dans de nombreuses espèces de mycobactéries. Celui de *Mabs* contient essentiellement des acides mycoliques de forme α et α' cis et trans éthyléniques, avec une faible proportion de fonction cyclopropyle. La forme majoritaire est issue de l'acylation de deux α mycolates C_{64} sur les positions 6 du tréhalose. L'hétérogénéité constatée provient du remplacement simple ou double par un α' mycolate C_{79}. Un mélange de deux autres composés provenant du TMM a pu être isolé montrant des acylations en positions 2 et 3 par des acides gras courts et polyméthylbranchés, mais qui n'ont pas été formellement identifiés.

III.E ETUDE DE LA FRACTION POLAIRE

III.E.1 DETERMINATION DE LA COMPOSITION EN MONOSACCHARIDES

Alors que la fraction apolaire ne présentait que du glucose, de nombreux monosaccharides ont été détectés dans les fraction polaires des morphotypes rugueux et lisse (Figure 29A & B). Le glucose (Tr = 28,3 min) est le monosaccharide majoritaire, puis le mannose (Tr = 27,7 min) et un hexose *O*-méthylé (Tr = 26 min) sont présents en quantité équivalente. Enfin, le *myo*-inositol (Tr = 27,7 min) se distingue juste avant le mannose ainsi que les traces d'un pentose (Tr = 20,7 min). Par contre, la détection possible de rhamnose en proportion infime dans la souche rugueuse par rapport à la souche lisse confirme questionne la production éventuelle de GPLs. (Figure 29B).

A



B

Monosaccharides		Me_2Rha	MeRha	6dTal	Rha	Pent	MeHex	Ins	Man	Glc
Temps de rétention (min)		16,5	19,3	20,3	20,4	20,7	26	27,7	27,8	28,3
Proportion	***Mabs* R**	ND	ND	ND	0,29%	1,63%	12,81%	17,43%	25,17%	42,96%
	***Mabs* S**	0,84%	0,51%	3,40%	1,70%	0,70%	8,11%	11,04%	16,65%	57,05%

Figure 29 : Composition en monosaccharides de la fraction polaire de *Mabs* S et R. Les dérivés itols acétates sont détectés en GC-MS après dérivation de la fraction polaire de *Mabs* S (A) et *Mabs* R. La différence principale réside dans l'absence de déoxyhexoses (B) excepté le rhamnose qu'il est possible de détecter en très petite proportion dans la souche *Mabs* R. Le *scyllo*-inositol est utilisé comme standard interne. Les pics marqués par un astérisque sont des dérivés lipidiques.

III.E.2 PURIFICATION DES GLYCOLIPIDES DE LA FRACTION POLAIRE

L'utilisation de plusieurs gradients de concentration en méthanol dans le chloroforme a permis la purification des GPLs autour de 5%, puis des DATs de 5 à 10% de méthanol, et enfin des PIMs au-

delà de 30% de méthanol (Figure 30A-C). Des purifications par CCM préparative et sur colonne de silice greffée C_{18} ont été utilisées pour améliorer la pureté.

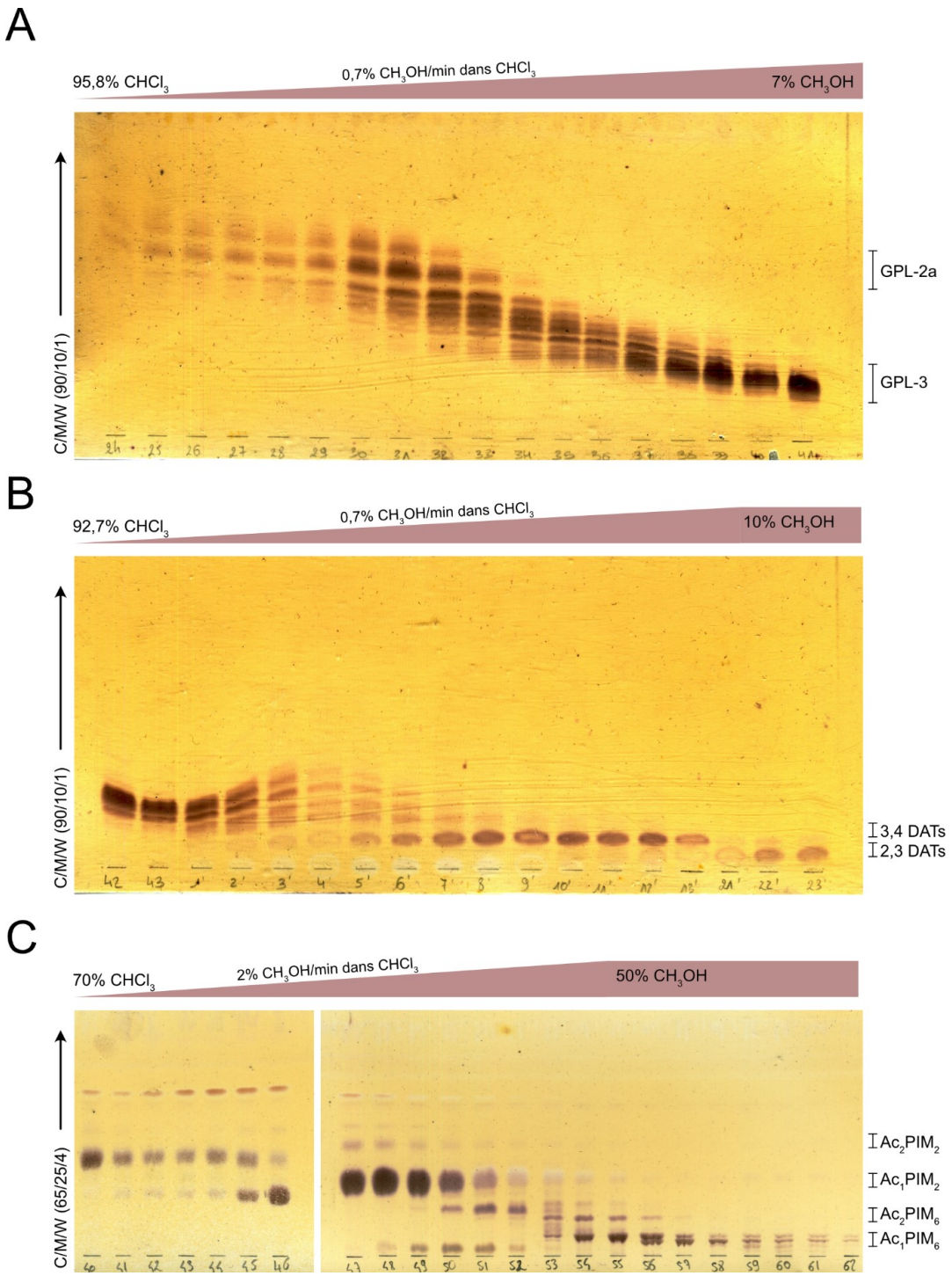


Figure 30 : Purification de la fraction polaire de *Mabs*. Séparation réalisée par chromatographie d'adsorption flash sur colonne de gel de silice à partir d'une fraction polaire de *Mabs* S (A & B) et de *Mabs* R (C) montrant les familles principales de glycolipides qui ont été étudiées : glycopeptidolipides (GPLs), di-*O*-acyl tréhalose (DATs) et phosphatidylinositol mannosides (PIMs). Les gradients de concentration en méthanol dans le chloroforme sont affichés au-dessus des contrôles CCM.

III.E.3 CARACTERISATION DES DI-*O*-ACYL TREHALOSE

En fonction de l'utilisation d'un milieu de culture 7H10 OADC ou 7H10 avec 0,2 % de glucose et de glycérol, le contenu en glycolipides de la forme rugueuse est différent. En effet, dans le cas du milieu 7H10 OADC, deux bandes de coloration violette après révélation à l'orcinol-sulfurique, apparaissent beaucoup plus intenses (Figure 31A). Après purification de ces deux glycolipides (Figure 31B), l'analyse GC-MS de la composition ne révèle que le glucose comme monosaccharide.

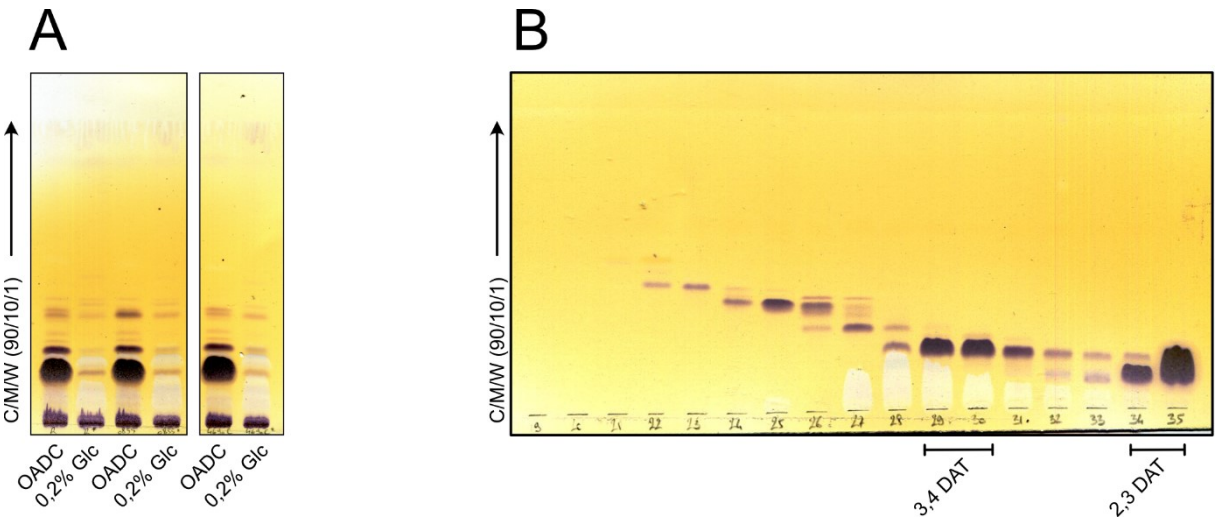


Figure 31 : Production des DATs selon le milieu de culture. Comparaison de l'expression des DATs contenus dans la fraction polaire de 3 souches de *Mabs* R cultivées sur milieu 7H10 OADC ou 7H10 + 0,2 % de glucose et de glycérol (A). Repérage des DATs préalablement à la caractérisation du tréhalose perméthylé.

De plus, suite à la perméthylation, le tréhalose a été identifié comme le disaccharide majoritaire à Tr = 35,15 min (Figure 32A). Concernant le DAT de R_f le plus élevé, il est accompagné d'esters méthyliques d'acides gras C_{16} et C_{18} saturé ou monoinsaturé (Figure 32A). Dans le cas du DAT ayant le plus petit R_f, deux pics chromatographiques intenses éluent après le tréhalose perméthylé à Tr = 46,36/47,3 min (Figure 32B). Leur spectre de fragmentation après impact électronique produit de nombreux ions communs avec celui du tréhalose, suggérant qu'il s'agit également de deux dérivés du tréhalose qui n'ont pas complétement subi la réaction de perméthylation.

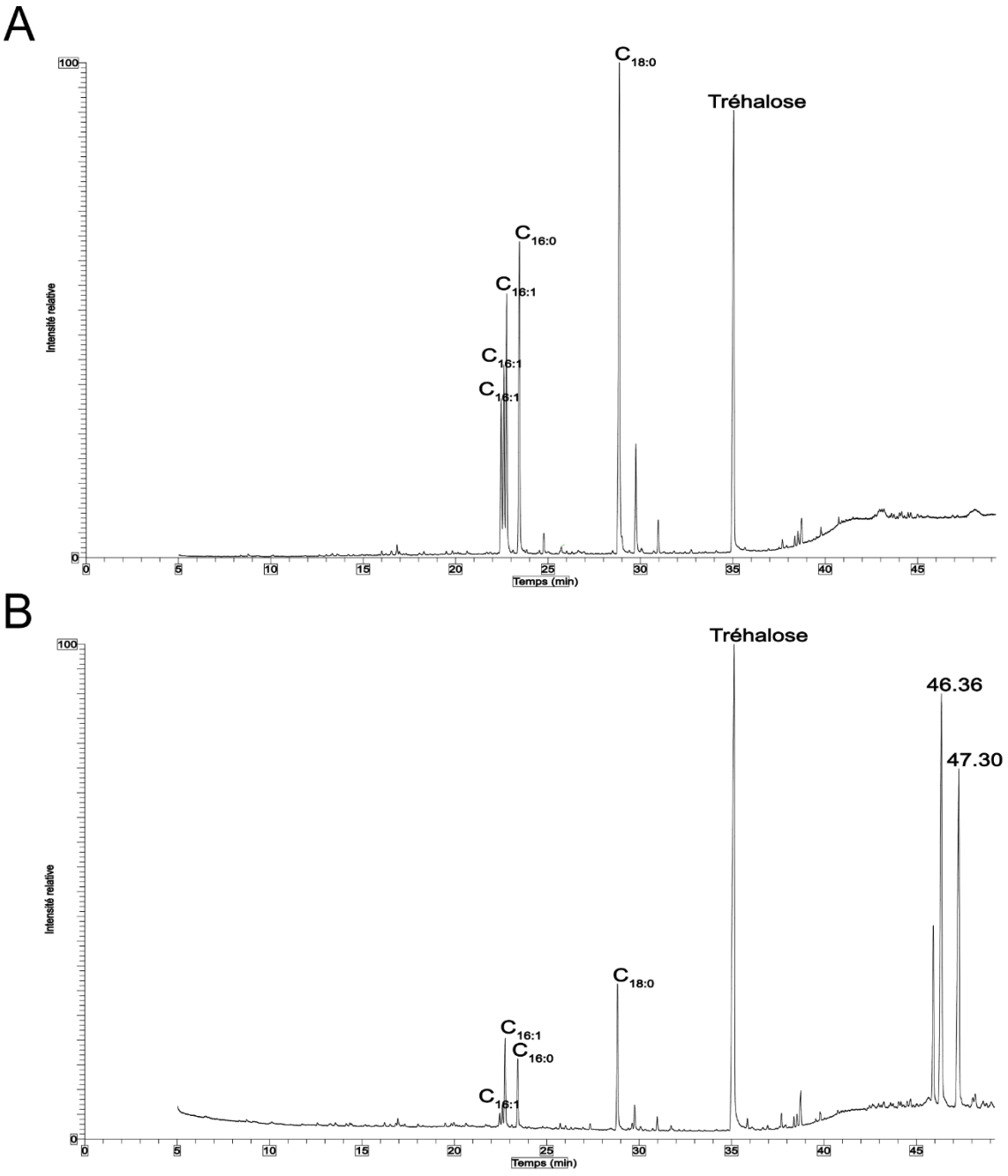


Figure 32 : Caractérisation du tréhalose dans les DATs. Identification du tréhalose sous forme perméthylé dans les 3,4 DAT (A) et 2,3 DAT (B) purifiés. Le traitement a produit un pic chromatographique très intense dont le Tr = 35,15 min et le spectre de fragmentation est identique à un standard de tréhalose. Le 3,4 DAT présente également des FAMEs de palmitoléate, palmitate ou stéarate. Le 2,3 DAT présente au moins 2 pics chromatographiques non identifiés à Tr = 46,36 et 47,30 min.

III.E.3.a Analyse du 3,4 DAT

III.E.3.a.i Spectres RMN du 3,4 DAT natif

Les déplacements chimiques du dérivé de tréhalose dont le R_f est le plus élevé sont présentés dans le tableau 11. En effet, deux signaux anomériques I-1 et I'-1sont visibles sur le spectre $^{1}H/^{13}C$ HSQC à 5,2/93,5 ppm et 5,14/93,9 ppm démontrant la présence de deux monosaccharides (Figure 33A). I-1 est corrélé à I-2 à 3,77/70,3 ppm puis avec I-3 et I-4 à 5,41/72,8 ppm et 4,95/69,3 ppm dont le fort déblindage est lié à leur acylation (Figure 33B). I-5 à 4,05/70,3 ppm est facilement repérée *via* la corrélation avec la position I-4. Il en est de même pour I-6 et 6' à 3,63 et 3,53/61,2 ppm qui sont corrélés entre eux ainsi qu'avec I-5 (Figure 33B). A partir de I'-1 à 5,14/93,9 ppm il est de même possible d'observer l'ensemble du système de spin (Tableau 11). La constante de couplage J^3 entre les protons I-1 et I-2 est de 3,7 Hz alors que les trois suivantes s'élèvent au-dessus de 9 Hz confirmant la nature α-glucosyl de ce résidu ainsi que du précédent.

Dans la zone plus blindée du spectre HSQC, on retrouve 4 signaux correspondants aux groupements méthylènes en α et en β du groupement carbonyle ainsi que de la fonction éthylénique. Son signal s'observe à 5,29/130,8 ppm. Ces signaux sont corrélés 2 à 2 sur le spectre $^{1}H/^{1}H$ COSY et finalement avec la chaine lipidique à 1,23/30,5 ppm qui est corrélé avec un groupement méthyl terminal, également visible à 0,83/14,7 ppm. Ces signaux témoignent de chaines lipidiques linéaires saturée ou monoinsaturée.

Les spectres présentent également 4 signaux dont 3 sont corrélés entre eux : un groupement méthine à 5,24/70,5 ppm et 2 protons appartenant à un groupement méthylène à 4,18 et 4,4/63 ppm attribuable à un résidu de glycérol appartenant à un lipide contaminant (Figure 33A & B).

Tableau 11 : Signaux RMN du tréhalose du 3,4 DAT. Déplacements chimiques et constantes de couplage attribuées aux 6 positions des 2 unités de glucose I et I' dans le tréhalose du 3,4 DAT.

		Glc -1	*Glc -2*	*Glc -3*	*Glc -4*	*Glc -5*	*Glc -6/6'*
I	**δ ^{1}H (ppm)**	5,2	3,77	5,41	4,95	4,05	3,63/3,53
	Multiplicité	D	M	T (Dd)	T (Dd)	M	Dd
	J^3 (Hz)	3,8		9,6/9,6	9,8/9,8		12,2
	δ ^{13}C (ppm)	93,5	70,3	72,8	69,3	70,3	61,2
I'	**δ ^{1}H (ppm)**	5,14	3,55	3,85	3,34	3,77	3,84/3,69
	Multiplicité	D	Dd	T (Dd)	T (Dd)	M	Dd
	J^3 (Hz)	3,7	3,7/9,9	9,3/9,3	9,4/9,4		11,8/5,5
	δ ^{13}C (ppm)	93,9	71,2	73,3	70,9	72,7	62

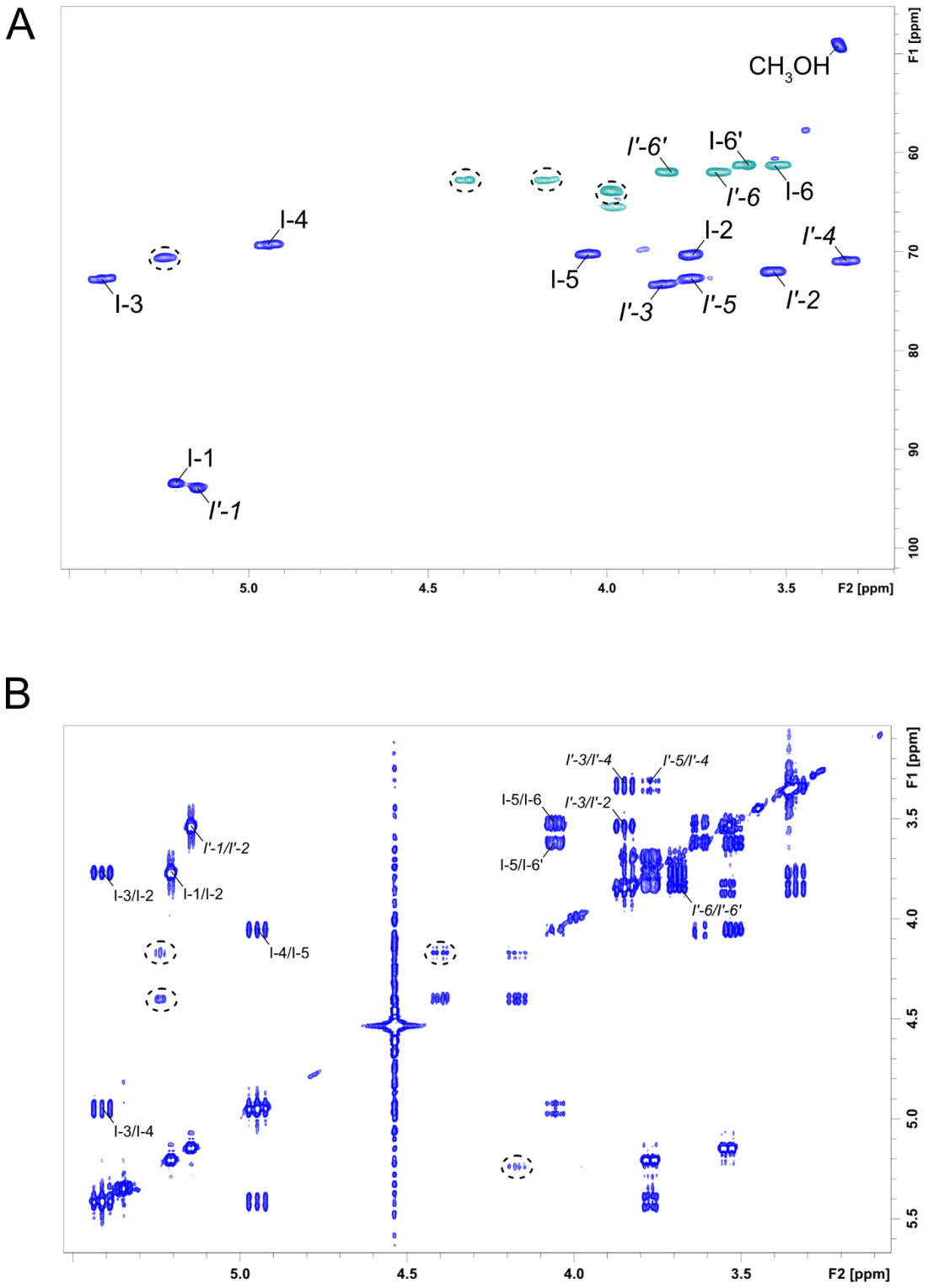


Figure 33 : Analyse RMN du 3,4 DAT. Spectres $^{1}H/^{13}C$ HSQC (A) et $^{1}H/^{1}H$ COSY (B) du 3,4 DAT purifié à partir de *Mabs* R, et mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Le spectre est calibré sur le méthanol, et les signaux en vert sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2. Les signaux entourés correspondent à un résidu de glycérol provenant d'un probable contaminant lipidique. Le fort déblindage du signal ^{1}H indique que I-3 et I-4 sont acylés.

III.E.3.a.ii Spectres MALDI-TOF du 3,4 DAT natif

Le spectre de masse du 3,4 DAT produit deux ions majoritaires de rapport *m/z* 841 et 869 et un ion minoritaire *m/z* 897 (Figure 34A, spectre rouge). L'intensité relative des ions *m/z* 839 et 867 implique qu'une insaturation peut être présente dans la chaine lipidique. La fragmentation du parent *m/z* 841 produit un ion fils majoritaire *m/z* 679 correspondant à la perte d'un hexose, ainsi qu'un second fragment *m/z* 585 correspondant au départ d'un résidu palmitoyl (Figure 34A, spectre bleu). Une nouvelle sélection de l'ion *m/z* 679 aboutit à l'apparition d'un ion fragment *m/z* 423 qui confirme la perte du résidu palmitoyl (Figure 34A, spectre vert). Il en est de même pour les ions parents 869 et 897 dont la fragmentation permet d'observer la perte d'un hexose ou d'un résidu d'acide gras C_{16} ou C_{18}.

III.E.3.a.iii Composition lipidique du 3,4 DAT natif

Finalement, l'analyse des acide gras par GC-MS montre principalement les formes saturées et monoinsaturées C_{16} et C_{18} (Figure 34B). Les acides palmitique (Tr = 24,3 min) et stéarique (Tr = 26,2 min) sont relativement minoritaires, 3 isomères de l'acide palmitoléique (Tr = 24-24,1 min) ont pu être séparés et l'acide oléique est le plus intense (Tr = 26 min). L'utilisation d'une hydrolyse alcaline suivie d'une transestérification a donné un résultat similaire. La distribution des lipides parmi les adduits sodium observés précédemment sur le spectre MALDI-TOF est donnée par le tableau 12.

Tableau 12 : Distribution lipidique du 3,4 DAT. Répartition des lipides dans les adduits sodium détectés par spectrométrie de masse MALDI-TOF en mode positif. Seul l'ion *m/z* 867 peut correspondre à deux combinaisons différentes.

Adduit	**Masse (u.m.a.)**				
$[M+Na]^+$	**Tréhalose**	**$C_{16:0}$**	**$C_{16:1}$**	**$C_{18:0}$**	**$C_{18:1}$**
839	342	239	237		
841	342	239x2			
867	342		237	267	
	342	239			265
869	342	239		267	
895	342			267	265
897	342			267x2	

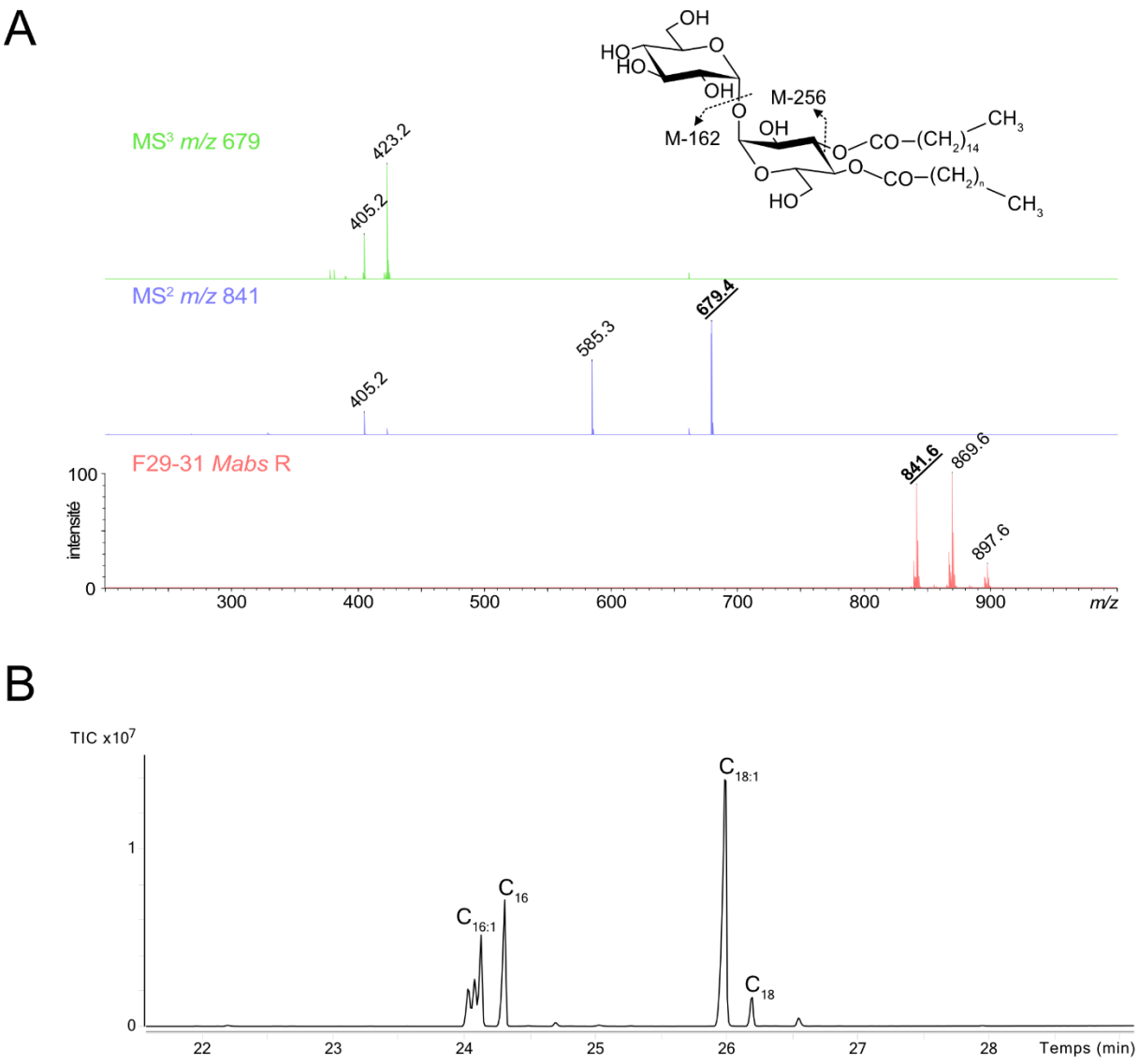


Figure 34 : Substitution lipidique du 3,4 DAT. Spectres MALDI-MS et MS^n du 3,4 DAT (A) et chromatogramme des esters méthyliques d'acides gras (FAMEs) générés par méthanolyse acide (B) de l'acyl tréhalose provenant de la souche clinique GS163 (B). Le spectre MS^2 du parent à *m/z* 841 produit des ions M-162 qui confirme la perte d'une hexose non acylé, et M-256 suggèrant la présence de palmitate comme indiqué sur le schéma de fragmentation.

III.E.3.b Analyse du 2,3 DAT

III.E.3.b.i Spectres RMN du 2,3 DAT natif

L'analyse des spectres ^{1}H/^{13}C HSQC et ^{1}H/^{1}H COSY du DAT avec le R_f le moins élevé indique la présence de deux monosaccharides de type α-glucosyl (Figure 35A, tableau 13). A partir du signal I-1 à 5,33/91,4 ppm, on observe sur le spectre ^{1}H/^{1}H COSY des corrélations avec I-2 à 4,88/71,1 ppm, puis I-3 à 5,48/72,2 ppm (Figure 35B). Cela implique que ces positions soient acylées. D'autre part, les

signaux des protons portés par I-6/6' des 2 monosaccharides demeurent superposés sur le spectre HSQC, rendant difficile l'attribution de chacun des 4 protons (Figure 35A). L'autre système de spin appartenant au résidu de glucose non acylé démarre de la position *I'*-1 à 5,14/94,1 ppm pour terminer sur la position *I'*-6 dans la même zone à $^{1}H/^{13}C$ 3,73/61,4 ppm.

Tableau 13 : Signaux RMN du tréhalose du 2,3 DAT. Déplacements chimiques et constantes de couplage attribuées aux 6 positions des 2 unités de glucose I et I' dans le tréhalose du 2,3 DAT.

		Glc -1	*Glc -2*	*Glc -3*	*Glc -4*	*Glc -5*	*Glc -6/6'*
I	δ ^{1}H (ppm)	5,33	4,88	5,48	3,56	3,95	3,85/3,73
	Multiplicité	D	Dd	T (Dd)	T (Dd)	M	Dd
	J^3 (Hz)	3,4	3,5/10,3	9,7/9,7	9,7/9,7		10,3/3,5
	δ ^{13}C (ppm)	91,4	71,1	72,2	69,5	72,6	61,5
I'	δ ^{1}H (ppm)	5,14	3,55	3,81	3,47	3,58	3,73
	Multiplicité	D	Dd	T (Dd)	T (Dd)	M	Dd
	J^3 (Hz)	3,7	3,7/9,9	9,9/9,9	9,9/9,9		12,4/4
	δ ^{13}C (ppm)	94,1	72	73,4	70,6	72,7	61,4

D'après le spectre $^{1}H/^{13}C$ HMBC, la corrélation réciproque des deux positions anomériques démontrent que les unités α-glucosyl sont bien reliées entre elles via les positions 1 (Figure 36A). D'autre part, les protons I-2 et I-3 sont bien corrélés avec les carbones quaternaires de deux groupements carbonyles dont le déplacement chimique ^{13}C est respectivement à 177,5 ppm et 177,6 ppm. Ces mêmes fonctions sont également corrélées avec deux fois trois signaux : deux groupements méthines à 2,48/37,4 ppm et 2,54/37,8 ppm, deux groupements méthylènes à 1,59/41,4 ppm et 1,68/41,3 ppm puis deux groupements méthyles à 1,09/17,4 ppm et 1,14/18 ppm (Figure 36B).

Sur le spectre ^{1}H, les groupements méthines ont l'apparence de deux sextuplets superposés avec une constante de couplage apparente J^3 d'environ 7 Hz et les groupements méthyles produisent deux doublets avec la même valeur de constante de couplage. Par ailleurs, ces signaux sont corrélés entre eux de façon intense sur le spectre $^{1}H/^{1}H$ COSY ainsi qu'avec le groupement méthylène. Enfin, trois signaux correspondant à des groupements méthines de ^{1}H 1,43 à 1,57 ppm ainsi que deux groupements méthyles à 1H 0,83 et 0,85 ppm apparaissent aussi sur le spectre $^{1}H/^{13}C$ HSQC (Figure 36A).

L'ensemble de ces observations montre que les deux acides gras portés par les positions 2 et 3 ont des chaines saturées polyméthylbranchées contrairement au 3,4 DAT.

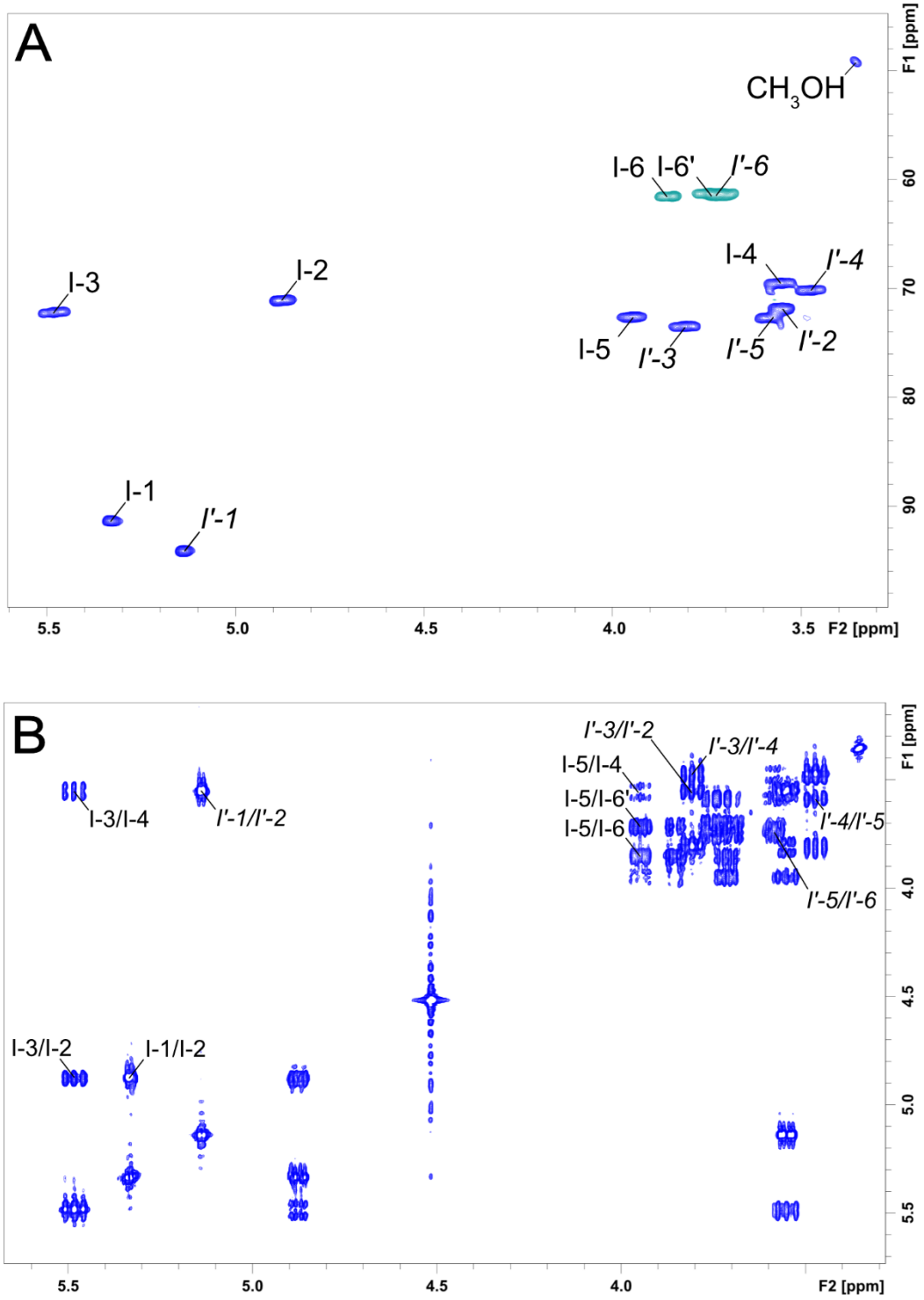


Figure 35 : Analyse RMN du 2,3 DAT. Spectres $^{1}H/^{13}C$ HSQC (A) et $^{1}H/^{1}H$ COSY (B) du 2,3 DAT purifié à partir de *Mabs* R, et mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Le spectre est calibré sur le méthanol, et les signaux en vert sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2. Le fort déblindage du signal ^{1}H indique que les positions I-2 et I-3 sont acylés.

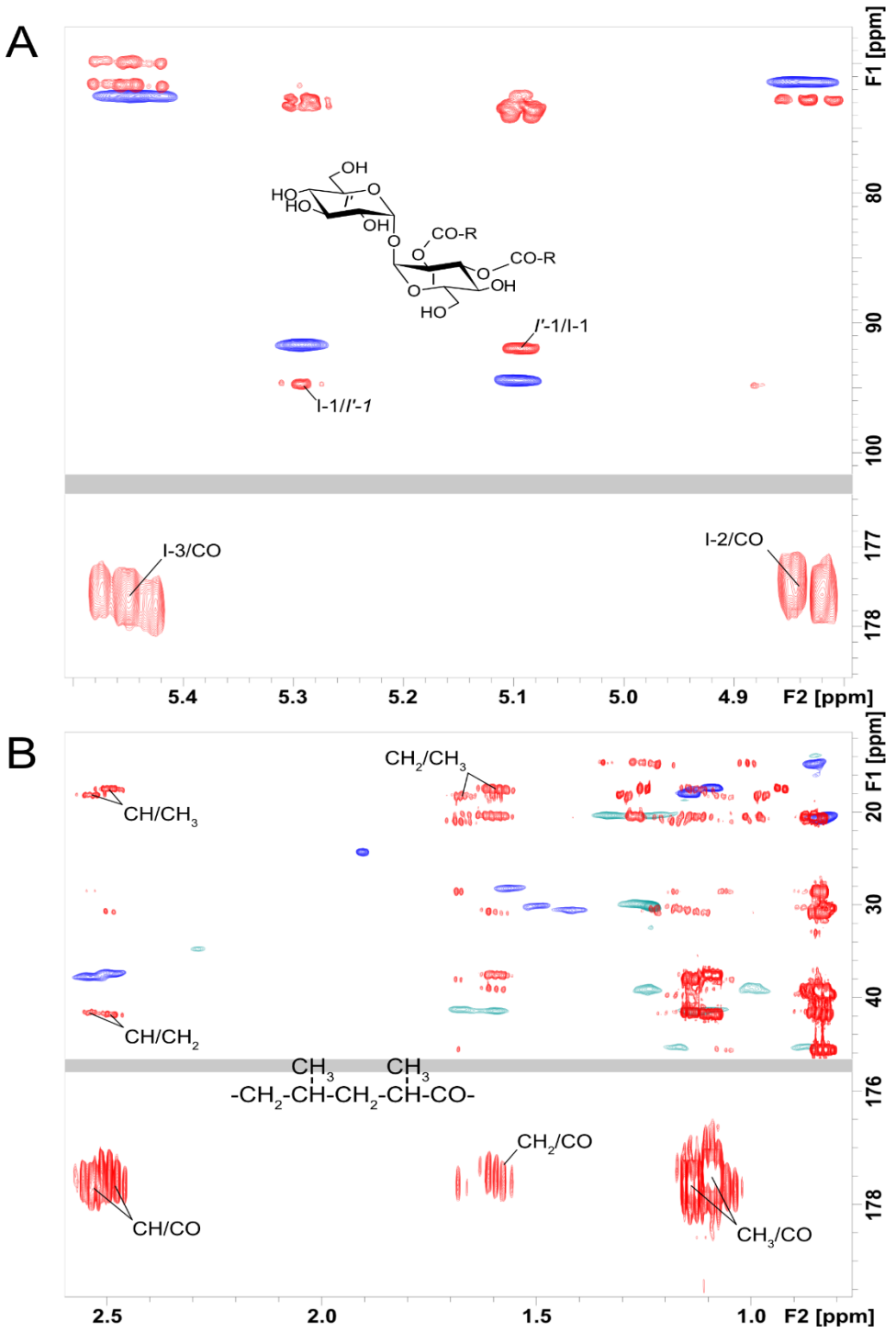


Figure 36 : Acylation du 2,3 DAT. Superposition des spectres $^{1}H/^{13}C$ HSQC (bleu) et HMBC (rouge) du 2,3 DAT purifié à partir de *Mabs* R, et mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Le spectre est calibré sur le méthanol, et les signaux en vert sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2. Le 1er zoom (A) montre la corrélation des positions anomères entre elles, et des positions I-2 et I-3 avec les carbonyles. Le 2ème zoom montre la corrélation des méthines, methylènes et méthyles avec ces mêmes carbonyles liés à des chaines lipidiques polyméthylbranchées. La correspondance avec la structure de la molécule est au centre du spectre.

III.E.3.b.ii Spectres MALDI-TOF du 2,3 DAT natif

L'analyse par spectrométrie de masse MALDI-TOF du 2,3 DAT s'est révélé infructueuse malgré les tentatives de dilution croissante avec la matrice DHB ; l'utilisation des matrices amino quinoline, HCCA ou THAP ; des essais d'élévation de l'énergie du LASER pour la désorption ; une détection en mode négatif plutôt qu'en mode positif ; et des purifications supplémentaires sur silice greffée C_{18}.

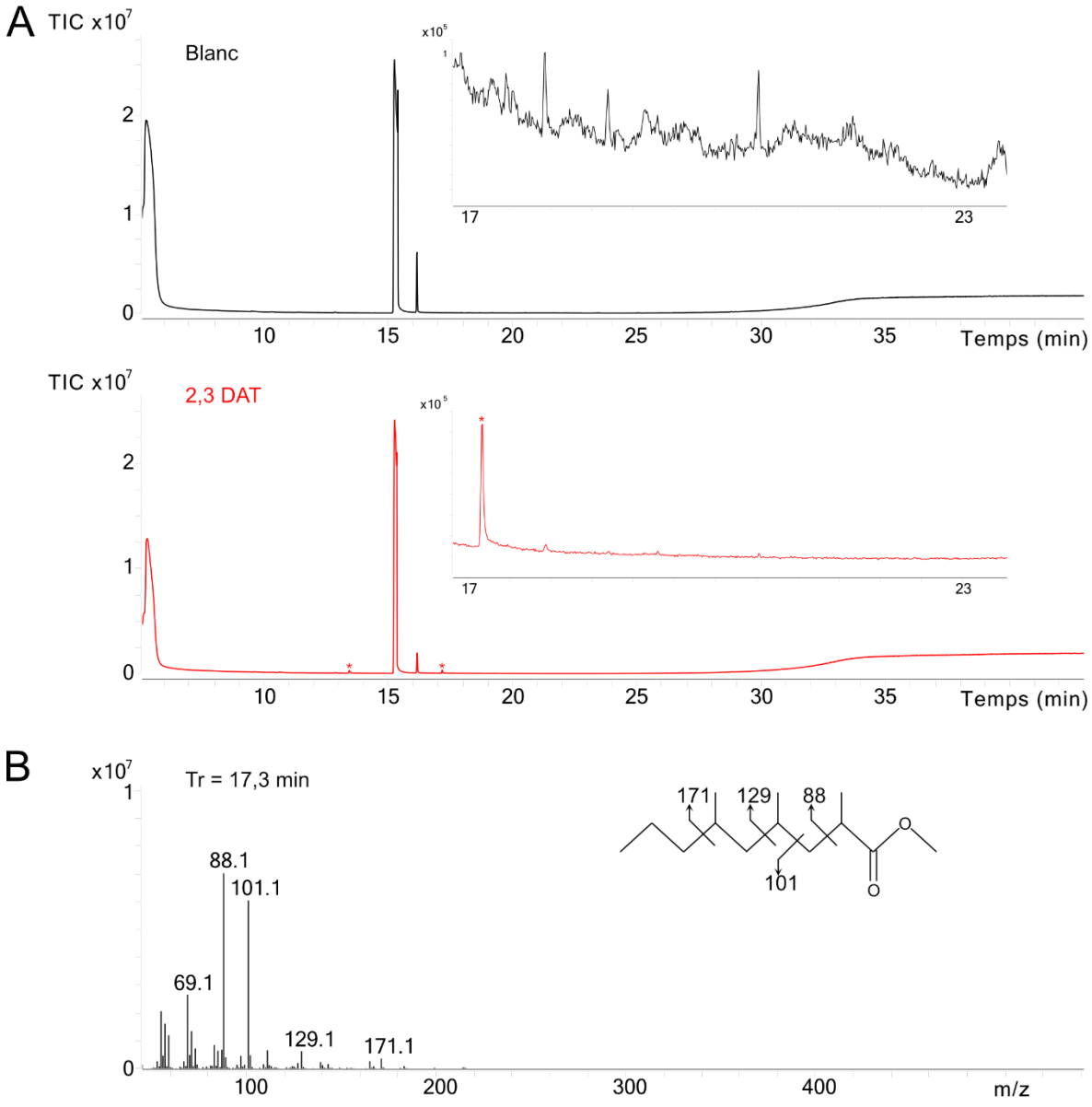


Figure 37 : Composition lipidique du 2,3 DAT par GCMS. Les esters méthyliques d'acides gras (FAMEs) ont été générés par hydrolyse alcaline suivie d'une méthylation (A). Cette dernière méthode a produit des contaminants identifiables dans le blanc. Les pics chromatographiques marqués d'un astérisque sont présents dans le 2,3 DAT. Le spectre de fragmentation de celui à Tr = 17,3 min révèle un acide 2,4,6 tri-*O*-méthyl nonanoïque (B).

III.E.3.b.iii Composition lipidique du 2,3 DAT

L'analyse des lipides sous forme d'esters méthyliques obtenus après hydrolyse alcaline montre la présence de deux pics résiduels à Tr = 13,4 et 17,3 min (Figure 37A) identifiés par GC-MS comme l'acide 2,4 diméthyl heptanoïque et l'acide 2,4,6 triméthyl nonanoïque respectivement (Figure 37B). Ces acides gras identifiés sont en accord avec les résultats de l'analyse RMN puisqu'ils possèdent une chaine polyméthylbranchée. Pour confirmer ce résultat, après deux purifications successives sur colonne de silice vierge puis sur colonne de silice greffée C_{18}, j'ai comparé les esters méthyliques obtenus suite à la méthanolyse acide ou à l'hydrolyse alcaline suivie d'une méthylation sur plusieurs fractions. Le suivi de l'intensité de l'ion *m/z* 88 montre que la méthanolyse acide produit principalement le palmitate et le stéarate de méthyle. Par contre, l'hydrolyse alcaline couplée à la méthylation aboutit bien à la détection d'acides 2,4 diméthyl heptanoïque et 2,4,6 triméthyl nonanoïque.

III.E.3.c Analyse d'un acyl tréhalose provenant de la souche clinique GS163

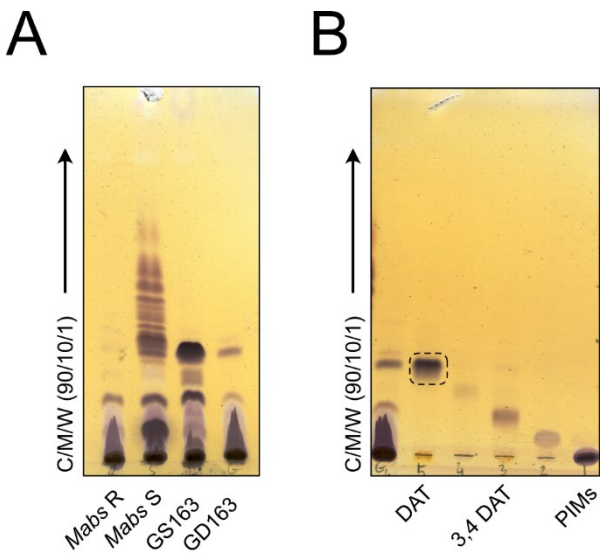


Figure 38 : Profil CCM de la souche clinique GD163. Comparaison des extraits polaires des souches de référence ATCC-19977 rugueuse (*Mabs* R) et lisse (*Mabs* S) au variant lisse (GS) observé après culture de la souche clinique parentale GD163. Le glycolipide avec le plus grand R_f s'accumulant a été purifié pour une analyse RMN puis MALDI. Les autres fractions ont été identifiées d'après les résultats de leurs spectres MALDI.

Durant mon projet de thèse, un 3ème dérivé du tréhalose a été isolé sur une souche clinique de *Mabs* isolée d'un patient atteint de mucoviscidose et qui ne répondait pas aux traitements antibiotiques. A l'origine, la souche GD163 était rugueuse chez le patient. La présence de variants

intermédiaires entre lisse et rugueux a été observée de façon très régulière sur boite de cultures. Nos collègues de l'IRIM ayant exclu une éventuelle contamination, nous souhaitions étudier cette réversion inédite vers un phénotype plus lisse. D'après les profils CCM, le variant GS ne présente ni le profil des souches de référence rugueuse et lisse, ni celle de la souche parentale GD163 (Figure 38A). Un glycolipide s'étant accumulé, il a donc été purifié par CCM préparative (Figure 38B).

III.E.3.c.i Spectres RMN de l'acyl tréhalose natif issu de la souche GS163

D'après le spectre $^{1}H/^{13}C$ HSQC et le spectre $^{1}H/^{1}H$ COSY, les signaux observés sont assez similaires au 3,4 DAT précédemment décrit, puisque deux signaux anomériques I-1 et I'-1 apparaissent à nouveau à 5,16/93,9 ppm et 5,07/94,5 ppm dont les systèmes de spin ont été identifiés (Fig. 34A). De manière remarquable I-3 à 5,35/73 ppm et I-4 5,02/68,7 ppm possèdent un déplacement chimique proton fortement déblindé probablement dû à la présence de groupements acylés. De même le fort déblindage de I-5 et I-6 par rapport au 3,4 DAT suggère une acylation supplémentaire qui n'a pas pu être confirmée par HMBC.

La zone plus blindée montre simplement la présence d'une chaine lipidique possiblement monoinsaturée à l'image du précédent 3,4 DAT décrit. L'individualisation de deux signaux méthyles couplés à des méthines à 0,92/23,1 ppm et 0,81/20,6 ppm témoignent de la possibilité de branchements.

Tableau 14 : Signaux RMN du tréhalose de l'acyl tréhalose isssu de la souche GS163. Déplacements chimiques et constantes de couplage attribuées aux 6 positions des 2 unités de glucose I et I' du tréhalose issu de l'acyl tréhalose purifié à partir du variant lisse GS de la souche clinique GD163. En comparaison avec le 3,4 DAT du tableau 11 les signaux I-4 à I-6 sont davantage déblindés.

		Glc -1	*Glc -2*	*Glc -3*	*Glc -4*	*Glc -5*	*Glc -6/6'*
I	**δ ^{1}H (ppm)**	5,16	3,75	5,35	5,02	4,24	4,14/4,04
	Multiplicité	D	M	T (Dd)	T (Dd)	M	Dd
	J^{3} (Hz)	3,7		9,7/9,7	9,8/9,8		11,2/3,8
	δ ^{13}C (ppm)	93,9	70,4	73	68,7	68	62,2
I'	**δ ^{1}H (ppm)**	5,07	3,48	3,78	3,31	3,74	3,79/3,67
	Multiplicité	D	Dd	T (Dd)	T (Dd)	M	Dd
	J^{3} (Hz)	4	3,5/9,7	9,0/9,0	9,7/9,7		
	δ ^{13}C (ppm)	94,5	72,1	73,5	71	72,9	61,9

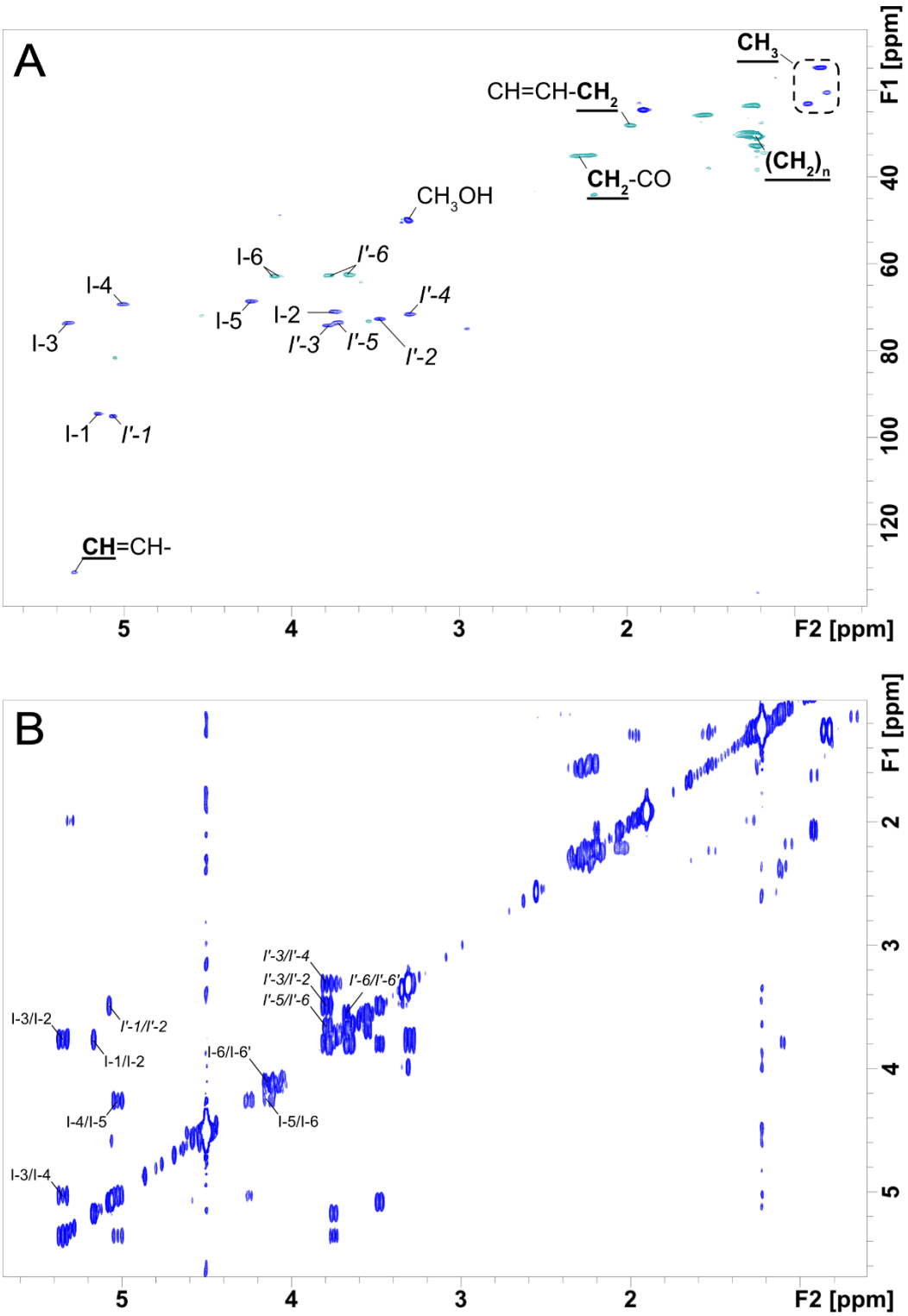


Figure 39 : Analyse RMN de l'acyl tréhalose de la souche clinique GS163. Spectres $^{1}H/^{13}C$ HSQC (A) et $^{1}H/^{1}H$ COSY (B) de l'acyl tréhalose purifié à partir du variant lisse GS163, et mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Le spectre est calibré sur le méthanol, et les signaux en vert sur le spectre $^{1}H/^{13}C$ HSQC correspondent à des CH_2. Les positions I-4 à I-6, sont davantage déblindés que dans le 3,4 DAT suggérant une acylation supplémentaire sur la position 6. Les signaux entourés correspondent à des groupements méthyles dont deux sont probablement des branchements de la chaine lipidique.

III.E.3.c.ii Spectres MALDI-TOF de l'acyl tréhalose natif issud de la souche GS163

Le spectre de masse du DAT de GS163 présente un massif s'étalant des ions *m/z* 911 à 1009 avec 2 ions majoritaires de rapport *m/z* 925 et 953. A nouveau, le spectre de fragmentation de l'ion *m/z* 925 produit essentiellement deux fragments *m/z* 763 et 669, correspondant respectivement à la perte d'un hexose et d'un résidu palmitoyl, qui est confirmé par la fragmentation additionnelle du fragment *m/z* 763 (Figure 40). La perte du résidu hexose s'observe nettement pour la série des ions *m/z* 939, 953, 967 ou 981, ainsi que la perte du résidu palmitoyl pour l'ion *m/z* 953 et 981.

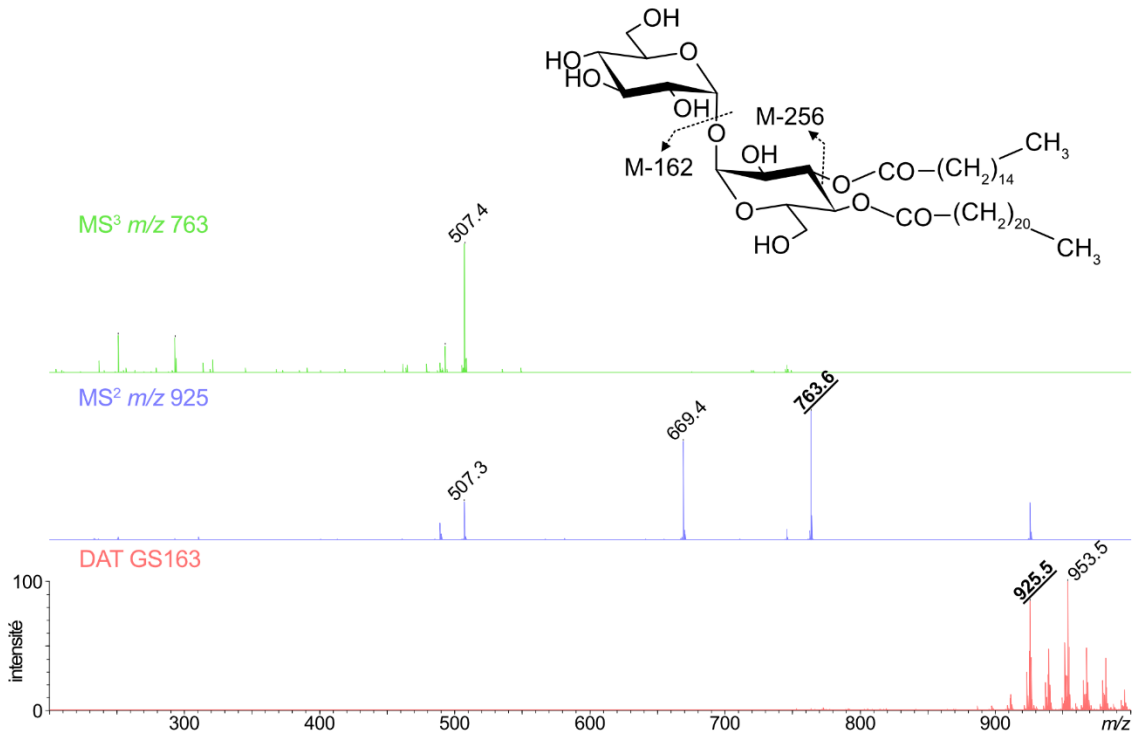


Figure 40 : Analyse MALDI-TOF de l'acyl tréhalose issu de la souche clinique GS163. Le spectre MS^2 du parent à *m/z* 925 produit des ions M-162 qui confirme la perte d'une hexose non acylé, et M-256 suggèrant la présence de palmitate, confirmé par le spectre MS^3 de l'ion à *m/z* 763.

Malgré l'acylation supplémentaire suggérée par l'analyse RMN, les résultats de fragmentation en spectrométrie de masse plaident en faveur d'un 3,4 di-*O*-acyl tréhalose dont l'une des chaines lipidiques serait plus longue. Compte tenu de la présence d'ions *m/z* 1095 ou 1391 de masse supérieure et d'intensité 15 fois inférieure, la fragmentation spontanée d'un tri-*O*-acyl tréhalose devenue moins détectable sous sa forme $[M+Na]^+$ n'est pas exclue.

III.E.3.d Conclusion : 3 acyl tréhaloses identifiés.

Grâce aux analyses effectuées, trois acyl-tréhaloses ont été mis en évidence dans la souche de référence *Mabs* ATCC19977 ainsi que dans la souche clinique *Mabs* GS163. Pour la première, le 3,4 di-*O*-acyl tréhalose (3,4 DAT) est substitué par des chaines lipidiques linéaires saturées ou monoinsaturées C_{16} et C_{18}, alors que le 2,3 di-*O*-acyl tréhalose (2,3 DAT), synthétisée en très grande quantité sur le milieu 7H10OADC, porte des acides gras courts avec une chaine polyméthylbranchées C_7 ou C_9. Les positions acylées et la nature des lipides sont identiques à celles décrites dans le dérivé di-*O*-acylé du TMM de la fraction apolaire. D'autre part, nous avons également retrouvé un acyl tréhalose apparenté au 3,4 DAT qui est majoritaire dans la souche clinique GS163 devenue lisse. Son comportement en CCM et son analyse RMN supposent qu'il pourrait s'agir d'un 3,4,6 tri-O-acyl tréhalose mais les résultats de l'analyse MALDI suggérent plutôt un 3,4 DAT portant un lipide plus long.

III.E.4 CARACTERISATION DES PIMs

III.E.4.a Purification des PIMs

Une séparation étendue des PIMs par chromatographie-flash sur colonne de silice a été réalisé au moyen d'un gradient de méthanol atteignant 50% dans le chloroforme. Le contrôle par CCM de chacune des fractions révèle des bandes de coloration violette possédant des R_f décroissant de la fraction 40 à la fraction 62 (Figure 30C). Les fractions 40 et 47 possèdent chacune une bande large et intense de coloration violette accompagnée d'une 2ème bande de R_f supérieur de coloration plus orangée. Les fractions 50 à 52 présentent deux autres bandes, l'une de coloration violette et l'autre de R_f inférieur grise. A partir de la fraction 57, deux fines bandes violettes d'intensité décroissante et avec le plus faible R_f subsistent. Sur les spectres de masse des fractions 40, 47, 53 et 57, la détection des ions *m/z* correspondants aux adduits $[M+2Na-H]^+$ en mode positif puis aux adduits $[M-H]^-$ en mode négatif permet la recherche des phospholipides dont les PIMs.

III.E.4.b Spectres MALDI-TOF des PIMs

Dans les spectres de masse de la fraction 53, plusieurs espèces sont détectées (Figure 41A). Le phosphatidyl-inositol (PI) correspond aux ions *m/z* $[M+2Na-H]^+$ 897 et $[M-H]^-$ 852 dont la composition lipidique théorique serait d'un acide palmitique C_{16} associé à un acide tuberculostéarique C_{19}. Le spectre MS^2 en mode négatif de l'ion parent *m/z* 852 présente un ion intense *m/z* 298 lié à la présence d'acide tuberculostéarique libre ainsi qu'un second fragment *m/z* 434 correspondant au motif lysophosphatidyl (Figure 41B).

Le PIM_1 est présent à l'état de trace puisque les ions correspondants *m/z* $[M+2Na-H]^+$ 1059 et $[M-H]^-$ 1014 sont moins intenses. Le décalage de 162 u.m.a. est lié à la présence du mannose supplémentaire. Ils sont accompagnés des ions *m/z* $[M+2Na-H]^+$ 1074 et $[M-H]^-$ 1028. Enfin, le PIM_2 est détecté avec une plus forte intensité sous forme de la paire d'ions *m/z* $[M+2Na-H]^+$ 1236 en mode positif et $[M-H]^-$ 1190 en mode négatif (Figure 41A). La fragmentation de ce dernier produit deux fragments *m/z* 298 et 434 communs avec le PI (Figure 41C). La présence des ions fragments *m/z* 934 et 636 montre la possibilité de perdre successivement les acides palmitique et tuberculostéarique. De plus, la rupture de la liaison phosphodiester avec le glycérol produit un ion fragment *m/z* 580 contenant l'inositol et les deux mannoses dont l'un est modifié. En effet, le décalage de masse entre le PI et le PIM_2 égale à 338 u.m.a. ne peut pas correspondre à l'addition de 2 hexoses de 162 Da.

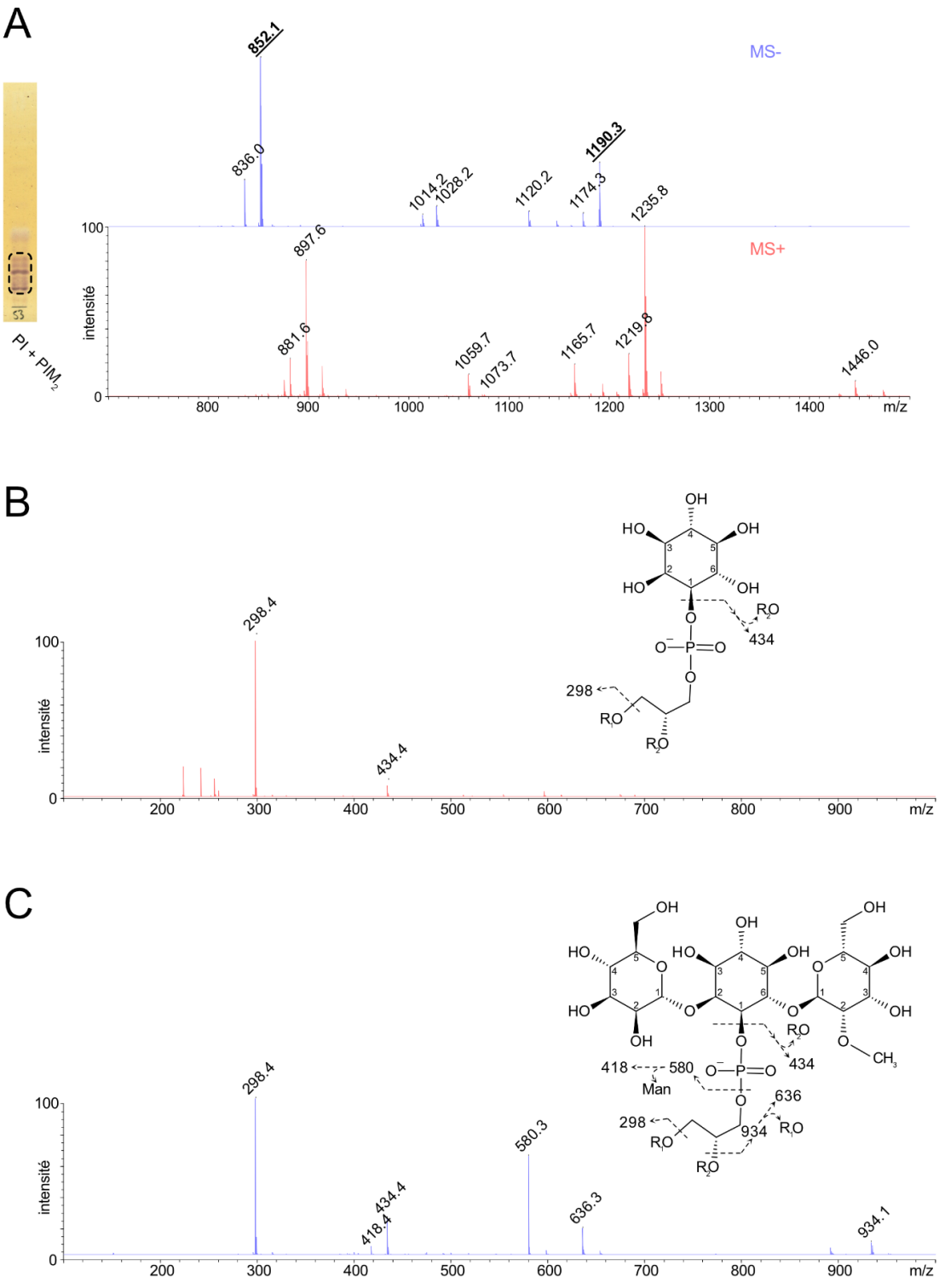


Figure 41 : Analyse MALDI-TOF du PI, PIM et PIM_2. Spectres en mode positif (rouge) et négatif (bleu) de la fraction 53 montrant les adduits $[M+2Na-H]^+$ et $[M-H]^-$ correspondant aux PI, PIM et PIM_2. L'ion *m/z* 852 est ensuite fragmenté en mode négatif afin de confirmer sa composition en acides gras C_{16} et C_{19} (B), de même que le PIM_2 à *m/z* 1190 qui produit de nouveaux fragments associés à la rupture de la liaison phosphodiester.

Dans la fraction 40, l'Ac_1PIM_2 est détecté par les paires d'ions *m/z* $[M+2Na-H]^+$ 1445 et $[M-H]^-$ 1400 sur les spectres positif et négatif (Figure 42A). Des ions décalés de 28 u.m.a. sont également visibles sur les deux spectres indiquant la présence possible de deux groupements méthylènes supplémentaires dans le lipide additionnel. D'autre part, le décalage de masse de 210 ou 238 u.m.a. entre le PIM_2 et l'Ac_1PIM_2 suggère l'addition probable d'une chaine d'acide myristique ou palmitique. De plus, les spectres MS^2 en mode positif des ions parents *m/z* 1446 produisent les ions fragments à *m/z* 1190 et 1148 (M-256 et M-298) correspondant la perte d'acide palmitique ou tuberculostéarique. L'ion fragment *m/z* 853 (M-594) est lié à la perte du groupement diacyl-glycérol, et les ions fragments *m/z* 437/479 à l'obtention d'un acide lysophosphatidique C_{16} ou C_{19} (Figure 42B).

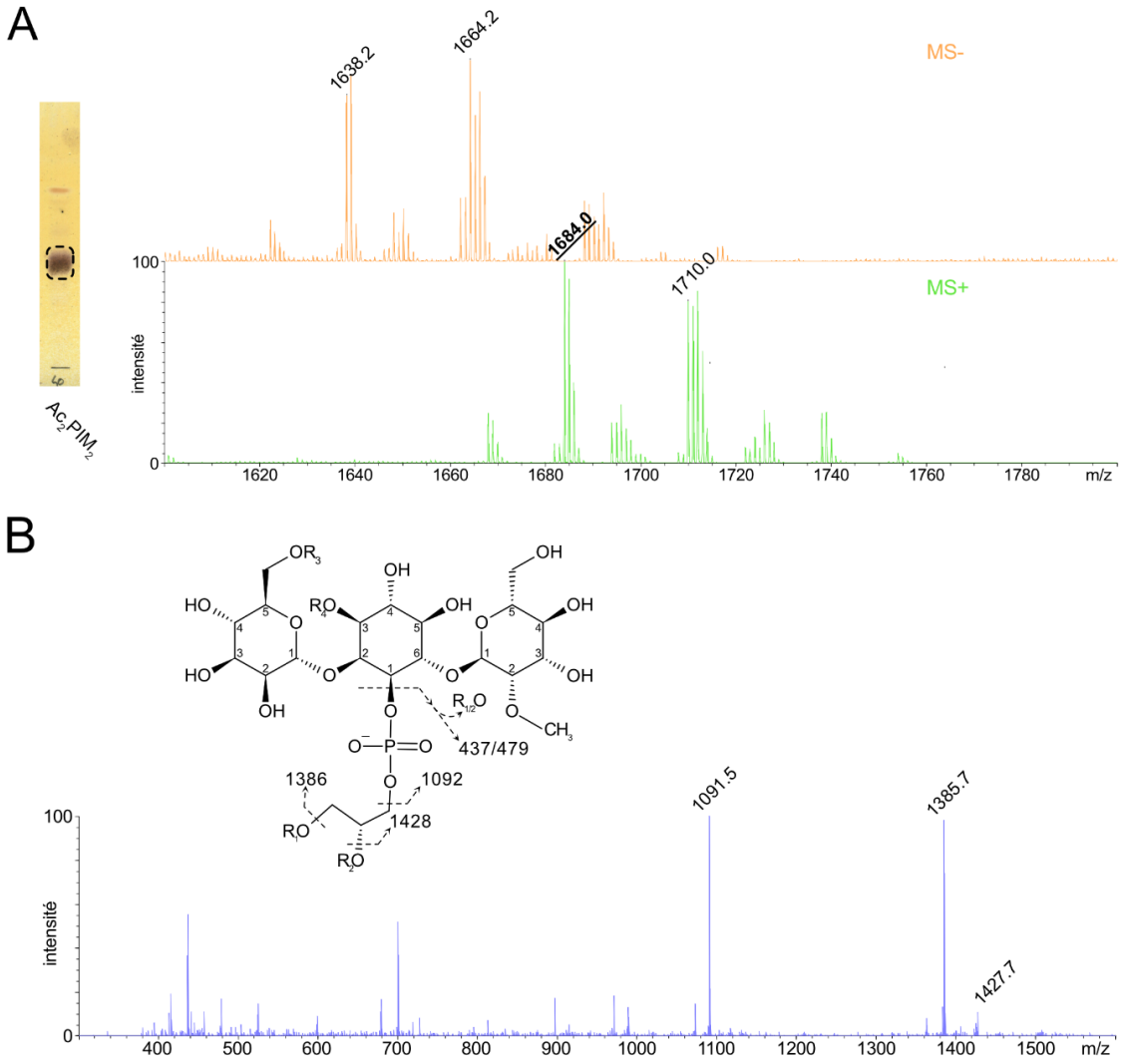


Figure 42 : Analyse MALDI-TOF de la fraction 40 contenant l'Ac_2PIM_2. La comparaison des spectres MS (A) en mode positif et négatif illustre les adduits $[M+2Na-H]^+$ 1684 et 1710 ainsi que $[M-H]^-$ 1638 et 1664. La sélection de l'ion *m/z* 1446 génère des fragments sur le spectre MS^2 (B) liés à la perte des groupements acyles (R_1 = tuberculostéarique, R_2 = palmitique, R_3 = myristique). Le résidu diglycéride peut également être perdu.

Pour la fraction 57, les spectres positif et négatif présentent les paires majoritaires d'ions *m/z* $[M+2Na-H]^+$ 2314 et $[M-H]^-$ 2359 ainsi que $[M+2Na-H]^+$ 2093 et $[M-H]^-$ 2048 correspondant aux Ac_2PIM_6 et Ac_1PIM_6 (Figure 43A & B). En outre, la différence de masse de 648 u.m.a. entre l'Ac_1PIM_2 et Ac_1PIM_6 est en accord avec l'ajout de 4 nouveaux résidus de mannose.

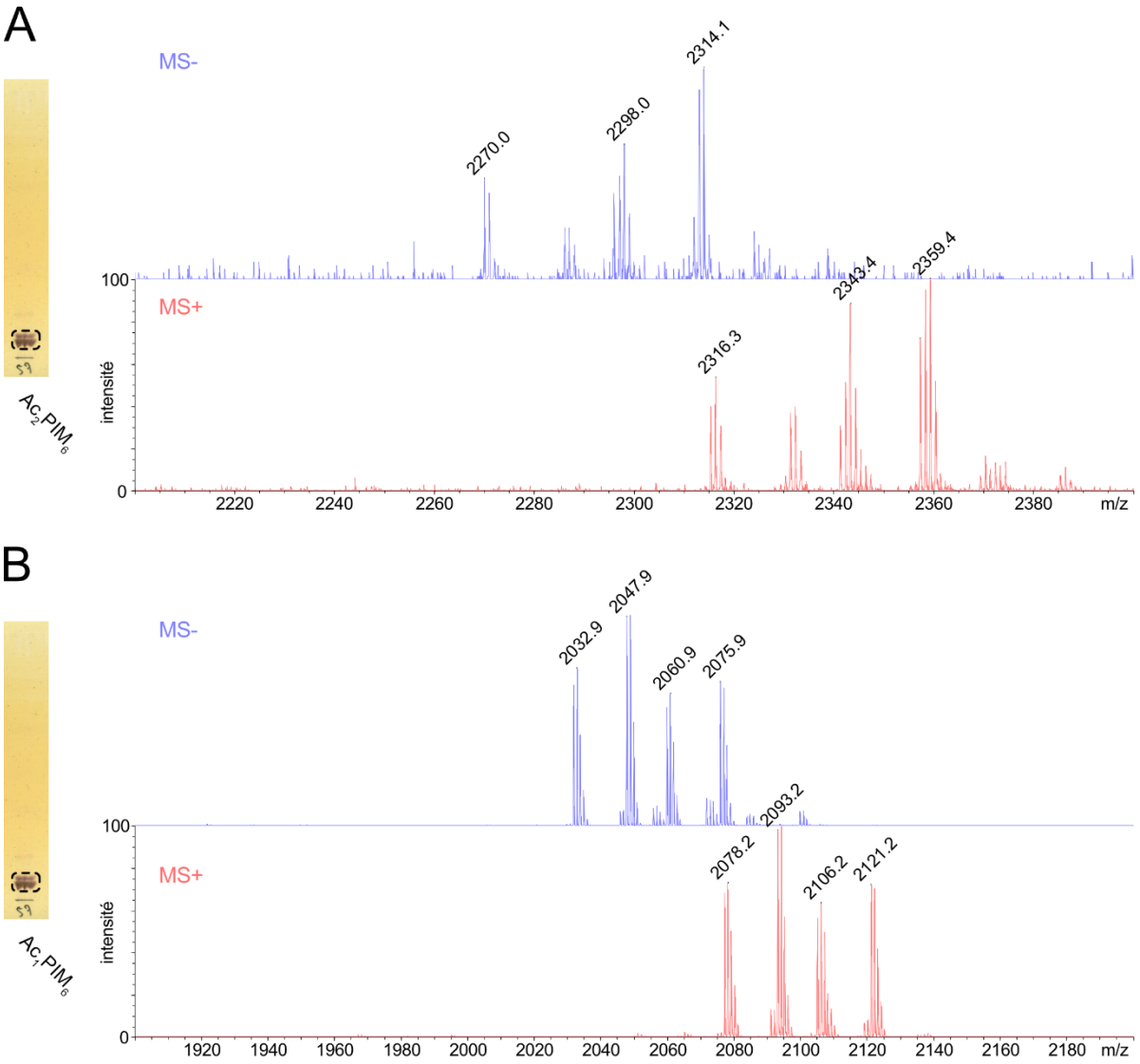


Figure 43 : Analyse MALDI-TOF de la fraction 57 contenant les PIM_6. La comparaison des spectres MS en mode positif et négatif illustre les adduits $[M+2Na-H]^+$ 2314 et $[M-H]^-$ 2359 correspondant à l'Ac_2PIM_6 (A) ainsi que [M+2Na-H]+ 2093 et [M-H]- 2048 correspondant à l'Ac_1PIM_6 (B). L'hétérogénéité des massifs autour de chaque ion montre à nouveau que la composition lipidique peut varier.

La purification et l'analyse des PIMs ont prouvé que les $AcPIM_2$ et les $AcPIM_6$ sont majoritaires. Les conditions de purification n'ont pas permis d'isoler des quantités suffisantes de PI, PIM_1 et PIM_2. D'une part, l'analyse MALDI-TOF a démontré que les acides gras C_{16} et C_{19} sont acylés sur le glycérol alors que des lipides C_{14} ou C_{16} le sont sur l'inositol et ou le mannose. D'autre part, les différences de

masse observées entre PI, PIM_1 et PIM_2 suggèrent que l'un des deux premiers mannoses portés par l'inositol est modifié, alors que les PIM_6 intègrent bien quatre résidus non modifiés.

III.E.4.c Composition en monosaccharides des PIMs

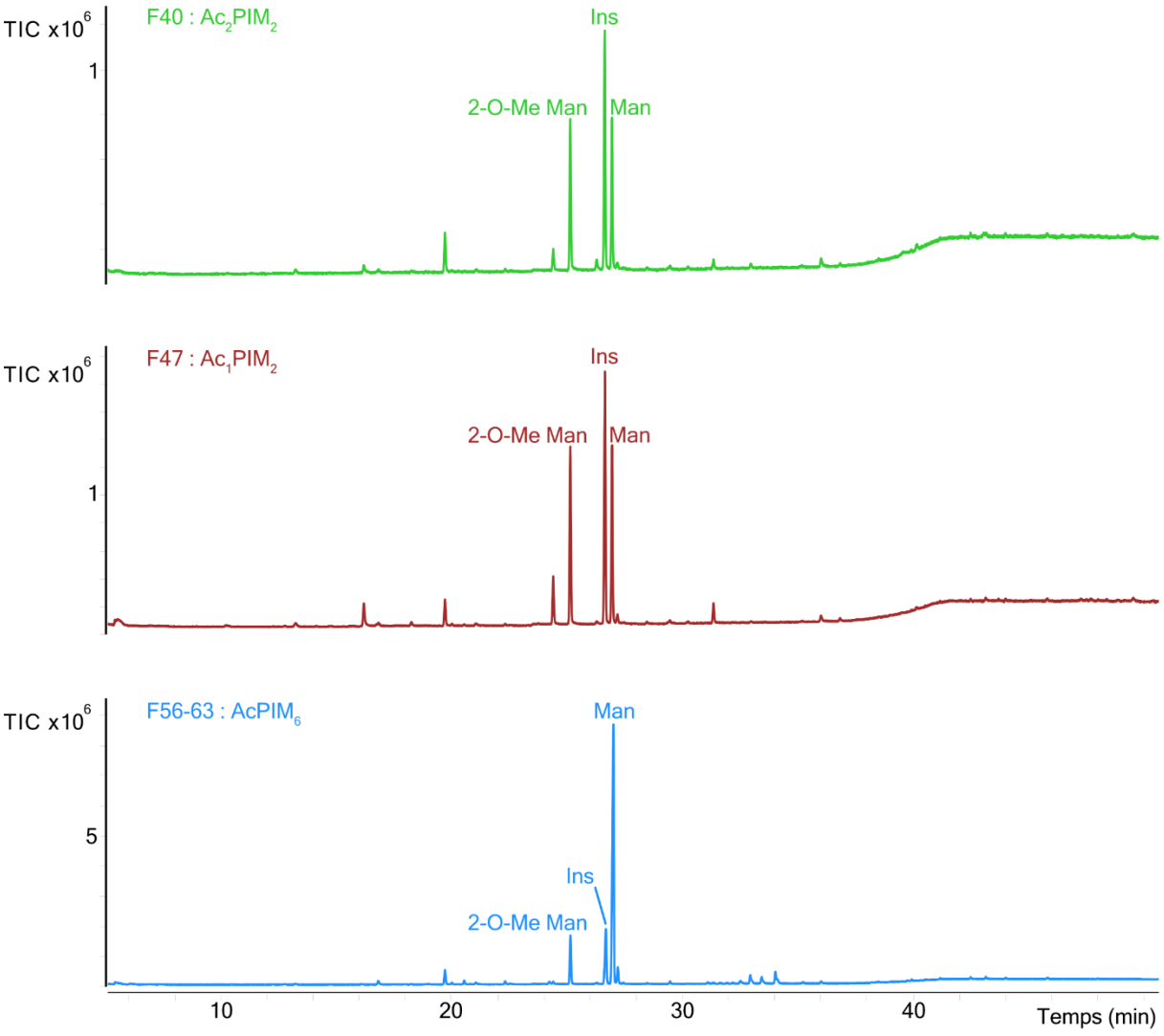


Figure 44 : Composition en monosaccharides des PIMs. Les fractions 40 (vert) et 47 (rouge) correspondant aux PIM_2 révèlent des proportions relatives similaires en 2-*O*-méthyl mannose (Tr = 25,15 min), *myo*-inositol (Tr = 26,64 min) et mannose (Tr = 26,96 min) alors que ce dernier est largement majoritaire dans le regroupement des fractions 56 à 63 correspondant aux PIM_6 (bleu).

L'analyse GC-MS des dérivés itols-acétates obtenus à partir des fractions 40, 47 ainsi que du regroupement des fractions 56-63 a permis de caractériser trois signaux chromatographiques majeurs (Figure 44). Le premier à Tr = 25,14 min a été identifié comme un résidu de 2-*O*-méthyl hexose d'après le spectre de fragmentation en impact électronique (Figure 45A). Le second pic à Tr = 26,64 min a été

attribué au *myo*-inositol d'après le temps de rétention du standard correspondant, et le spectre de fragmentation (Figure 45B). Enfin, le troisième pic chromatographique à Tr = 26,96 min et dont la proportion relative augmente significativement par rapport aux deux premiers dans la fraction 56-63, a été identifié comme un résidu de mannose conformément au temps de rétention et au spectre de fragmentation observé pour son standard (Figure 45C).

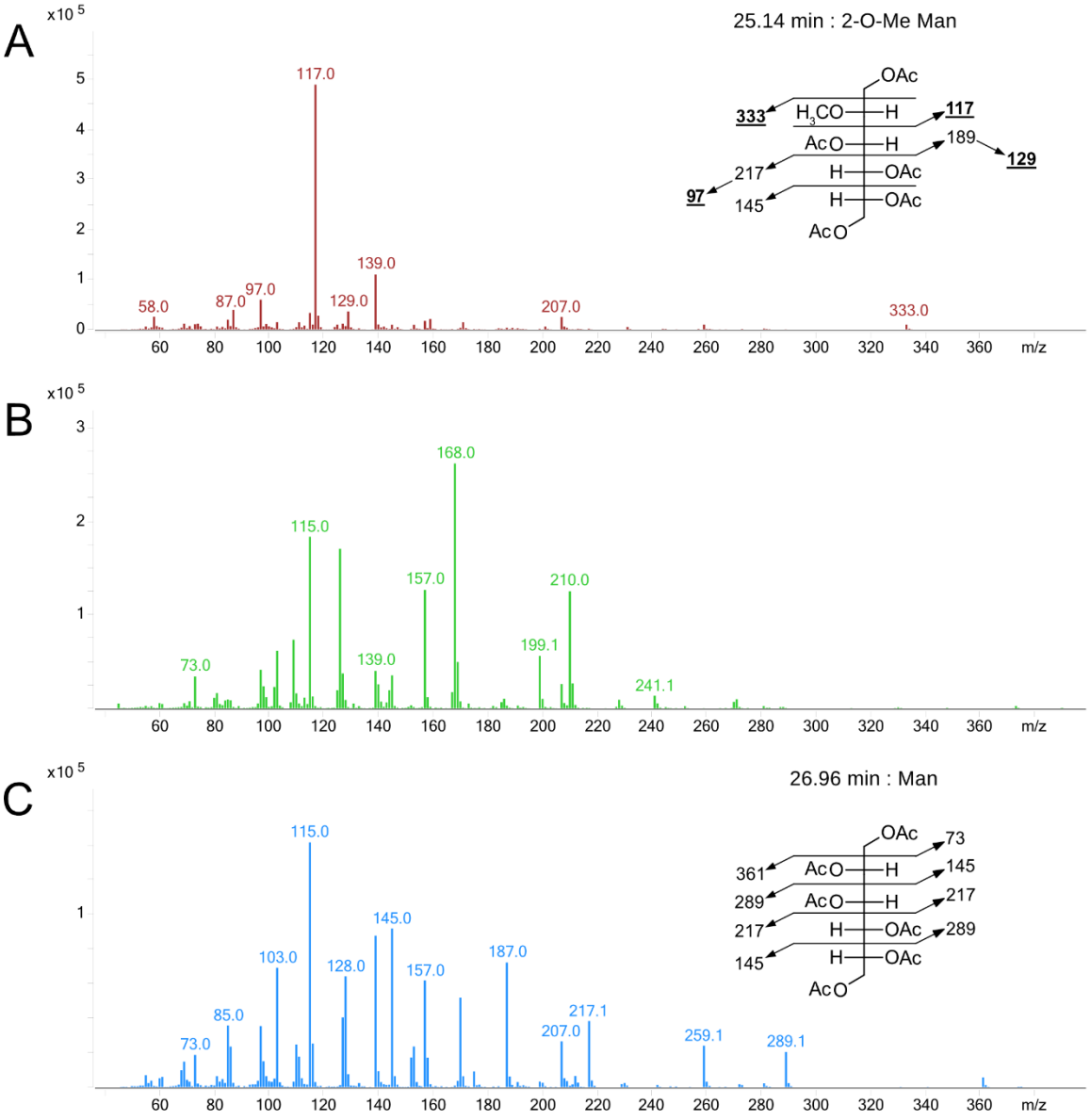


Figure 45 : Spectres EIMS des monosaccharides issus des PIMs. Fragmentation obtenue par impact électronique à 70 eV après séparation GC des monosaccharides contenus dans la fraction 40. Le 2-*O*-Me Man produit un ion très intense *m/z* 117 accompagné d'un second moins intense *m/z* 333 qui sont spécifiques d'une rupture au voisinage du groupement 2-*O*-méthyl (A). L'identification du *myo*-inositol et du mannose sont confirmés par le temps de rétention de leurs standards respectifs (B-C). Les spectres des 3 signaux majoritaires des fractions 47 et 56-63 sont identiques.

La nature identique des monosaccharides retrouvés dans les PIMs purifiés est cohérente avec la littérature indiquant que la biosynthèse des PIM_6 se fait à partir des PIM_2. Enfin, l'identification univoque du 2-*O*-méthyl hexose nécessiterait l'injection d'un standard de 2-*O*-méthyl mannose.

Tableau 15 : Contenu glycannique des PIMs identifiés de *Mabs*. Nature et nombre de mannoses retrouvés dans les PIMs après déduction des analyses MALDI-TOF et GC-MS de la composition en monosaccharides.

Adduit $[M+2Na-H]^+$	PIMs	2-O-Me Man	Ins	Man
897	PI	0	1	0
1059	PIM_1	0	1	1
1235	PIM_2	1	1	1
1446	Ac_1PIM_2	1	1	1
1684	Ac_2PIM_2	1	1	1
2094	Ac_1PIM_6	1	1	5
2332	Ac_2PIM_6	1	1	5

III.E.4.d Composition lipidique des $AcPIM_2$

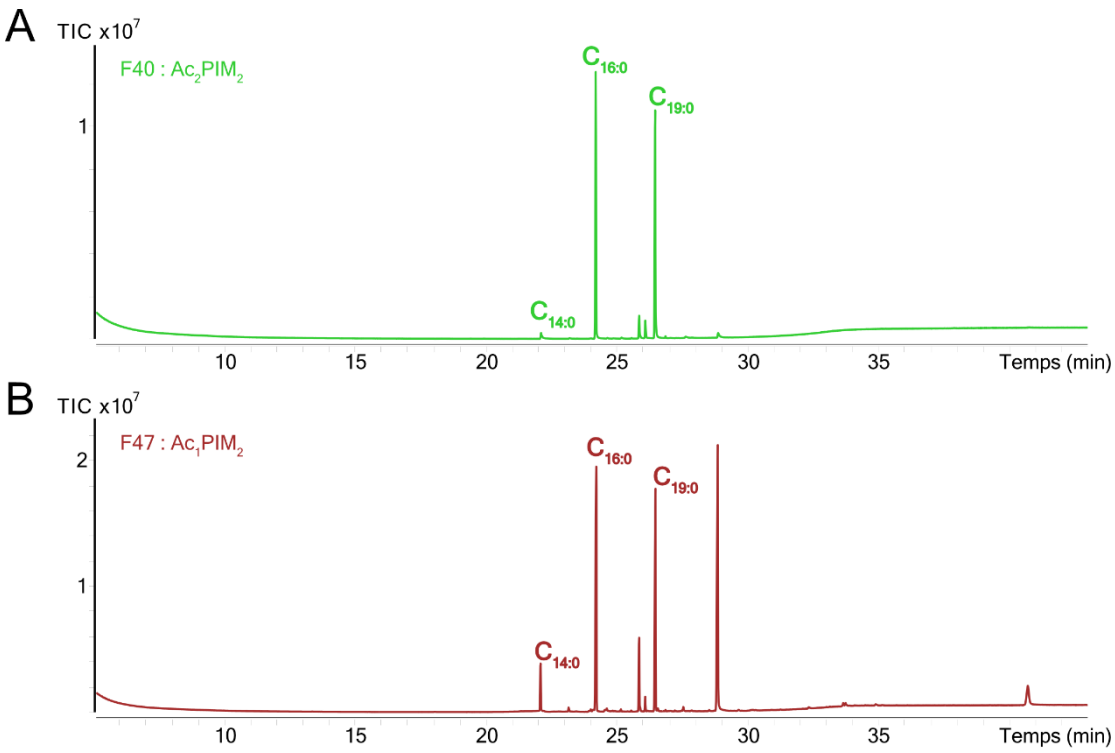


Figure 46 : Composition lipidique des $AcPIM_2$. Les FAMEs provenant des fractions 47 (A) et 40 (B) ont été repérés par l'ion spécifique *m/z* 74. De cette manière, les acides myristique (Tr = 22,08 min), palmitique (Tr= 24,19 min) et tuberculostéarique (Tr = 26,46 min) ont été caractérisés notamment grâce à leur spectre de fragmentations EIMS en figure 47.

Après avoir préparé les esters méthyliques d'acides gras sur les fractions correspondantes aux Ac_1PIM_2 et Ac_2PIM_2, leur repérage sur le chromatogramme obtenu s'est fait *via* l'extraction de l'ion caractéristique *m/z* 74. Dans le cas de l'Ac_1PIM_2, trois signaux sont apparus comme les plus intenses (Figure 46A). Le premier d'entre eux à Tr = 22,08 min est le moins intense, et il a été identifié comme de l'acide myristique C_{14} d'après son spectre de fragmentation (Figure 47A) et la comparaison de son temps de rétention avec un mélange de standards. Le plus intense est le second pic à Tr = 24,22 min qui a été identifié comme de l'acide palmitique C_{16}. Il est suivi à Tr = 26,46 min par de l'acide tuberculostéarique C_{19}. L'identification de ces deux acides gras résulte à nouveau de l'interprétation de leur spectre de fragmentation, et du temps de rétention dans le cas du palmitate (Figure 47B & C). Pour l'Ac_2PIM_2, les mêmes acides gras ont pu être identifiés même si la proportion relative de de l'acide myristique a largement diminué (Figure 46B).

Tableau 16 : Contenu lipidique des PIMs identifiés de *Mabs*. Nature et nombre des lipides retrouvés dans les PIMs après déduction des analyses MALDI-TOF et GC-MS de la composition en FAMEs.

Adduit $[M+2Na-H]^+$	**PIMs**	**$C_{14:0}$**	**$C_{16:0}$**	**$C_{19:0}$**
897	PI	0	1	1
1059	PIM_1	0	1	1
1235	PIM_2	0	1	1
1446	Ac_1PIM_2	1	1	1
1474		0	2	1
1684	Ac_2PIM_2	1	2	1
1712		0	3	1
2094	Ac_1PIM_6	1	1	1
2122		0	2	1
2332	Ac_2PIM_6	1	2	1
2360		0	3	1

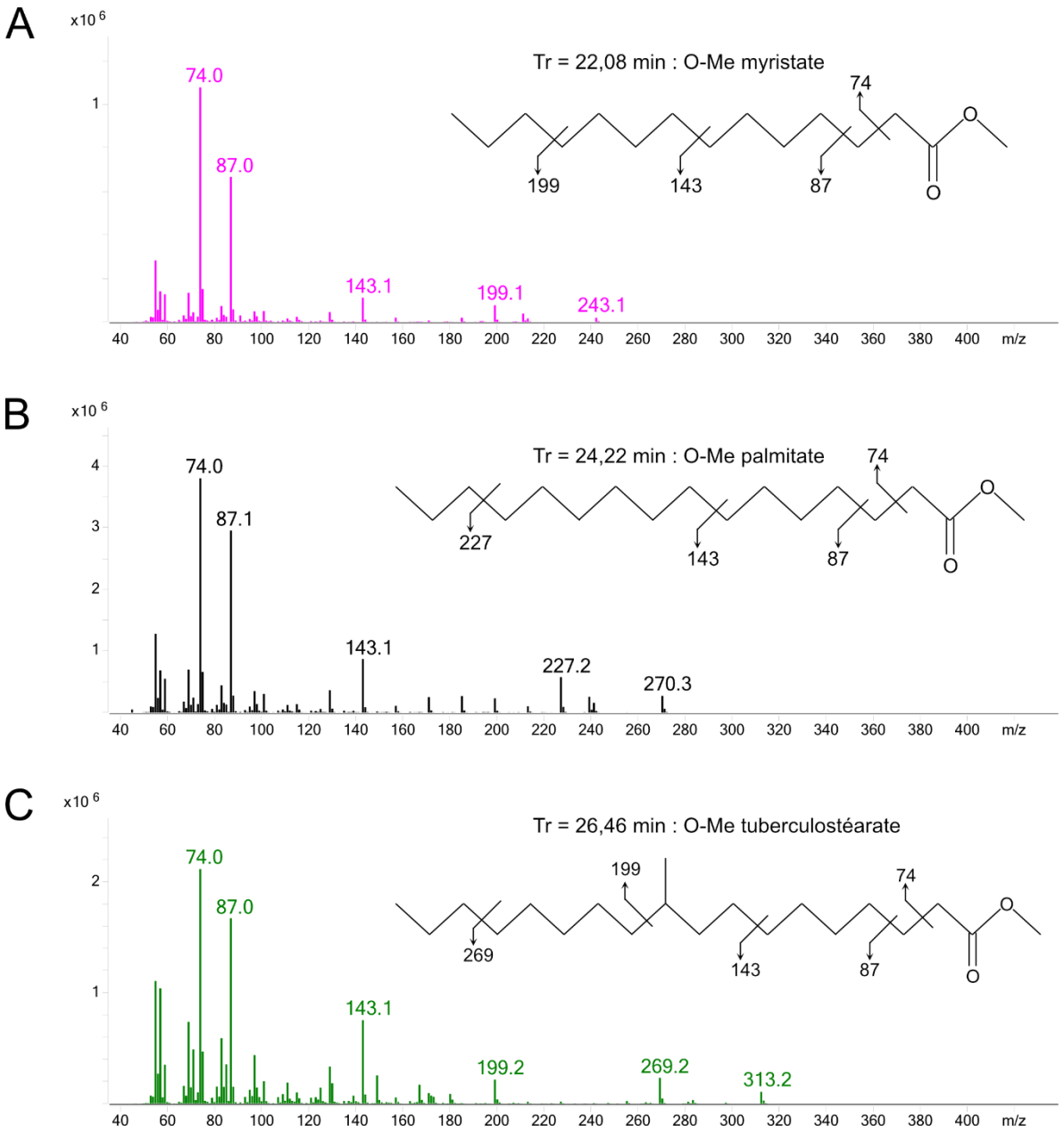


Figure 47 : Spectres EIMS des FAMEs issus des $AcPIM_2$. Les acides myristique (A), palmitique (B) et tuberculostéarique (C) ont été caractérisés notamment grâce à leur spectre de fragmentations EIMS dans les fractions 40 et 47.

Les acides gras identifiés par l'analyse de leurs esters méthyliques renforcent l'interprétation précédente des spectres de fragmentation MALDI. En effet, ils confirment la présence majoritaire d'acide palmitique et tuberculostéarique, et dans une moindre proportion d'acide myristique.

III.E.4.e Spectres RMN de l'Ac_2PIM_2

Le spectre $^1H/^{13}C$ HSQC de l'Ac_2PIM_2 révèle deux signaux anomères distincts à 5,11/102 ppm (M'-1) et 5,16/99,1 ppm (M-1) correspondant à chacun des deux mannoses supposés (Figure 48A). M'1 présente une corrélation avec M'-2 à 3,97/70,5 ppm identifiable sur le $^1H/^1H$ COSY (Figure 48B), alors que la corrélation suivante avec M'-3 à 3,71/71,2 ppm se distingue mieux sur le spectre $^1H/^1H$ TOCSY.

Tableau 17 : Signaux RMN de l'Ac_2PIM_2. Déplacements chimiques $^1H/^{13}C$ attribués à l'ensemble des positions identifiés de l'Ac_2PIM_2 ainsi que les corrélations protons carbones. La position 2 du 2-*O*-méthyl mannose (M-2) présente un signal ^{13}C déblindé et une corrélation longue distance avec le méthoxy. De même, les corrélations HMBC réciproques confirment la liaison du *myo*-inositol en 2 du mannose et en 6 du 2-*O*-méthyl mannose. Enfin, même si les corrélations avec le groupement carbonyle ne sont pas visibles pour les M'-6 et G-2, ces dernières sont acylées comme les positions G-1 et I-3.

	δ ^1H (ppm)	δ ^{13}C (ppm)	HMBC	HSQC/TOCSY
M-1	5,16	99,1	79,8/73,6	80,4
M-2	3,74	80,4	59,3	
M-3	3,83	71		80,4/73,4/67,9
M-4	3,54	67,9		73,6/71
M-5	3,92	73,6		71/68/62,1
M-6	3,80/3,68	62,1		
M'-1	5,11	102	76,3/71,2	70,5/64,2
M'-2	3,97	70,5		67,9
M'-3	3,71	71,2		70,3/67,6/64,1
M'-4	3,54	67,7		71,2
M'-5	3,96	71,2		71,2/67,7
M'-6	4,36/4,21	64,1		
I-1	4,14	76,8		
I-2	4,31	76,3	102	
I-3	4,73	71,9	174,1	79,8/76,8/73,9/71,2
I-4	3,71	71,2		79,8/73,9/
I-5	3,33	73,9		79,8/76,8/71,9/71,2
I-6	3,75	79,8	99,2	
G-1	4,38/4,13	63,1	174,6	70,8/64
G-2	5,2	70,8		64/63,1
G-3	3,97	64		70,8/63,1

Le second signal anomèrique M-1 corrèle avec M-2 à 3,74/80,4 ppm (Figure 48B). Le signal ^{13}C fortement déblindé suggère que cette position porte un groupement méthoxy identifié à 3,49/59,3 ppm qui présente une corrélation longue distance sur le spectre $^1H/^{13}C$ HMBC avec M-2 (tableau 17), ainsi que des corrélations spatiales sur le spectre $^1H/^1H$ NOESY avec M-1 et M-2 (Figure 48B). Le lien avec M-3 à 3,83/71 ppm se confirme par une nouvelle corrélation sur le spectre $^1H/^1H$ TOCSY. A partir de celle-ci, on identifie M-4 à 3,54/67,9 ppm et M-5 à 3,92/73,5 ppm sur le spectre $^1H/^1H$ COSY (Figure 48B) et avec les carbones correspondants sur le spectre HSQC/TOCSY (Figure 49A).

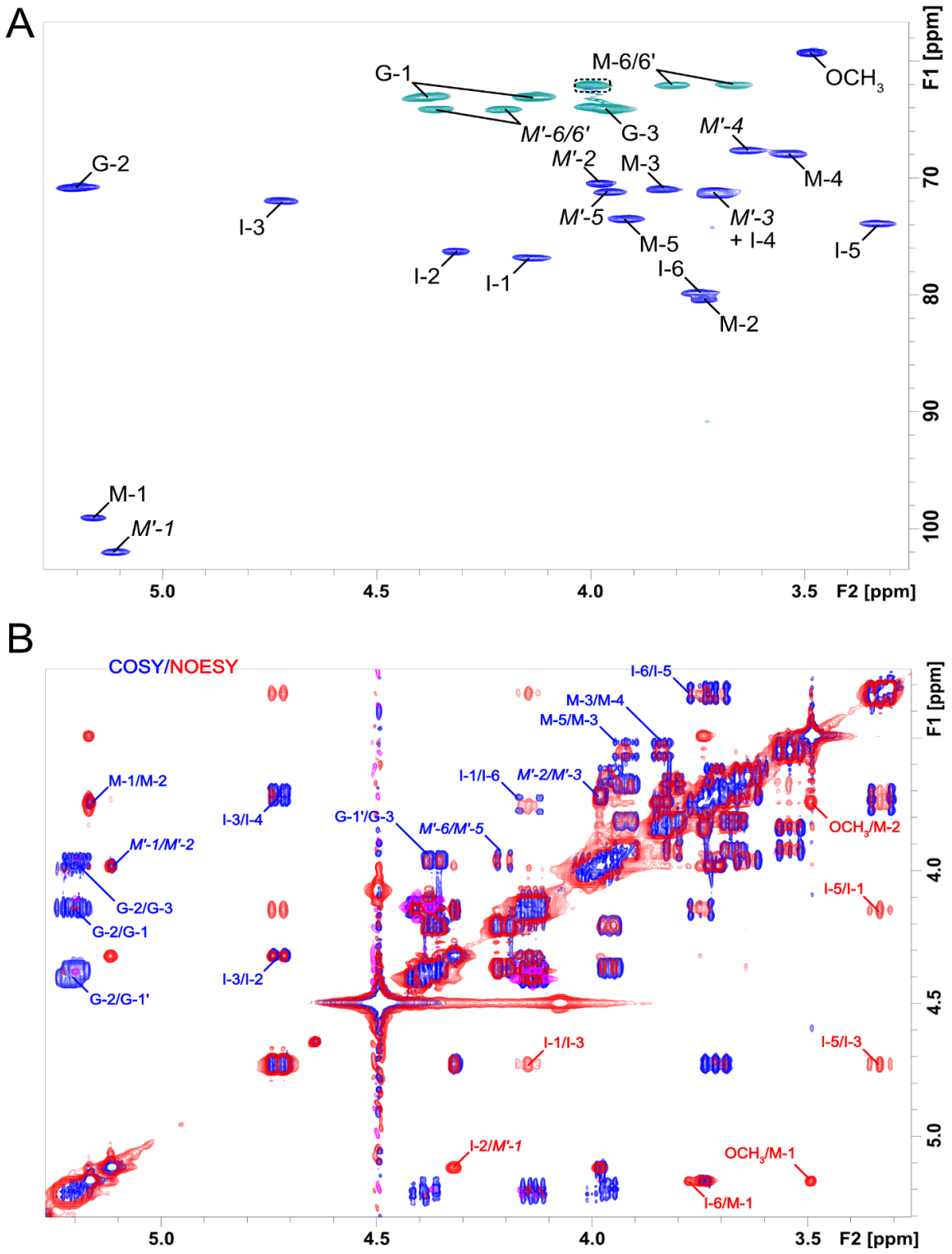


Figure 48 : Analyse RMN de l'Ac_2PIM_2. Spectres mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K, calibré avec le méthanol. Sur le spectre $^1H/^{13}C$ HSQC (A), les signaux en vert correspondent à des CH_2 et celui entouré est un probable contaminant phosphoéthanolamine. Sur les spectres $^1H/^1H$ COSY/NOESY superposés (B), les corrélations spatiales sont affichées en rouge et les corrélations ^3J scalaires en bleu. La dénomination est rappelée dans le tableau 17 et sur la structure de la figure 50.

Le spectre $^{1}H/^{1}H$ COSY affiche également les corrélations entre les 3 positions du résidu de glycérol (Figure 48B). Le groupement méthine (G-2) est le plus fortement déblindé à 5,20/70,8 ppm. Il est couplé d'une part au groupement méthylène (G-3) à 3,97/64 ppm, et d'autre part aux 2 signaux non équivalents du groupement méthylène (G-1) à 4,38 et 4,13/63,1 ppm. Ce dernier signal possède une corrélation longue distance avec un groupement carbonyle à 174,6 ppm sur le spectre HMBC.

Le spectre ^{31}P a montré 2 signaux à 0,341 et - 0,523 ppm mais c'est ce dernier qui a produit les corrélations scalaires protons/phosphore cohérentes avec l'identification du système de spin de l'inositol (Figure 49B). Le premier a été attribué à une possible molécule de phospho-éthanolamine contaminant notre fraction. Il est corrélé avec le signal d'un groupement méthylène $^{1}H/^{13}C$ à 4/62 ppm lui-même corrélé sur le spectre $^{1}H/^{1}H$ COSY avec un second méthylène à 3,08/41 ppm.

L'attribution des positions du groupement *myo*-inositol a été déduite des corrélations $^{1}H/^{1}H$ COSY/NOESY et des signaux méthines visibles sur le spectre $^{1}H/^{13}C$ HSQC à partir de la position 3 (Figure 48A & B). En effet, I-3 à 4,73/71,9 ppm est fortement déblindé du fait de son acylation : la corrélation est d'ailleurs visible avec le groupement carbonyle à 174,1 ppm sur le spectre $^{1}H/^{13}C$ HMBC. De plus, elle est directement corrélée sur le spectre $^{1}H/^{1}H$ COSY avec I-4 à 3,71/71,2 ppm et I-2 à 4,31/76,3 ppm (Figure 48B). Cette dernière est glycosylée par le résidu de mannose M'-1 car elle possède un signal carbone plus déblindé ainsi qu'une corrélation spatiale $^{1}H/^{1}H$ NOESY avec M'-1.

Compte tenu de la stéréochimie du *myo*-inositol, les corrélations spatiales $^{1}H/^{1}H$ NOESY de la position 3 ont permis d'identifier I-1 à 4,14/76,8 ppm et I-5 à 3,33/73,9 ppm qui sont du même côté du cycle à 6 carbones (Figure 48B). I-1 étant lié au groupement phosphate, son signal ^{13}C est très déblindé. De plus, les corrélations scalaires longue distance protons/phophore avec I-3, I-4 et I-5 sont nettement visibles sur le spectre $^{1}H/^{31}P$ HSQC à - 0,524 ppm (Figure 49B). Le déblindage du carbone de I-2 et sa corrélation NOE avec M-1 démontre que cette position est glycosylée par le résidu de mannose méthoxylé.

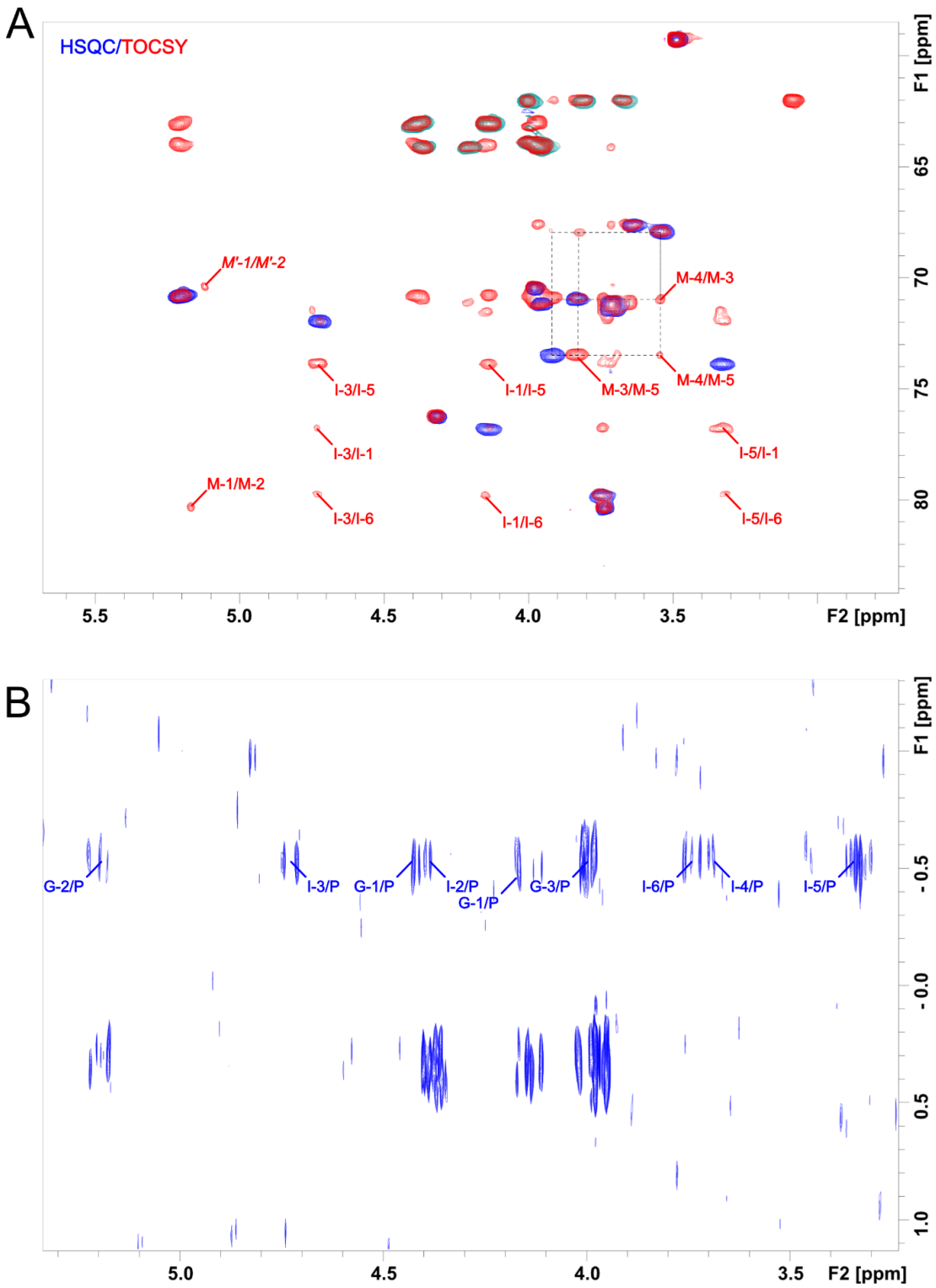


Figure 49 : Attribution des signaux RMN de l'Ac_2PIM_2. Spectres mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K. Les spectres $^1H/^{13}C$ HSQC et HSQC-TOCSY (A) superposés et calibrés sur le méthanol ont permis d'attribuer les positions M-3 à M-5. D'autre part, l'alignement simultanée des corrélations du signal ^{31}P à -0,523 ppm sur le spectre $^1H/^{31}P$ HSQC avec les signaux protons du glycérol et du *myo*-inositol confirment l'attribution de ce phosphore à l'Ac_2PIM_2.

Le déblindage de M'-6/M'-6' à 4,36 et 4,21/64,1 ppm démontre que cette position est acylée (Figure 48A) mais le groupement carbonyle correspondant n'a pas pu être déterminé sur le spectre $^{1}H/^{13}C$ HMBC. Ces signaux sont corrélés entre eux sur le spectre $^{1}H/^{1}H$ COSY ainsi qu'avec M'-5 à 3,96/71,2 ppm. A partir de cette position, le spectre $^{1}H/^{13}C$ HSQC-TOCSY montre la corrélation avec M'-4 à 3,54/67,7 ppm et M'-3 précédemment décrite (Figure 49A). Finalement, M-6/M-6' à 3,80 et 3,68/62,1 ppm est attribué à la position 6 du second résidu de mannose méthoxylé.

S'agissant des lipides, on retrouve les signaux caractéristiques d'une fonction éthylènique, des groupements méthyléniques en α du carbonyle ou de l'insaturation, de la chaine lipidique linéaire saturée se terminant par un groupement méthyle, d'un groupement méthyle branché issu de l'acide tuberculostéarique.

III.E.4.f Conclusion : structure résolue de l'Ac_2PIM_2

L'ensemble de ces données démontre la structure du motif 2,3-di-*O*-acyl-glycéro-phosphatidyl-(2,6-di-*O*-mannopyranosyl) *myo*-inositol dont l'originalité réside dans la présence d'un groupement méthyl sur le résidu de mannose lié à la position I-6 de l'inositol (Figure 50). Cette nouvelle caractéristique structurale de l'Ac_2PIM_2 de *Mabs* a été notamment confirmée en 2020 par Palcekova et collaborateurs[212]. Nos données montrent aussi qu'elle était présente sur l'Ac_1PIM_2 et se répercute sur les PIM_6 contenant également le 2-*O*-méthyl mannose.

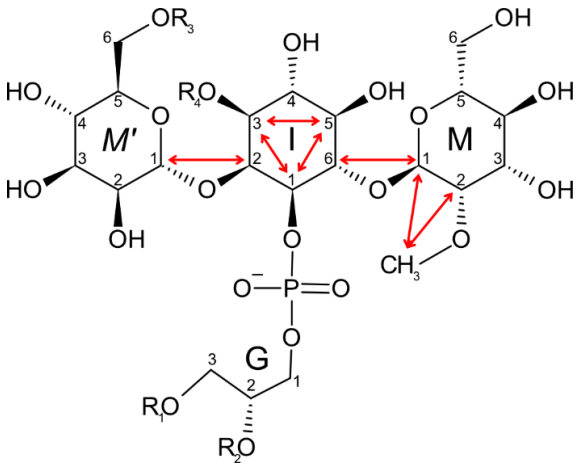


Figure 50 : Structure de l'Ac_2PIM_2 purifié de *Mabs* déduite de l'analyse RMN et affichant la stéréochimie des résidus mannoses et *myo*-inositol, ainsi que les corrélations $^{1}H/^{1}H$ NOESY établissant les positions des liaisons glycosidiques et du groupement méthoxy. D'après l'analyse MALDI-TOF et la composition lipidique, R_1 est arbitrairement un acide tuberculostéarique alors que R_2, R_3 et R_4 sont des acides palmitiques mais R_3 peut être remplacé par un acide myristique.

III.E.5 ETUDE DE LA BIOSYNTHESE DES GPLs DE *Mabs* S

En collaboration avec l'équipe du Dr Laurent KREMER qui a développé des méthodes de biologie moléculaire adaptées aux mycobactéries afin de cibler la délétion des gènes de la biosynthèse des GPLs, j'ai initié l'étude des GPLs sauvages ainsi que de nouvelles structures identifiées chez les mutants de la forme lisse de *Mabs* dans le cadre du projet ANR SunLive. La plupart des gènes étudiés ont été caractérisés notamment chez *Msmeg*, mais pas directement chez *Mabs* (Figure 11). Ce travail a donc commencé par l'étude du rôle du gène *fmt* sur la modification du lipide associé aux GPLs ainsi que sur l'impact de cette modification dans l'observation de phénotypes nouveaux comparativement à la souche sauvage. Dans un deuxième temps, j'ai étudié le profil des GPLs de sept mutants simples et multiples pour les gènes *gtf1*, *gtf2* et *gtf3* codant des glycosyltransférases impliqués dans la glycosylation des GPLs. Enfin, compte tenu de la duplication de deux gènes *atf1* et *atf2* codant des acétyltransférases dans le génôme de *Mabs* par rapport à celui de *Msmeg*, nous avons conclu ce projet sur leurs rôles possibles dans l'acétylation séquentielle des GPLs via l'étude des structures de neuf nouveaux mutants codant plusieurs acétyltransférases.

III.E.5.a Analyse comparative des GPLs de *Mabs* S et du mutant *Δfmt* (1er article)

C'est travaux ont donné lieu à la publication d'un article en 2020 dans le journal *ACS Infectious Diseases* dont je suis co-premier auteur [390]. Les résultats seront présentés dans la version libre de droits de l'article et commentés ci-dessous, en faisant référence aux figures originales de l'article.

Le morphotype lisse sauvage de *Mabs* produit dans l'extrait polaire total de grande quantité de GPL dont deux glycoformes principales nommés GPL-2a et GPL-3. Ces composés sont facilement détectables par spectrométrie de masse MALDI-TOF sous la forme de deux paires d'ions *m/z* 1258-1286 et 1404-1432 (Fig. 3C). L'analyse respective de leur spectre de fragmentation confirme la structure décrite des motifs glycanniques ainsi que leur position puisque les ions *m/z* 1404 et 1432 perdent soit un déoxyhexose relié à l'alaninol, soit un déoxyhexose di-*O*-acétylé relié à la thréonine et respectivement identifiés comme un résidu de Rha et de 6dTal (Fig. S3A). L'hydrolyse alcaline entraine d'ailleurs la perte des deux groupements acétyles et l'observation sur les spectres de masse d'un décalage de 84 u.m.a. pour le GPL-2a à *m/z* 1174-1202 et le GPL-3 à *m/z* 1320-1348 (Fig. S4A). L'analyse de la composition en monosaccharides de chacun des deux GPLs purifiés révèle que les résidus de 3,4 di-*O*-méthyl rhamnose, le rhamnose et le 6-déoxy-talose sont présents dans des proportions équimolaires pour le GPL-3 (Fig. S2). Par contre, le GPL-2a ne contient pas de rhamnose (Fig. S1). L'étude du spectre $^{1}H/^{13}C$ HSQC confirme non seulement le nombre de monosaccharides dans chacun

des 2 GPLs mais aussi la présence de trois groupements méthoxy dont deux appartenant au 3,4 di-*O*-méthyl rhamnose, et le troisième au lipide acylé côté N-terminal (Fig. 3D).

En parallèle, l'étude du profil des GPLs du mutant *Δfmt* a révélé que la composition en monosaccharides des GPL-3 et GPL-2a était identique aux GPLs de la souche sauvage (Fig. S1 & S2). Par contre, les bandes les plus intenses révélées en CCM ont des R_f plus faibles, les spectres de masse MALDI-TOF des GPLs natifs et saponifiés présentent des paires d'ions majoritaires décalés de 14 u.m.a. et les spectres $^{1}H/^{13}C$ HSQC illustrent la disparition du signal du groupement méthoxy associé au lipide (Fig. 3C & D). Afin de confirmer la position de cette méthoxylation, j'ai réalisé une double analyse des esters méthyliques d'acides gras, avant et après triméthylsilylation, à partir des GPLs purifiés de la souche sauvage et du mutant *Δfmt*. La comparaison de la fragmentation obtenue en impact électronique d'un standard de stéarate de méthyle β-méthoxylé a confirmé la présence des mêmes dérivés C_{26} et C_{28} chez la souche sauvage via la présence d'un ion spécifique *m/z* 117 (Fig. 4C), alors que le mutant *Δfmt* n'en contenait pas. A l'inverse, après triméthylsilylation, je suis parvenu à détecter les acides gras C_{26} et C_{28} β-triméthylsilylés en plus grande quantité à partir des GPLs du mutant *Δfmt* via la présence d'un ion spécifique *m/z* 175 (Fig. 4D). Enfin, l'étude approfondie des spectres de fragmentation MS^2 et MS^3 des GPLs purifiés natifs et saponifiés du mutant *Δfmt* a révélé la présence d'un mécanisme de fragmentation associé à une rupture de la liaison C-C au voisinage du groupement β-hydroxy, et qui n'était pas mis en évidence pour les GPLs de la souche sauvage (Fig. S5). Cette caractéristique a été utilisée pour rapidement caractériser les GPLs non méthoxylés également retrouvés au sein de trois mutants ciblés dans le domaine d'interaction avec la S-adénosyl méthionine (Fig. S6). Ainsi le rôle biochimique du gène *fmt* chez *Mabs* est identique à celui de *Mtf2* chez Msmeg[341]

Parallèlement à mon étude structurale, l'étude phénotypique a montré que l'absence de méthoxylation n'avait pas d'impact sur l'aspect des colonies ou leur croissance. Par contre, la partition dans l'hexadécane indique une baisse d'hydrophobicité de l'enveloppe et la microscopie de force atomique montre une diminution des forces d'adhésion. De plus, le mutant *Δfmt* s'est révélé moins phagocyté par les macrophages THP-1 que la souche sauvage. Finalement, quelle que soit la structure du lipide, les motifs glycanniques n'étaient pas modifiés entre les 2 souches, suggérant que la biosynthèse du lipide n'avait pas d'impact sur les étapes de glycosylation du noyau peptidique.

O-Methylation of the glycopeptidolipid acyl chain defines surface hydrophobicity of *Mycobacterium abscessus* and macrophage invasion

Wassim Daher[1,4,&], Louis-David Leclercq[2,&], Albertus Viljoen[1,3], Jona Karam[1], Yves F. Dufrêne[3], Yann Guérardel[2], and Laurent Kremer[1,4,#]

[1]Centre National de la Recherche Scientifique UMR 9004, Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France.

[2]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.

[3]Louvain Institute of Biomolecular Science and Technology, Université Catholique de Louvain, Croix du Sud, 4-5, bte L7.07.07., B-1348, Louvain-la-Neuve, Belgium.

[4]INSERM, IRIM, 34293 Montpellier, France.

[#]To whom correspondence should be addressed: Tel: (+33) 4 34 35 94 47;

E-mail: laurent.kremer@irim.cnrs.fr

[&]These authors contributed equally to that work.

Mycobacterium abscessus, an emerging pathogen responsible for severe lung infections in cystic fibrosis patients, displays either smooth (S) or rough (R) morphotypes. The S-to-R transition is associated with reduced levels of glycopeptidolipid (GPL) production and is correlated with increased pathogenicity in animal and human hosts. While the structure of GPL is well established, its biosynthetic pathway is incomplete. In addition, the biological functions of the distinct structural parts of this complex lipid remain elusive. Herein, the *fmt* gene encoding a putative *O*-methyltransferase was deleted in the *M. abscessus* S variant. Subsequent biochemical and structural analyses demonstrated that methoxylation of the fatty acyl chain of GPL was abrogated in the *Δfmt* mutant and this defect was rescued upon complementation with a functional *fmt* gene. In contrast, introduction of *fmt* derivatives mutated at residues essential for methyltransferase activity failed to restore GPL defects, indicating that *fmt* encodes an *O*-methyltransferase. Unexpectedly, phenotypic analyses showed that *Δfmt* was more hydrophilic than its parental progenitor, as demonstrated by hexadecane-aqueous buffer partitioning and atomic force microscopy experiments with hydrophobic probes. Importantly, the invasion rate of THP-1 macrophages by *Δfmt* was reduced by 50% when compared to the wild-type strain. Together, these results indicate that Fmt *O*-methylates the lipid moiety of GPL and plays a substantial role in conditioning the surface hydrophobicity of *M. abscessus* as well as in the early steps of interaction between the bacilli and macrophages.

Keywords: *Mycobacterium abscessus*, glycopeptidolipid, *O*-methyltransferase, hydrophobicity, phagocytosis, atomic force microscopy.

Infections caused by non-tuberculous mycobacteria (NTM) are increasing globally and are usually difficult to treat due to their inherent resistance to many common antibiotics [1]. Among NTM, *Mycobacterium abscessus* is a fast-growing species and an emerging human pathogen that causes nosocomial skin and soft tissue infections [2] but also pulmonary infections, especially in patients with cystic fibrosis (CF) and other lung disorders [3,4]. Like other NTM, *M. abscessus* exhibits either smooth (S) or rough (R) colony morphotypes, which are associated with distinguishable *in vitro* and *in vivo* phenotypes. This colony-based distinction relies on the production of high (in S) or low (in R) levels of surface-associated glycopeptidolipids (GPL) [5–7]. Comparative genomic and transcriptomic studies have been conducted to better understand the molecular mechanisms responsible for the S-to-R transition in *M. abscessus*, revealing multiple insertions or deletions occurring in in the R variants, mainly in the *mps1*, *mps2*, *gap*, and *mmpL4b* genes that are involved either in the synthesis or transport of GPLs [8]. Epidemiological surveys showcased the prominence of the *M. abscessus* R strain in patients with severe pulmonary infections [9,10] and with chronic colonization of the airways in CF patients [10], underlining that the distinction between S and R morphotypes is of paramount importance and clinically relevant. Indeed, numerous studies have emphasized that the presence or loss of GPL conditions important physiological and physiopathological aspects of infection, among which changes in sliding motility, biofilm formation and susceptibility to antibiotics have been reported [5,7,11–13]. Other properties related to the GPL content have been disclosed, such as bacterial surface hydrophobicity [14,15], aggregation leading to cord formation [5,13,14,16], interaction with and intracellular trafficking in macrophages [17,18] and induction of a pro-inflammatory response [19], all influencing the clinical outcome of the disease. This has also been raised by recent studies using various cellular and animal models, confirming the increased pathogenesis of the R over the S form [13,16,20,21]. In particular, the zebrafish model of infection has been proposed as a relevant and genetically tractable host-pathogen conjugate for dissecting *M. abscessus* interactions with host cells [22], which clearly showed that the S-to-R transition is associated with exacerbation of the bacterial burden, the formation of massive serpentine cords, abscess formation and increased larval killing [16]. The mycobacterial cell envelope comprises, from the inside to the outside, a plasma membrane surrounded by the peptidoglycan layer, covalently connected to an arabinogalactan layer, which itself is attached to the outermost membrane bilayer, designated the mycomembrane [23,24]. The inner leaflet of this outer membrane is predominantly composed of mycolic acids while the outer leaflet contains a mixture of free lipids, such as triacylglycerol (TAG), diacylglycerol (DAG) and a series of complex lipids, including trehalose mono- and di-mycolate as well as GPL. From a structural point of view, GPL are composed of a lipopeptide ornamented with a flexible arrangement of glycosylation that is built from *O*-methylated and *O*-acetylated deoxy hexoses [18,25,26] **(Figure1A)**. The peptidic core consists of a D-phenylalanine-D-*allo*threonine-D-alanine tripeptide linked to an L-alaninol. This D-Phe-D-*allo*Thr-D-Ala-L-alaninol sequence is assembled by two non-ribosomal peptide synthetases, Mps1 and Mps2, and acylated with a 3-hydroxy/methoxy C_{24}-C_{33} fatty acid by an unidentified acyltransferase [26–28]. The lipopeptide core is

glycosylated with the *allo*-Thr linked to a 6-deoxy-α-L-talose (6-dTal), while the alaninol is linked to an α-L-rhamnose (Rha), producing the less-polar diglycosylated GPL species. In addition to the diglycosylated GPL that contain a 3,4-di-*O*-acetylated 6-dTal and a 3,4-di-*O*-methylated or 2,3,4-tri-*O*-methylated Rha [18,26,29], *M. smegmatis* and *M. abscessus* also produces more polar GPL by the addition of a 2,3,4-tri-hydroxylated Rha to the alaninol-linked 3,4-di-*O*-methyl Rha [26,30]. Although being structurally identical, triglycosylated GPL are more abundant in *M. abscessus* than in *M. smegmatis* [26]. GPL are heterogenous in structure and vary according to the fatty acyl chain length and the degree of hydroxylation or *O*-methylation of the glycosidic moieties through the action of various *O*-glycosyltransferases, *O*-methyltransferases and *O*-acetyltransferases **(Figure1A)** [31]. Interestingly, a recent study revealed unexpected structural features of the GPL content in *M. abscessus* grown on CF sputum or synthetic CF medium, consisting of an increase in the amount of triglycosylated forms and possible replacement of the classical carboxy terminal alaninol by other branched-chain aminol alcohol, valinol or leucinol [32]. Despite the relevance of GPL in NTM of clinical significance, the GPL biosynthetic pathway, which has been well delineated in *M. smegmatis* and *M. avium* [25,33-36], is only partly described in *M. abscessus*. The individual role of several genes compulsory for GPL production, particularly modifications of the lipid part of GPL, remains to be experimentally investigated. In addition, whether changes in the acyl chain influences the surface properties of *M. abscessus* and eventually its early interaction with host cells remain undetermined. In this study, we questioned as to whether *fmt*, a gene located within the GPL biosynthetic locus of *M. abscessus* **(Figure1B)** participates in GPL biogenesis. Construction of a genetically-deficient *fmt* mutant was used to demonstrate that *fmt* encodes an *O*-methyltransferase. Importantly, deletion of *fmt* abrogated *O*-methylation of the fatty acyl chain of GPL, which strongly impacted the bacterial surface properties and phagocytosis of *M. abscessus* by human macrophages. Overall, this study advances our understanding of the genetic requirements for the biosynthesis and biological functions of the high-GPL producing variant of this understudied human pathogen.

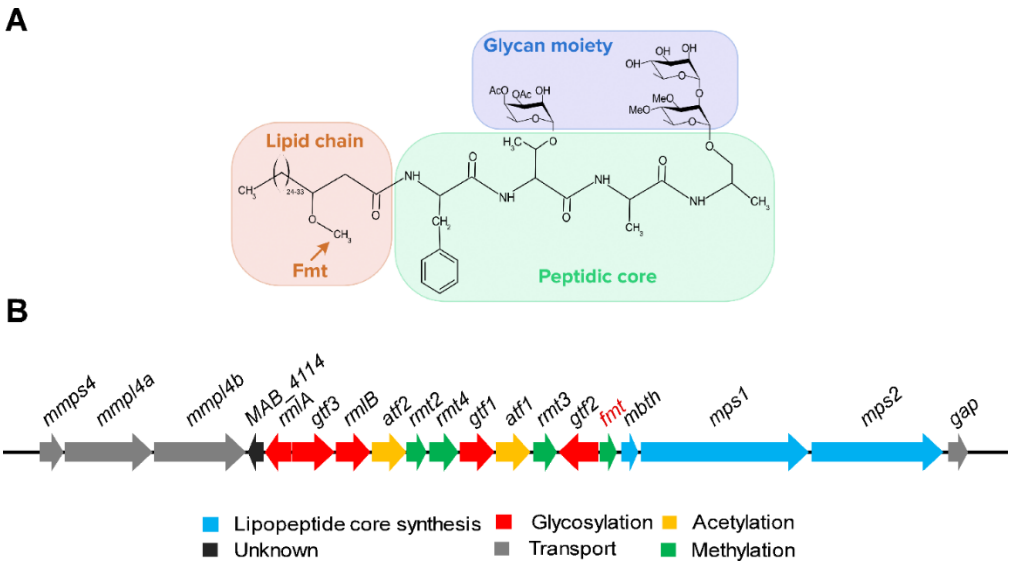


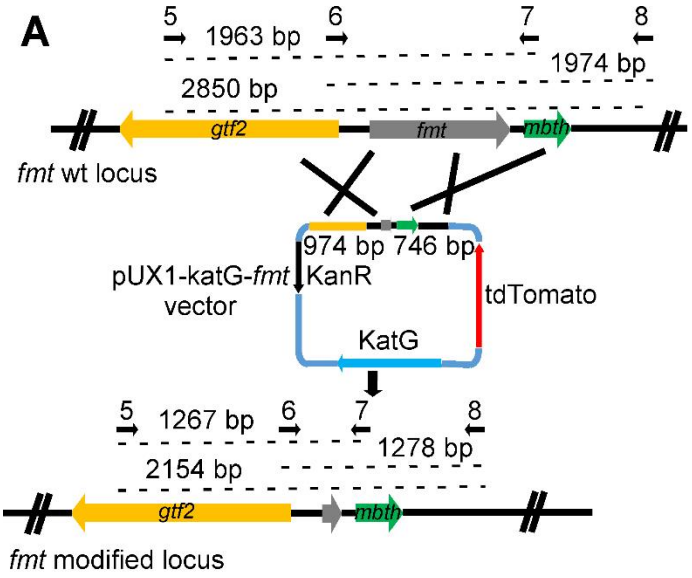
Figure 1. The biosynthetic locus and structure of *M. abscessus* GPL. (A) The lipid, peptidic and glycan moieties of GPL are indicated by different colors. The schema illustrates the triglycosylated form of GPL (GPL-3) consisting of a D-Phe-D-*allo*Thr-D-Ala-D-alaninol peptidic core (green) of which the D-*allo*Thr is glycosylated with a 3, 4 di-*O*-acetyl 6-deoxy-L-talose residue and the alaninol is diglycosylated with rhamnose and 3,4-di-*O*-methylated rhamnose residues. The three monosaccharides are presented in the blue area. The orange arrow shows the methyl group, presumably transferred by the fatty acid *O*-methyltransferase enzyme (Fmt) onto position 3 in the lipid chain (orange area). **(B)** The GPL locus encodes multiple components required for biogenesis, modification and transport of GPL. *fmt* is indicated in orange.

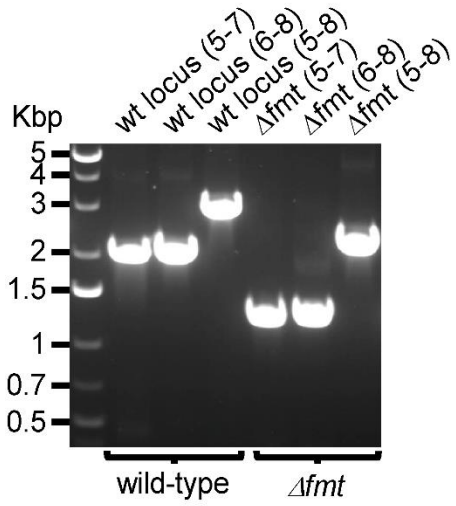
RESULTS

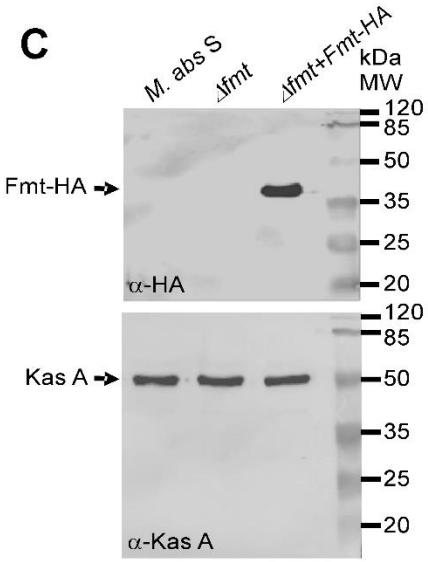
Loss of *fmt* affects neither growth nor colony morphology of *M. abscessus in vitro*. To get insights into the function of the Fmt protein, conventional gene inactivation was performed in the high-GPL producing *M. abscessus* S variant using an unmarked deletion system [37]. The strategy involves double homologous recombination leading to the removal of the *fmt* open reading frame from the GPL locus **(Figure 2A)** using the pUX1-*katG* suicide construct, which carries a kanamycin resistance (*kan*^{R}) cassette, a tdTomato red fluorescence marker and a *katG* cassette that confers sensitivity to isoniazid, an antitubercular to which *M. abscessus* is naturally resistant [37]. Two DNA sequences flanking *fmt* were cloned adjacently to each other into pUX1-*katG* effectively leading to an *fmt*-deleted allele in the resulting pUX1-*katG-fmt* vector. After transformation, clones that integrated pUX1-*katG-fmt* into their chromosome *via* a first homologous recombination were isolated by their resistance to kanamycin and their red fluorescence. A subsequent second homologous recombination step resulting in the loss of pUX1-*katG-fmt* allowed isolation of potential *Δfmt* clones by their resistance to isoniazid, loss of red fluorescence and kanamycin sensitivity. PCR and sequencing screens using primers listed in **Table S1** on genomic DNA confirmed the genotype of a *Δfmt* mutant **(Figure 2B).** To confirm that *fmt* alone was responsible for any phenotypic defects observed in the *Δfmt* strain, the mutant was complemented by site-specific integration at the L5 mycobacteriophage chromosomal attachment site [38] of a C-terminally HA-tagged *fmt* copy under control of its endogenous promoter **(Figure 2C)**. Westernblotting was carried out by loading an equal amount of total protein extracts (confirmed by the KasA protein internal loading control). A band with the approximate molecular weight expected for the Fmt-HA protein was detected in the complemented strain, designated *ΔfmtC*.

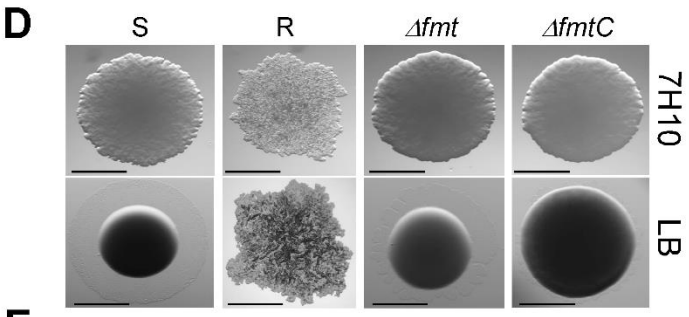
Since it is well established that the absence or the presence of GPL conditions the morphotype of *M. abscessus* colonies on agar medium, we next assessed whether disruption of *fmt* influences the morphology of the strain. Serial dilutions of cultures were plated on Middlebrook 7H10 or on LB agar and single colonies were observed after 4-5 days of incubation. No visual differences in the morphology were noticed between the WT (S morphotype), *Δfmt* and *ΔfmtC* colonies, regardless of the medium used; and these strains did not show the typical serpentine cords typifying the R variant **(Figure 2D)**. Moreover, *Δfmt* failed to aggregate in liquid culture but produced homogenous suspensions similar to *M. abscessus* S cultures **(Figure 2E)**. The *in vitro* growth curves in 7H9 broth at 37°C were also similar for the S, *Δfmt* and *ΔfmtC* strains, indicating that

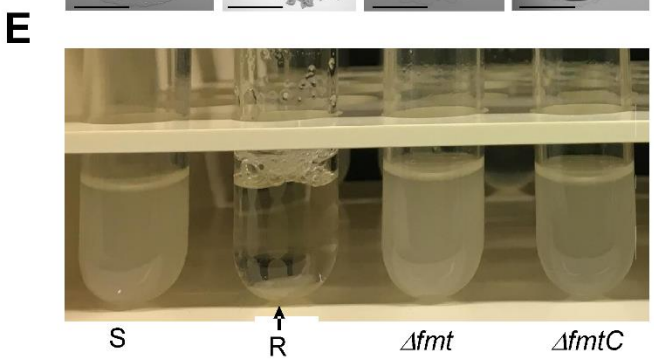
deletion of *fmt* does not impact on the replication rate of *M. abscessus* inplanktonic culture **(Figure2F).** In the same line, like the S progenitor, *Δfmt* produced fragile pellicles at the air-liquid interface **(Figure2G)**.

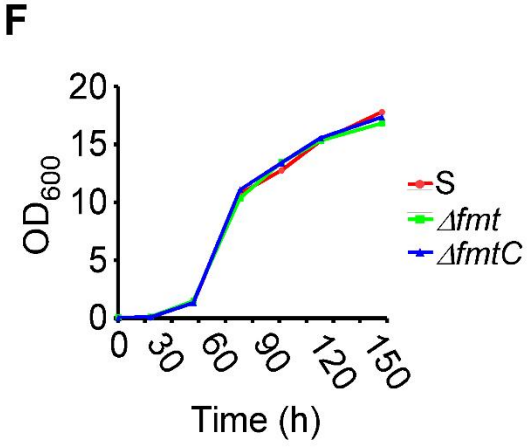


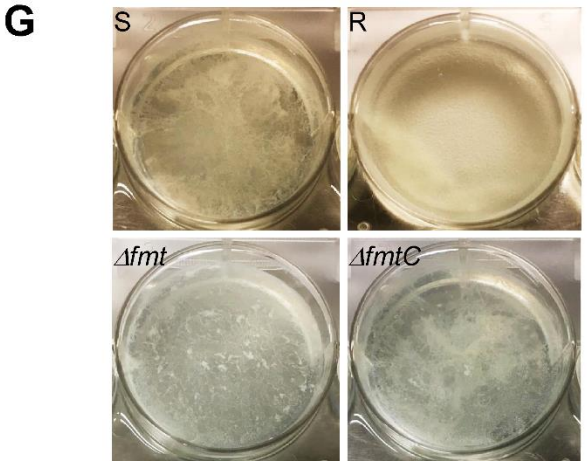












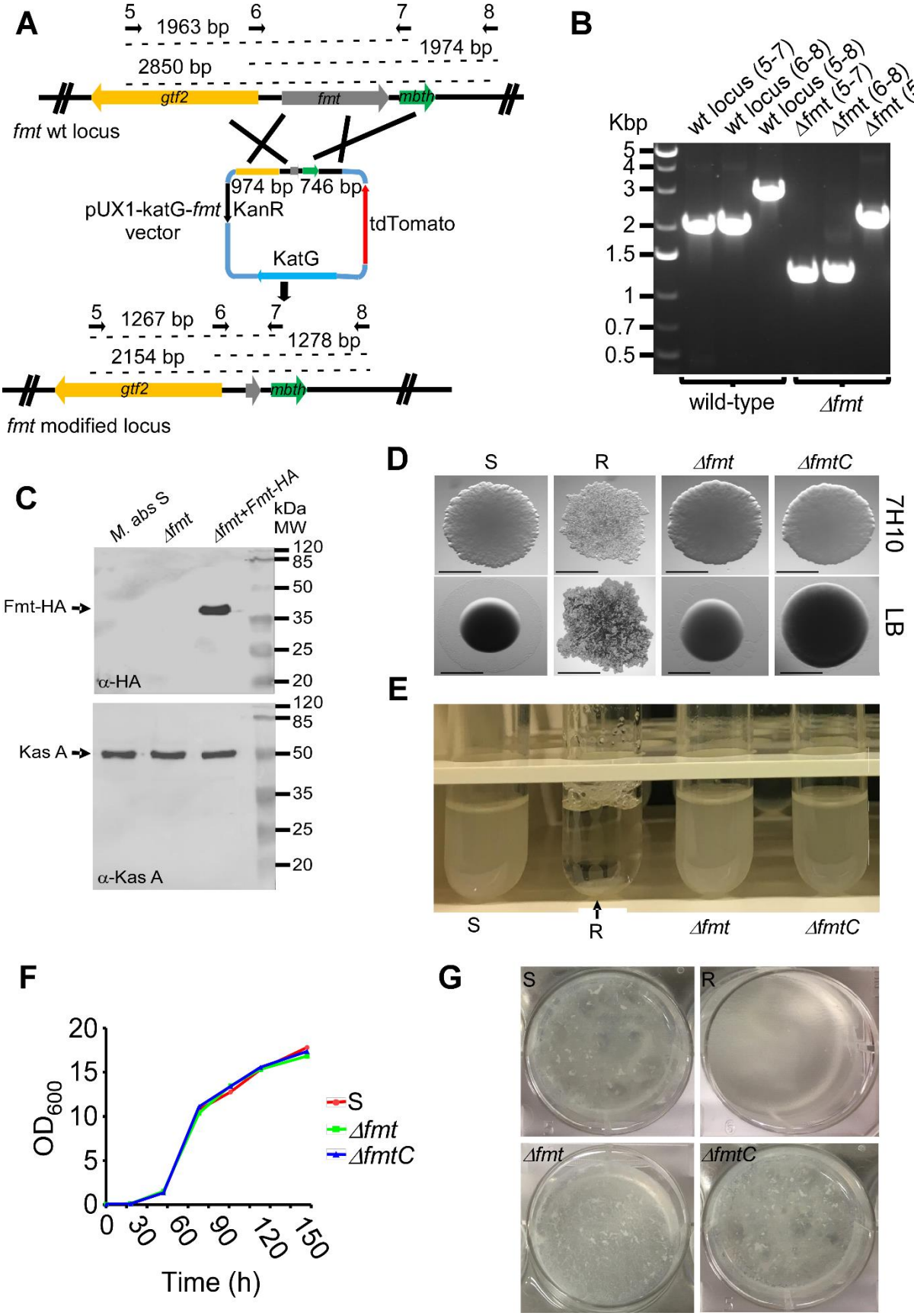


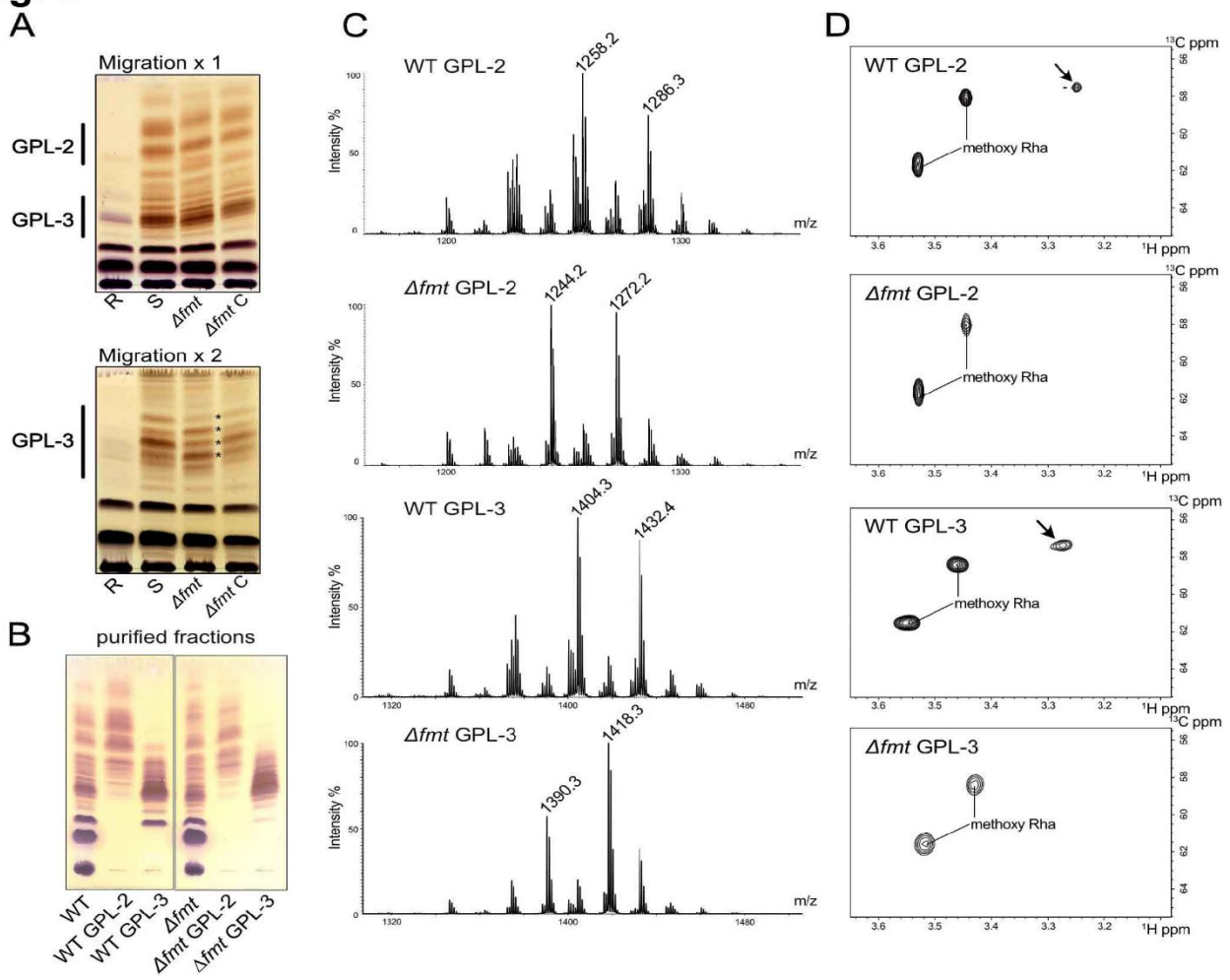
Figure2. Generationofanunmarked*fmt* deletion mutant in *M. abscessus*.(A)The*fmt* gene is neighbored by the genes *gtf2* and *mbth*. To generate *Δfmt*, two sequences of 974 bp and 746 bp located upstream and downstream of *fmt* were cloned into pUX1-*katG*.The resulting suicide plasmid (lacking motifs for episomal replication or mobile elements promoting chromosomal integration) was used to transform *M. abscessus*,

in which it could only propagate *via* homologous recombination between the cloned sequences and their chromosomal homologs. A first recombination event was favored in the absence of antibiotic and clones in which a cross-over occurred were selected in the presence of kanamycin. A single red fluorescent tdTomato-expressing clone was selected and subjected to the second round of recombination. Clones in which a second recombination event occurred were selected on isoniazid and screened for double cross-over phenotypes, *i.e.* loss of red fluorescence, sensitivity to kanamycin and resistance to isoniazid. Dotted lines represent the size (indicated above each line) of the expected PCR products in *M. abscessus* WT and *Δfmt*. Black arrows represent the primers used for PCR analysis. **(B)** PCR analysis demonstrating deletion of *fmt* in the mutant strain. Genomic DNA from WT bacteria was used to amplify the intact *fmt* locus. Amplicons were subjected to sequencing to confirm deletion of *fmt*. **(C)** Western blot analysis of the *ΔfmtC* strain expressing the Fmt protein in fusion with an HA tag under the control of the native *fmt* promotor. KasA was included as a loading control. **(D)** Colony morphology of the S, R, *Δfmt* and *ΔfmtC* strains either on $7H10^{OADC}$ or LB plates. No difference in colony morphology between the progenitor S and *Δfmt* strains was found, regardless of the media used. Experiments were carried out twice. **(E)** Like the parental S strain, but unlike the R variant, *Δfmt* does not sediment rapidly in liquid culture. The arrow indicates the sedimented bacterial aggregates. **(F)** Growth curve analysis of parental S, *Δfmt* and *ΔfmtC* strains in 7H9 medium at 37°C. Data is representative of three independent experiments. **(G)** Pellicle-forming capabilities of *M. abscessus* S, R, *Δfmt* and *ΔfmtC* strains.

Disruption of *fmt* alters the overall GPL profile in *M. abscessus*. GPL are a heterogenous family of polar glycolipids whose structures have been well established in several NTM, including *M. abscessus*[18,25,26]. Individual GPL comprises three structural parts, a Phe-*allo*Thr-Ala-alaninol peptidic core, a lipid chain and a glycan moiety. In *M. abscessus*, a C_{24}-C_{33} methoxylated 3-hydroxy fatty-acyl chain esterifies the amino group of the Phe, while the *allo*Thr and alaninol are substituted by a 6-dTal and by one or two methylated Rha residues, respectively **(Figure1A)**. The presence of one or two Rha residues differentiates diglycosylated (GPL-2) from triglycosylated (GPL-3) GPL species. The impact of *fmt* deletion on the overall GPL profile was first analyzed by lipid extraction and separation by thin-layer chromatography (TLC). As anticipated, GPL production was abrogated in the *M. abscessus* R morphotype whereas the S variant produced large amounts of highly heterogenous di- and triglycosylated GPL **(Figure3A**, upper panel**)** [26]. Major differences were noticed in the migration profile between the S variant and *Δfmt*, presumably resulting from small chemical changes in individual molecules. A second round of migration showed that the modified GPL pattern in *Δfmt* comprises GPL species that were either produced in lower amounts or accumulating **(Figure3A**, bottom panel**)**. Importantly, the wild-type (WT) GPL profile was restored in *ΔfmtC*, indicating that the altered GPL profile cannot be attributed to a polar effect caused by the gene deletion in *Δfmt.*

To qualitatively analyze GPL-2 or GPL-3 from WT S and *Δfmt* strains, an in-depth analysis was carried out. Partial purification of the GPL-2 and GPL-3 fractions was accomplished by liquid flash-chromatography on a silica gel column and monitored by TLC analysis **(Figure3B)**. Starting from the lower part of the TLC to the solvent front, two intense and sharp bands, tentatively identified as trehalose-containing lipids with unsaturated C_{16} and C_{18} fatty acids (data not shown) were observed along with multiple bands with higher R_f values, corresponding to GPL-3 and GPL-2, based on their respective mobilities. These fractions were next subjected to thorough structural analyses.

Fig. 3


A
Migration x 1
GPL-2
GPL-3
R
S
Δfmt
Δfmt C
Migration x 2
GPL-3
B
purified fractions
WT
WT GPL-2
WT GPL-3
Δfmt
Δfmt GPL-2
Δfmt GPL-3
C
WT GPL-2
1258.2
1286.3
Δfmt GPL-2
1244.2
1272.2
WT GPL-3
1404.3
1432.4
Δfmt GPL-3
1418.3
1390.3
Intensity %
m/z
D
WT GPL-2
methoxy Rha
Δfmt GPL-2
WT GPL-3
Δfmt GPL-3
13C ppm
1H ppm

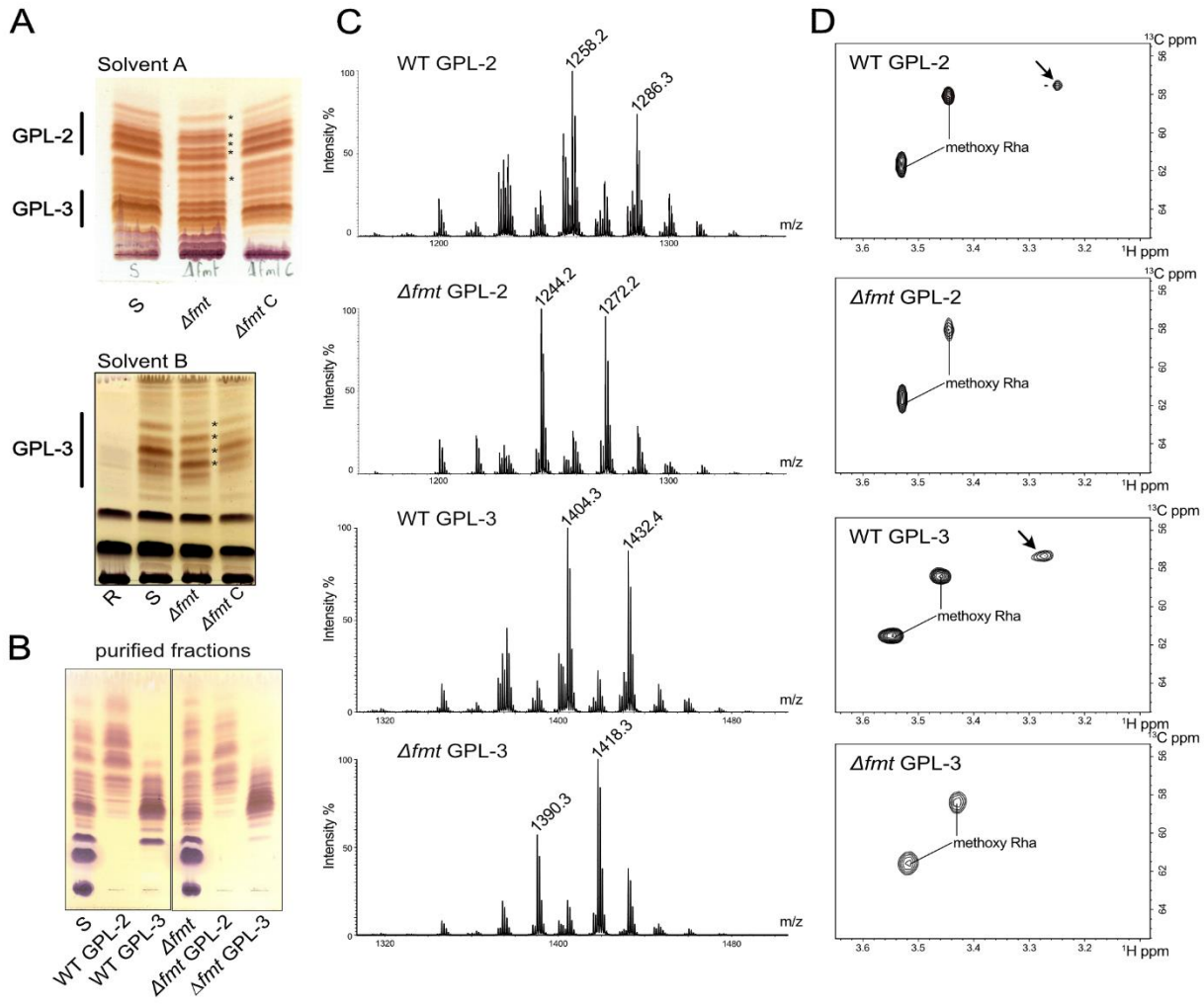


Figure3. Lossoffatty acidmethylation on GPLfrom *Δfmt*. (A) TLC analysis of the crude lipid fractions of the parental, mutant and complemented strains. GPL-2 and GPL-3 were separated using either solvent A ($CHCl_3$/MeOH 95:5, v/v) or solvent B ($CHCl_3$/MeOH 9:1, v/v) and revealed with orcinol staining after either one or two runs of migrations. GPL are expressed at similar levels in the parental (S), *Δfmt* and *ΔfmtC* strains but are absent in the R variant. Mobility of diglycosylated and triglycosylated GPL species differ in *Δfmt* as compared to the WT and *ΔfmtC* strains (upper panel). A second round of migration allowed to better visualize differences in the triglycosylated GPL profile as indicated by asterisks (lower panel). **(B)** Diglycosylated fractions (GPL-2) and triglycosylated fractions (GPL-3) were separated by flash chromatography and visualized on TLC. **(C)** MS spectra of GPL-2 and GPL-3. The two upper panels correspond to GPL-2 fractions while the two lower panels correspond to GPL-3 fractions. A decrease of 14 m.u. is observed between WT and *Δfmt* GPL fractions. **(D)** $^{1}H/^{13}C$ HSQC spectra of purified GPL fractions show methoxy functions substituting either the Rha or the lipid moiety. The methoxy group on the lipid moiety is exclusively found on WT GPL-2 and GPL-3 at δ 3.27/57.4 ppm, as indicated by an arrow.

Theglycanmoietyof GPLisnot altered in the *Δfmt* strain. The monosaccharide composition determined by GC/MS of WT GPL-2 showed the presence of 3,4 di-*O*-methyl Rha and 6-dTal in an approximate 1:1 ratio. Very similar results were obtained with the *Δfmt* GPL-2 fraction **(FigureS1)**.Both WT GPL-3 and *Δfmt* GPL-3 showed the presence of 3,4 di-*O*-methyl Rha, 6-dTal and Rha in approximate 1:1:1 ratio **(FigureS2)**.In addition, 3-*O*-methyl Rha was observed in much lower quantities on all chromatograms, suggesting the

presence of minor glycoforms. These results not only confirm that GPL-3 lipids correspond to extended GPL-2 lipids with an additional Rha residue but also imply that loss of *fmt* does not influence the monosaccharide composition of GPL. Mass spectrometry analysis of GPL-2 and GPL-3 from WT and *Δfmt* showed complex patterns of signals dominated by two major signals with 28 m.u. increments, tentatively attributed to two CH_2 groups **(Figure3C)**. The 146 Da increment between WT GPL-2 and WT GPL-3 major signals at *m/z* 1258/1286 and at *m/z* 1404/1432 was tentatively attributed to a deoxyhexose, confirming that WT GPL-3 corresponds to WT GPL-2 with an extra Rha residue, in agreement with GC/MS analysis and previous reports [26,31]. The fine sequence of individual WT GPL was further delineated by MS^2 fragmentation of the intense ions at *m/z* 1258 **(Figure4A)**and 1404 **(FigureS3A)**.Thepresence of terminal Me_2Rha and Ac_2dTal residues in WT GPL-2 was established owing to primary fragment ions at *m/z* 1084 and 1010, respectively, and confirmed by secondary ion products at *m/z* 836 and 454 (**Figure4A**). In contrast, WT GPL-3 was characterized by a terminal Rha and Ac_2dTal residues owing to primary fragment ions at *m/z* 1258 and 1156 and to a number of secondary ions products (**FigureS3A**),confirmingthat WT GPL-3 was further elongated by a single Rha residue on the Me_2Rha as compared to WT GPL-2. Similar results were obtained on 28 m.u. higher parent ions for GPL-2 and GPL-3 at *m/z* 1286 and 1432 corresponding to a longer lipid chain (data not shown). The presence of the two acetyl groups on dTal residues to form Ac_2dTal was confirmed by MS and MS/MS analyses of mild alkali treated GPL that led to an 84 m.u. decrease for the molecular ion as well as for all fragments containing the dTal residue. This is exemplified by the MS analysis of saponified WT GPL-3 that shows two signals at *m/z* 1319 and 1347 **(FigureS4A)**ascompared to ions at *m/z* 1404 and 1432 for the non-treated fraction **(Figure3C)**and confirmed by the corresponding MS/MS fragmentation patterns **(Figure S4B)**.

As for WT, the MS spectra from *Δfmt* GPL-2 and *Δfmt* GPL-3 showed complex patterns dominated by major ions with 146 m.u. increments at *m/z* 1244/1272 and *m/z* 1390/1418, respectively, exhibiting a 14 m.u. decrease compared to WT GPL-2 and WT GPL-3 **(Figure3C)**,tentativelyassigned to the loss of a methyl group on *Δfmt* GPL. MS^2 fragmentation patterns of the major GPL species **(Figure4B;FigureS3B)**and of their alkali-treated equivalents (data not shown) established that the glycan moieties of *Δfmt* GPL-2 and *Δfmt* GPL-3 were identical to those of WT GPL-2 and WT GPL-3. These results clearly indicate that deletion of *fmt* does not alter the glycan moiety of GPL.

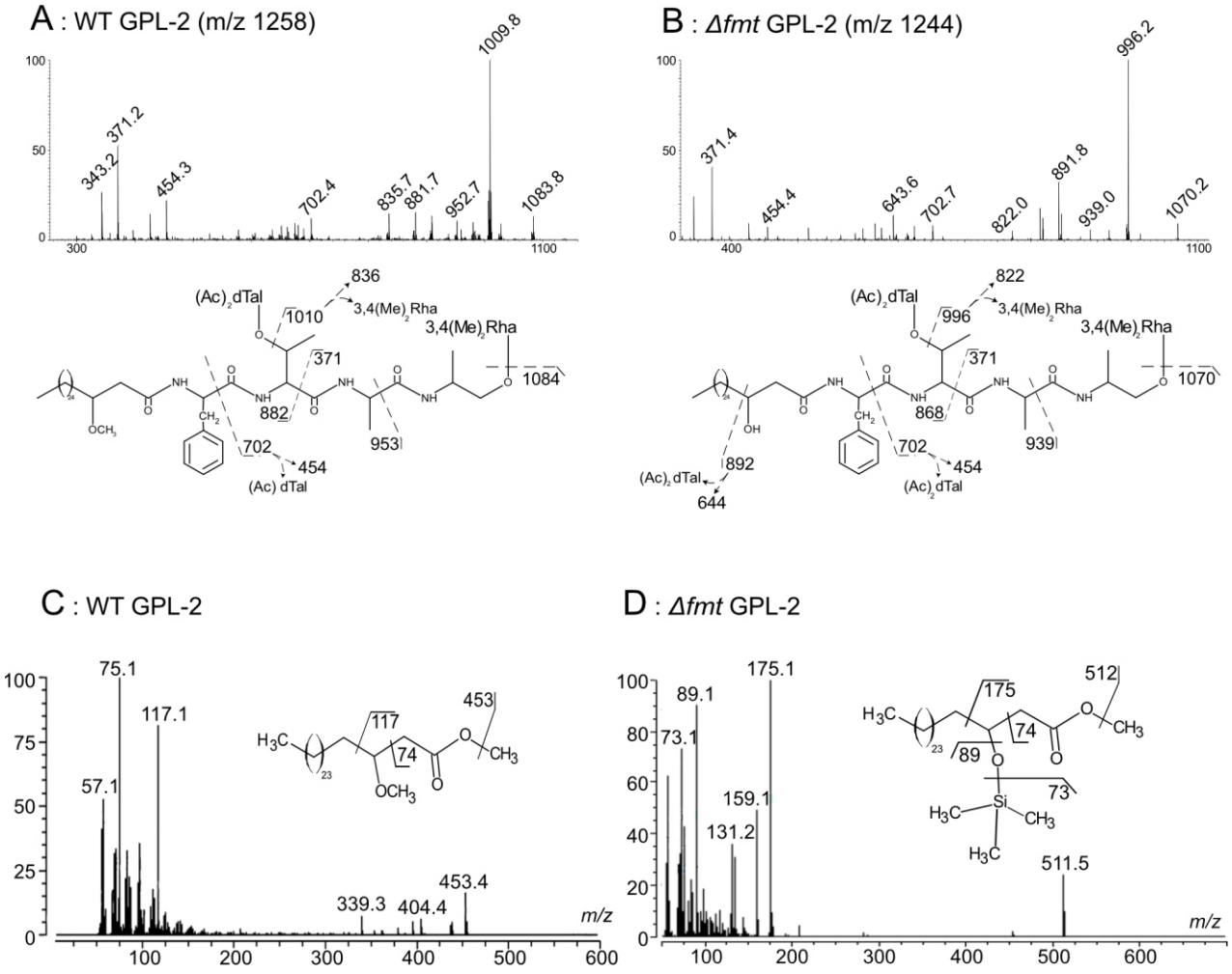


Figure4. StructuralanalysisofGPL-2. MS^2 fragmentation spectra of **(A)** the MS signal at *m/z* 1258 of WT GPL-2 and **(B)** at *m/z* 1244 of *Δfmt* GPL-2**.** EI-MS spectra of 3-methoxylated C_{28} FAME isolated from **(C)** WT GPL-2 and its **(D)** 3-trimethylsilylated C_{28} equivalent from *Δfmt* GPL-2.

The fatty acyl chain of GPL lacks a 3-methoxy function in *Δfmt*. NMR was subsequently used to unambiguously identify the origin of the 14 m.u. decrease in *Δfmt* samples. In particular, $^{1}H/^{13}C$ HSQC heteronuclear experiments confirmed that one signal in the region of *O*-CH_3 chemical shifts at δ 3.27/57.4 ppm ($^{1}H/^{13}C$) detected for WT GPL-2 and GPL-3 was absent in the corresponding *Δfmt* fractions **(Figure3D)**. This signal was clearly identified as a methoxy group associated with the fatty acid chain, according to its chemical shifts reported in *M. smegmatis*[39,40]. In contrast, the two less shielded signals which were similarly observed in all samples were assigned to the *O*-Me of Me_2Rha, in agreement with previous work [41]. This confirmed that *fmt* has no influence on the methylation of Rha residues. To further certify the loss of methoxylated fatty acids in *Δfmt*, the fatty acid composition of WT GPL and *Δfmt* GPL was investigated by GC/MS. As expected, a mixture of methoxylated fatty acids dominated by a C_{28} octacosanoyl form was identified in WT GPL [40]. Electronic-impact mass spectrometry (EI-MS) fragmentation **(Figure4C)** ofthe fatty acid methyl esters (FAME) highlighted the presence of a methoxy group in the 3-position owing to the presence of a specific fragment ion at *m/z* 117. Similar experiments on *Δfmt* GPL did not permit to identify

similar methoxylated fatty acids. To identify potential non-methoxylated fatty acids, the free hydroxyl groups of FAME extracted from *Δfmt* GPL were labelled by trimethylsilylation prior to GC/MS analysis. EI/MS of the reaction products from *Δfmt* GPL showed the presence of C_{28} octacosanoyl FAME hydroxylated in the 3-position owing to the intense fragment ion at *m/z* 175 **(Figure4D)**.Finally, exhaustive analysis of MS data of *Δfmt* GPL-3 revealed that MS^2 spectra of ions at *m/z* 1390 and 1418 generated a specific fragment ion at *m/z* 1038, resulting from the cleavage of the fatty acid C3-C4 bond next to the hydroxyl group that could be further fragmented by MS^3 **(FigureS5A)**.Thisspecific fragment ion was not detected in WT GPL, although an isobaric fragment ion resulting from the cleavage of glycosidic bonds was found by MS^3 analysis **(FigureS5B)**.These data support the view that WT GPL are substituted by 3-methoxylated fatty acids while *Δfmt* GPL are substituted by their hydroxylated equivalents.

The GXGmotif is essentialfor the methyltransferaseactivity of Fmt. To investigate whether *fmt* encodes a methyltransferase, the amino acid sequence of Fmt was aligned with other validated methyltransferases using CLUSTALW (https://www.genome.jp/tools-bin/clustalw). The aligned sequences include MT0146/CbiT from *Methanobacterium thermoautotrophicum* [42], the corynebacterial MtrP protein (NCg12764) assisting the transport of trehalose mycolates [43], the methyltransferase catalyzing the transfer of a methyl group onto the lipid moiety of phthiotriol and glycosylated phenolphthiotriol dimycocerosates in *M. tuberculosis*(Rv2952) [44] and the Fmt orthologue in *M. smegmatis*Mtf2 [45], for which active site residues have been identified. In particular, two glycine residues are conserved in *S*-adenosyl methionine (SAM)-dependent methyltransferase and known to be involved in binding to the methyl donor. These correspond to Gly86 and Gly88 in Fmt, which are also conserved in MT0146/CbiT, NCg12764, and Mtf2 **(Figure5A)**. However, only the second glycine residue is conserved in Rv2952 **(Figure5A)**. We reasoned that if replacement of these crucial residues by Ala would result in loss of methoxylation, trans-complementation of *Δfmt* with the Ala mutated *fmt* alleles would fail to restore the WT GPL pattern. Thus, *Δfmt* was transformed with constructs allowing expression of Fmt variants (HA-tagged) in which Gly86 and Gly88 were either individually or simultaneously substituted by Ala. Introducing pMV306-*fmt*-G86A, pMV306-*fmt*-G88A or pMV306-*fmt*-G86A/G88A into *Δfmt* resulted in similar expression levels of the different Fmt variants, as judged by probing the proteins using anti-HA antibodies **(Figure5B)**. The TLC profile of the extracted lipids clearly indicates that the introduction of the single or double mutations in *fmt* failed at functionally complementing the GPL profile **(Figure5C)**.

To support these results, GPL-2 and GPL-3 were purified from Δ*fmt* harboring either pMV306-*fmt*-G86A, pMV306-*fmt*-G88A or pMV306-*fmt*-G86A/G88A and analyzed by MS. As depicted in **FigureS6**,and in agreement with the TLC pattern, the MS profiles in these three strains remain identical to the one in *Δfmt*. In each of the three *Δfmt* strains expressing a catalytically-inactivated Fmt variant, GPL-3 fractions displayed a set of intense bands with lower R_f due to a more polar behavior **(FigureS6A,**black arrow**)** and identical

MS spectra to *Δfmt,* with two main ions at *m/z* 1390 and 1418 **(FigureS6C)**.MS^3 analysis of the fragment ion 1418 -> 1038 generated very similar fragmentation pattern to those observed for the *Δfmt* GPL-3 **(Figure S5A)**,strongly suggesting the presence of a β-hydroxyl group on the fatty acid **(FigureS7)**.In addition, the two major ions at *m/z* 1244 and 1272 were found in the GPL-2 fractions of all three strains **(FigureS6B)**. These results establish the absence of methoxylated GPL in *Δfmt* producing the catalytically-inactivated Fmt proteins. Collectively, this indicates that both Gly86 and Gly88 are required for restoring the GPL profile to WT level and confirm that Fmt is a canonical methyltransferase.

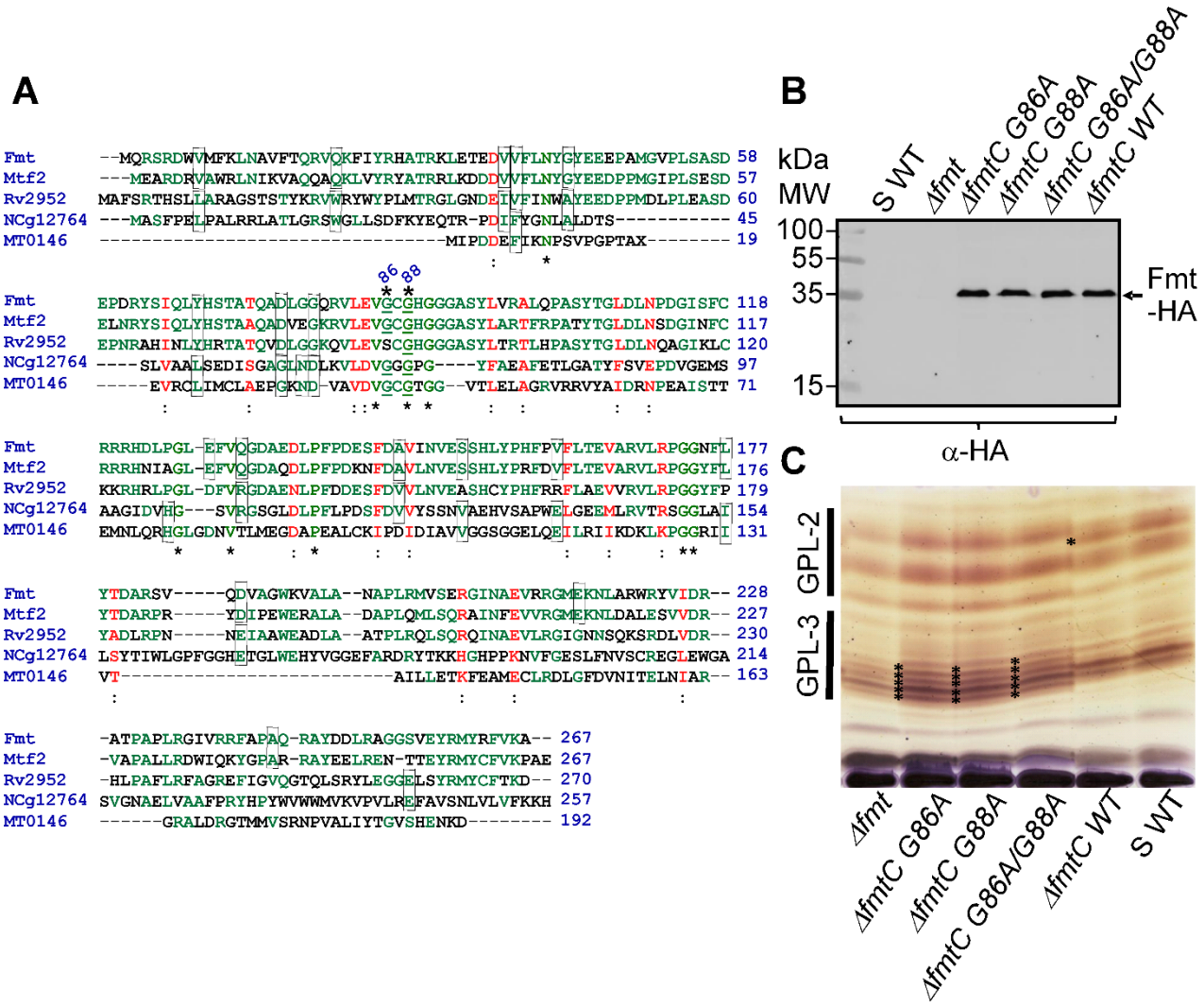


Figure5. The GXGmotif is essentialfor Fmt activity. (A) Multiple sequence alignment of Fmt with the *Methanobacteriumthermoautotrophicum*protein MT0146, the corynebacterial protein NCg12764, the *M. tuberculosis*protein Rv2952 and its orthologue in *M. smegmatis*(Mtf2) using CLUSTALW. Similar residues are shown in red and identical residues are in green. Wherever possible, the two types of identical amino acids were surrounded within the same vertical aligned column. Conserved glycine residues (Gly86 et Gly88) within the GXG motif and mutated in this study are numbered. **(B)**Western blot analysis of the *Δfmt* strain expressing either the Fmt(WT)-HA, Fmt(G86A)-HA, Fmt(G88A)-HA or Fmt(G86A/G88A)-HA fusion proteins under the control of the *fmt* native promotor. **(C)**TLC analysis of the crude lipid fractions of the parental, mutant and the various complemented strains. GPL were separated using $CHCl_3$/MeOH (9:1, v/v) and revealed after spraying with orcinol and charring. Black asterisks indicate modified GPL species.

Hydrophobicity of *M. abscessus* S is dependent on GPL methylation. Partitioning of mycobacterial cultures between hexadecane and an aqueous buffer has been used as a quantitative marker of hydrophobicity in *M. tuberculosis* evolution and pathogenicity, ranging from the hydrophilic environmental low-pathogenicity ancestors *Mycobacterium kansasii* and *Mycobacterium canettii* to highly hydrophobic virulent tubercle bacilli [46]. To address whether the methylation of GPL affects the surface hydrophobicity of *M. abscessus*, hexadecane-aqueous buffer partitioning was applied to the S, R, *Δfmt* and *ΔfmtC* strains. *M. kansasii* Hauduroy (ATCC 12478) was included as a highly hydrophilic control strain [46]. Consistent with previous findings, the S variant is more hydrophilic than the R variant while *M. smegmatis* mc²155 exhibited an intermediate phenotype [14]. Unexpectedly, *Δfmt*, while producing normal levels of GPL but lacking the methoxy on the fatty acid core, was significantly more hydrophilic than the S variant, almost reaching levels of *M. kansasii*, while complementation restored the wild-type hydrophobic properties of the S strain **(Figure 6A)**. This suggests that lack of *O*-methylation is responsible for the increased hydrophilicity of *Δfmt*.

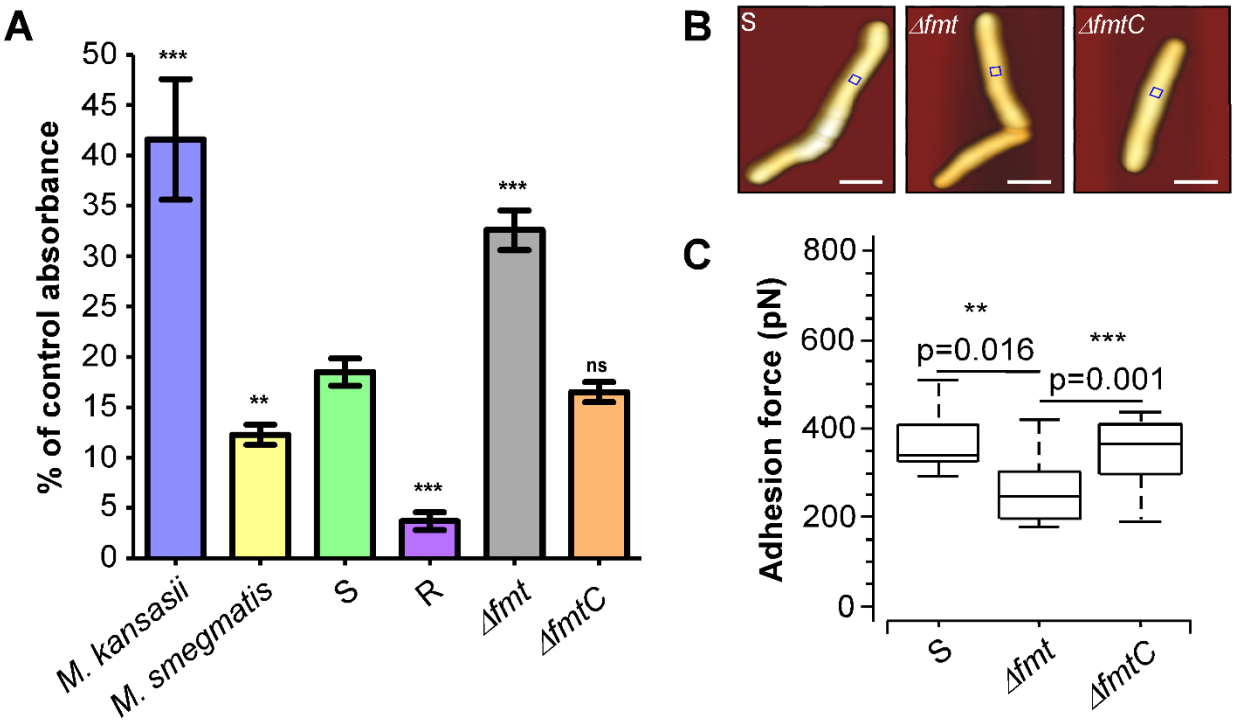


Figure 6. Decreased cell-surface hydrophobicity of *Δfmt*. (A) *M. abscessus* R is more hydrophobic than S, while the *Δfmt* mutant strain's hydrophilicity is significantly higher than S, as assessed by hexadecane partitioning and shown as hydrophobicity index (in percentage of aqueous phase OD prior to partitioning). Histograms and error bars are means ± SD of three independent experiments. Differences between means were analyzed for significance using a two-tailed Student's *t*-test employing Welch's correction for unequal variances. ns, non-significant, $^{**}p < 0.01$, $^{***}p < 0.001$. **(B)** 3D projections of height images obtained in quantitative imaging (QI) mode of *M. abscessus* S variant wild type (WT), *Δfmt* and *ΔfmtC* cells. The blue squares (250 nm × 250 nm) in each image indicate where adhesion force maps were collected for these representative cells, shown in **Fig. S8** (WT #2, *Δfmt* #5 and *ΔfmtC* #1). Scale bar = 1 µm. **(C)** Boxplots of adhesion forces for *M. abscessus* S wild type (WT), *Δfmt* and *ΔfmtC* cells. Whiskers indicate the range, the bottom and tops of the boxes, the 1st and 3rd quartiles, respectively, and the thick black bar in the middle

of the boxes represents the sample median. The different groups were compared using one-way Mann Whitney tests. n = 17 (WT), n = 18 (*Δfmt*), n = 11 (*ΔfmtC*).

Hydrophobic forces are weaker within hydrophobic nanodomains on single *Δfmt* cells. Atomic force microscopy (AFM) with chemically modified (hydrophobic) AFM tips was employed to assay at high spatial resolution the magnitudes, frequencies and surface distributions of hydrophobic adhesive forces on the surface of *Δfmt* [15,47]. Low resolution topographic images (5 µm × 5 µm) were first recorded and then high-resolution adhesion maps were obtained (250 nm × 250 nm, **Figure6B**,blue squares) on top of the bacteria to avoid edge artefacts. In line with previous work [15], heterogenous distributions of hydrophobic adhesive forces were observed, *i.e.* hydrophobic nanodomains surrounded by hydrophilic areas on the surface of *M. abscessus* S cells **(FigureS8A)** as well as on *Δfmt* **(FigureS8B)** and *ΔfmtC* cells **(FigureS8C)**.The frequencies of detectable hydrophobic forces on *Δfmt* (24 ± 3%, n = 19 cells) did not differ significantly from those on WT bacteria (16 ± 4%) or *ΔfmtC* cells (23 ± 4%), indicating that abrogation of GPL methylation by Fmt does not change the proportion of hydrophobic *versus* hydrophilic areas on the cell surfaces.

To test whether changes in the methylation pattern of GPL affect hydrophobic adhesion within nanodomains, we assessed the average magnitudes of adhesion forces within these domains only (*i.e.* excluding non-adhesive events). The average adhesion force within hydrophobic nanodomains on *Δfmt* was 281 ± 28 pN (mean ± SEM, n = 17, 481 curves from 18 cells), while in WT cells it was 375 ± 37 pN (mean ± SEM, n = 11, 476 curves from 17 cells), a moderate (~25%) but significant difference (p=0.016). In *ΔfmtC*, adhesive forces with hydrophobic nanodomains were restored to WT levels (371 ± 20 pN, mean ± SEM, n = 10, 495 curves from 11 cells). These results shown as boxplots in **Figure6C** indicate that alteration of the GPL methoxylation pattern correlates with changes in the surface hydrophobic properties of *M. abscessus* S cells. It also supports the view that more hydrophobic classes of GPL define the hydrophobic nanodomains exhibited on the surfaces of S variant cells (at least in contribution with other apolar lipid classes). However, that the ratio of hydrophobic to hydrophilic area did not change in *Δfmt* indicates that methoxylation is not the only determinant of GPL hydrophobicity and that other chemical differences between the class of GPL that define the hydrophilic nanodomains and those that define the hydrophobic ones likely exist.

Reduced invasion of THP-1 macrophages by *Δfmt*. Because cell surface hydrophobicity plays important roles such as favoring cell-surface interactions with host tissues [48], the first step leading to infection, we next evaluated the possible impact of GPL methylation on adhesion and invasion of *M. abscessus* by human THP-1 macrophages. Cells were infected with the tdTomato-expressing *M. abscessus* S, *Δfmt* or *ΔfmtC* strains for 3 hrs at a multiplicity of infection (MOI) of 2:1, prior to assessing phagocytosis and rate of replication over time. Following infection, macrophages were treated with 250 µg/mL amikacin for 2 hrs and then amikacin was maintained at 50 µg/mL to prevent extracellular bacterial growth. At 0, 1, 3 and 5 dpi, macrophages were lysed and plated to determine the intracellular bacterial burden. At day 0 (3 hrs

post-infection), the invasion rate of *Δfmt* was reduced by almost 50% as compared with the control strain and this effect was rescued with *ΔfmtC* **(Figure7A)**.This defect was maintained at 1, 3- and 5-days post-infection **(Figure7A)**.In parallel, at 0, 1, 3- and 5-days post-infection, macrophages were stained with DAPI and observed under an epifluorescence microscope. Microscopic observations and quantification of the infected cells revealed a pronounced reduction in the number of infected THP-1 cells infected with *Δfmt* just after phagocytosis. Again, this effect was maintained over time as compared to the macrophages infected with either the WT progenitor or *ΔfmtC*strains **(Figures7Band7C)**.Incontrast, no major effect on bacilli multiplication was observed within the cells, suggesting that *fmt* is important for bacterial invasion rather than for intracellular survival.

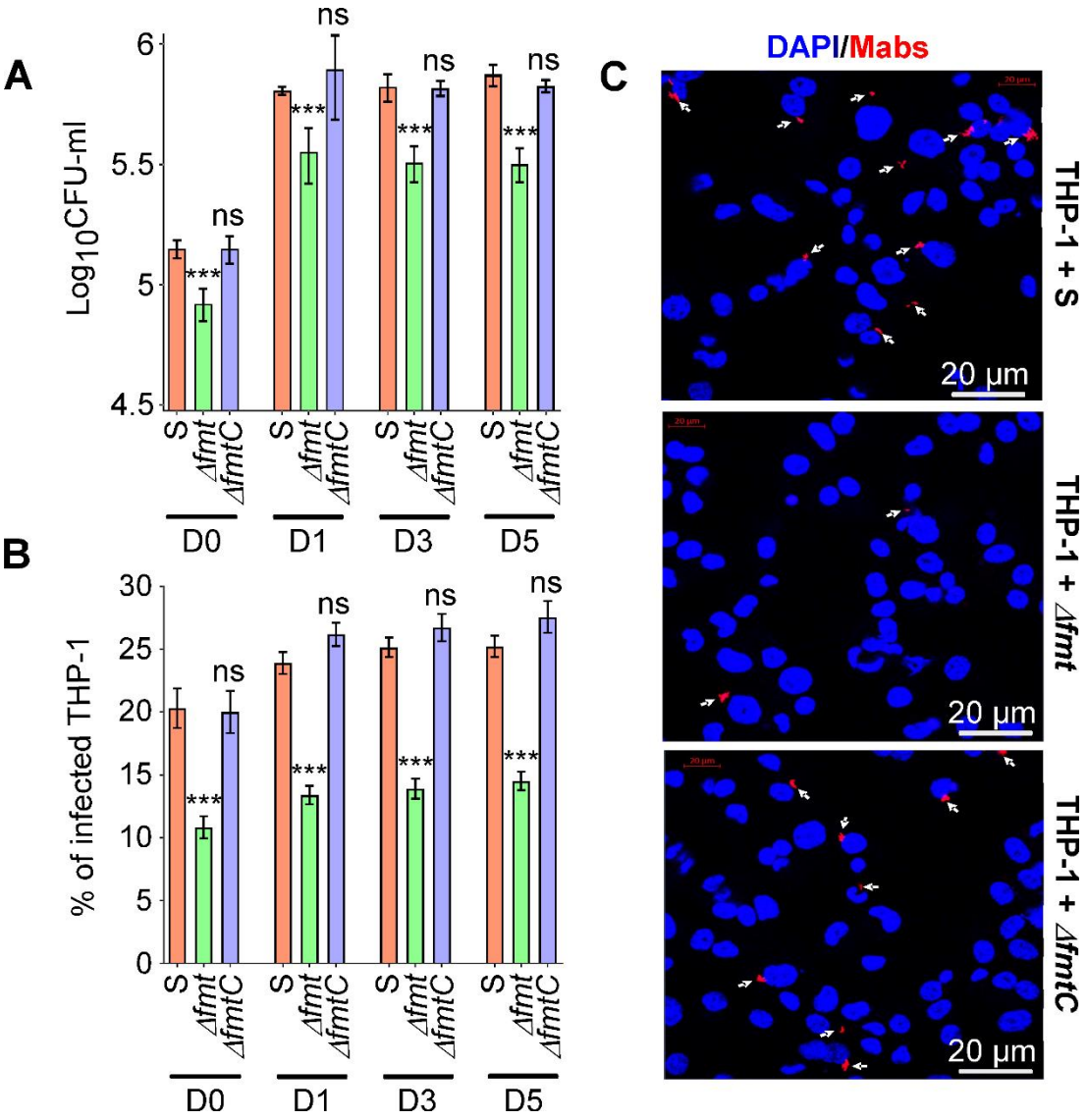


Figure 7. Fmt participates in invasion of human macrophages by *M. abscessus*.Macrophages were infected with the *M. abscessus*S,*Δfmt*and *ΔfmtC*strains expressing tdTomato (MOI of 2:1). **(A)** CFU were determined at days 0, 1, 3 and 5 post-infection. Data are mean values ± SD for three independent

experiments. One-tailed Mann Whitney test. ns, non-significant, $^{***}p < 0.001$. **(B)** Percentage of infected THP-1 macrophages at days 0, 1, 3 and 5 post-infection. Data are mean values ± SD for three independent experiments. One-tailed Mann Whitney test. ns, non-significant, $^{***}p < 0.001$. **(C)** Three immuno-fluorescent fields were taken after one-day post-infection at a 40X magnification (using a confocal microscope), showing the nuclei of macrophages infected with the various strains. The nuclei are shown in blue. White arrows indicate mycobacteria within macrophages. Scale bars represent 20 µm. Experiments were done three times independently.

DISCUSSION

This work provides a first glimpse into the structure-function relationship of GPL fatty acid methoxylation. It unambiguously demonstrates that *fmt* encodes a *O*-methyltransferase that catalyzes the transfer of a methyl to the 3-hydroxyl group of the C_{24}-C_{33} fatty acyl chain, with which the amino group of the peptidic core is esterified: i) unmarked deletion of *fmt* leads to the production of GPL carrying a fatty acyl chain lacking a 3-methoxy group while the glycan moiety remains unaltered; ii) substitution of the Gly86 and/or Gly88 residues of the conserved GXG motif, known to bind to the methyl donor in other *O*-methyltransferases [43], resulted in the loss of *O*-methylation of the lipid domain of GPL; iii) Fmt is the orthologue of Mtf2 which methylates GPL in *M. smegmatis* [45]. However, *M. smegmatis* being a non-pathogenic species, the impact of lipid *O*-methylation in the physiopathology of infection cannot be investigated. In addition, we procure strong evidence that *O*-methylation of the GPL fatty acid increases cell surface hydrophobicity of *M. abscessus* and increases internalization of the bacilli by macrophages. By substituting the endogenous promoter of the *mmpS4-mmpL4a-mmpL4b* locus involved in the transport of GPL with a leaky acetamidase promoter from *M. smegmatis* in *M. abscessus* S, we previously reported a strain producing low-GPL levels with a rough appearance, similar to R strains [14]. This rough low-GPL producing mutant was less hydrophilic than the S variant but more hydrophilic than the R variant, indicating that lowering the production of GPL leads to increased hydrophobicity. Herein, we show that *O*-methylation of the lipid core of GPL represents another way for *M. abscessus* to regulate its surface hydrophobicity properties. However, these alterations in the GPL profile were not associated to changes in the susceptibility to cefoxitin, imipenem, bedaquiline, amikacin or rifabutin (data not shown), indicating methoxylation of the fatty acyl chain does not contribute to the envelope permeability and drug sensitivity of *M. abscessus*.

Earlier comparative studies between the S and R morphotypes of *M. abscessus* underscored the positive correlation between low GPL production and increased hydrophobicity and virulence in animal models [14]. These features are reminiscent of the role that hydrophobicity has played in *M. tuberculosis* evolution towards greater pathogenicity, where more modern pathogenic strains exhibit a higher proportion of less polar lipids in the outer membrane that corresponds with greater surface hydrophobicity [46]. Importantly, increased hydrophobicity may enhance the capability for aerosol transmission, affecting virulence and pathogenicity [46,49]. Conversely, *M. kansasii*, which is much less pathogenic than *M.*

tuberculosis and less virulent than *M. abscessus* in the zebrafish model of infection [50], produces large quantities of hydrophilic lipids and is characterized by mycomembrane lipids that are less apolar than those found in *M. tuberculosis* [46]. Our results indicate that *M. abscessus* S is less hydrophilic than *M. kansasii* but more than the R morphotype. Remarkably, that *⊿fmt* is more hydrophilic than the S parental strain, indicates that a simple change in the methoxylation pattern of GPL can lead to pronounced phenotypic changes in surface hydrophobicity. However, keeping in consideration the wide heterogeneity of the microbial population in planktonic cultures, comprising a mixture of bacteria with various degrees of hydrophilicity, we aimed at imaging and quantifying surface hydrophobicity at the single cell level using AFM. As noticed previously, the *M. abscessus* S cell surface is characterized by hydrophilic nanodomains, the latter being completely absent in R variant cells that express homogenous surface hydrophobicity [15], like in other GPL-defective mycobacteria [47,51,52]. While the homogenous hydrophobicity on the R variant cell surface is easily explained by the exposure of more hydrophobic lipid classes, like mycolic acids, in the absence of GPL, two explanations for the presence of both hydrophobic and hydrophilic nanodomains on S variant cells were posited: i) mycolic acids that are hydrophobic (most likely along with other hydrophobic lipids) define the hydrophobic nanodomains, while GPL, which expose relatively hydrophilic head groups, define the hydrophilic nanodomains; ii) more hydrophobic GPL classes define the hydrophobic nanodomains, while hydrophilic nanodomains are defined by more hydrophilic classes [15]. That the average adhesion force within hydrophobic nanodomains on *Δfmt* was lower than in the parental and *ΔfmtC* strains supports the latter hypothesis. It is possible that more apolar GPL-2 define the hydrophobic nanodomains and that a loss of fatty acid methoxylation leads to decreased hydrophobic adhesive forces measured specifically within hydrophobic nanodomains of *Δfmt*. Ensuing studies may explore the contribution of GPL glycosylation profile on surface hydrophobicity.

The fatty acid core of GPL is thought to interact with other lipids present within the outer leaflet of the outer membrane, implying that the methoxy substituent is likely to be buried within the membrane and, presumably, not exposed on the bacterial surface. One cannot, however, exclude that the overall reduced hydrophobicity of *⊿fmt* GPL relies simply on direct consequence of the lack of the methoxy group. It is possible that this methoxy group is important for the GPL to interact with other membrane lipids. The loss of the methyl group in *⊿fmt* may, therefore, result in "abnormal" interactions with lipids, favoring the exposition of other polar lipids at the bacterial surface, thus influencing the overall cell surface chemical composition and hydrophobic properties. Interestingly, methoxylation has been ascribed to the phenolglycolipids (PGL) and phthiocerol dimycocerosate (PDIM) in *M. tuberculosis*, two other well-described, complex, virulence lipids [53]. In these lipids, the transfer of a methyl group onto the free hydroxyl group of the lipid domain of PDIM and PGL is catalyzed by the methyltransferase Rv2952 [44], whose gene is found within the PDIM/PGL locus, similarly to how *fmt* is located within the GPL locus. Multisequence alignments indicated an overall conservation between Fmt and Rv2952, albeit the GXG motif is not fully

conserved in Rv2952 as the first Gly residue is substituted by a Ser. Based on our findings with GPL, it would be interesting to investigate whether the lack of lipid methoxylation in PDIM and PGL translates into altered biological properties and influences pathogenesis of *M. tuberculosis*.

The outermost layer of *M. abscessus* characterizes the interface between bacilli and their host macrophages. It is now clear that the hydrophobic effect is a factor in adhesion of numerous pathogens [48]. Mycobacterial exposed adhesins are recognized by specific cell surface receptors that allow bacteria to reach and invade their cellular targets. A few adhesins have been shown to play a role in the interaction between mycobacteria and host cells, but their surface position and their approachability to their cognate receptors remain poorly understood. This work shows that the decreased cell surface hydrophobicity of *Δfmt* correlates with reduced invasion of macrophages by *M. abscessus*, while the intracellular growth rate of the mutant was not altered as compared to the one of the WT progenitor. These findings are consistent with previous studies demonstrating that internalization of *M. smegmatis* by human macrophages is more efficient in a mutant lacking GPL (*mps*-inactivated) than in a control strain producing GPL [29,54]. However, whether the reduced adhesion and internalization of *Δfmt* in macrophages is a direct or indirect effect of the lack of GPL methoxylation deserves future investigations. One possibility could be that GPL are direct ligands of macrophage receptors and that the methoxy of the GPL fatty acid chain participates in the ligand-receptor recognition. Improved adherence to target cells is directly correlated with increased bacterial surface hydrophobicity in Neisseria and Streptococci [55,56]. In another study, methylation was found to increase hydrophobicity of the outer surface of the flagellar filament in *Salmonella* and, consequently, enhancing bacterial adhesion and host cell invasion [57]. At this stage, future work is required to understand how bacterial surface hydrophobicity can alter the adhesion strength between bacteria and host cells.

CONCLUSION

This study underscores the contribution of GPL methoxylation in the biological functions of this important class of lipids. Because GPL are present in numerous NTM, these results suggest that the activity of Fmt participates in the surface properties and internalization of other NTM, highlighting possible therapeutic interventions against these opportunistic mycobacterial pathogens. Future studies should enlighten the contribution of the other parts of this complex lipid molecule with respect to their biological functions.

METHODS

Mycobacterial strains, growth conditions and reagents. All bacterial strains are listed in **Table S2.** Rough (R) and smooth (S) variants of *M. abscessus* CIP104536^T were usually grown in Middlebrook 7H9 broth (BD Difco) supplemented with 0.05% Tween 80 and 10% oleic acid, albumin, dextrose, catalase (OADC enrichment; BD Difco) (7H9$^{T/OADC}$) at 37°C in the presence of antibiotics, when required. Electro-competent mycobacteria were transformed using a Bio-Rad Gene pulser (25 µF, 2500 V, 800 ohms). For bacterial

selection, media were supplemented either with 1 mg/ml hygromycin for strains carrying pTEC27 (Addgene, plasmid 30182), allowing tdTomato expression or with 250 µg/ml kanamycin when harboring the pMV306 derivatives. On plates, colonies were selected either on Middlebrook 7H10 agar (BD Difco) supplemented with 10% OADC enrichment ($7H10^{OADC}$) or on LB agar. Antibiotics were purchased from Sigma-Aldrich.

***In vitro* growth, colonymorphologyassaysandsedimentation.** Growth was examined by inoculating the mid-log phase cultures into fresh 7H9 at an OD_{600} of 0.05. Cultures were incubated at 37°C with shaking and OD_{600} was monitored for 150 hrs using a Synergy H1 hybrid reader (BioTek). To evaluate colony morphology, bacteria from a log phase bacterial culture (OD_{600}=1) were resuspended in phosphate buffered saline (PBS) and 2 µl of cell suspension were spotted onto $7H10^{OADC}$ or LB agar. Plates were then incubated for 4-5 days at 37°C and colonies were imaged using a Zeiss microscope equipped with a Zeiss Plan Neo Fluor Z 1×/0.25 FWD objective. Images were acquired with an Axiocam503 monochrome (Zeiss) camera and processed using ZEN 2 (blue edition). For sedimentation experiments, cultures were diluted to an OD_{600} of 1 and 2 ml of culture were transferred to a glass tube. Images were taken after 5 min.

Pellicleassays. Pellicleformation was assessed by inoculating 10 µl of mid-log phase mycobacterial cells (OD_{600}, 0.8–1.0) onto medium consisting of M63 supplemented with 10% of glucose, 1 mM of $CaCl_2$ and 1 mM of $MgSO_4$ and incubated without agitation at 30°C for 5 days.

Hexadecanepartitioning. Exponentially growing bacteria were washed twice in PUM buffer (100 mM K_2HPO_4, 54 mM KH_2PO_4, 30 mM urea, 0.8 mM $MgCl_2$) and suspended to an OD_{600} of 0.7. Aliquots (3 ml) were transferred to glass tubes and hexadecane (2.4 ml) added. After a brief mixing, samples were incubated for 8 min at 37°C and phase separation allowed to occur at 22°C for 15 min. The hydrophobicity index was defined as aqueous phase OD_{600}, expressed as a percentage of that of the bacterial suspension in PUM buffer alone [46].

Constructionof*Δfmt*. The suicide vector pUX1-*katG* [37] was used to generate pUX1-*katG-fmt* **(TableS3)**, which was subsequently used to generate an unmarked deletion mutant in the S variant of *M. abscessus.* Briefly, the left and right arms (LA and RA, respectively) were PCR-amplified using genomic DNA, Q5 polymerase (New England Biolabs) as well as primers 1 and 2 (LA) and primers 3 and 4 (RA) **(TableS1)**.The purified LA and RA amplicons were restricted with PacI/MfeI and MfeI/NheI, respectively, and ligated to the PacI-NheI-linearized pUX1-*katG*, yielding pUX1-*katG-fmt*, designed to delete 696 bp (87%) of the *fmt* gene. Electrocompetent *M. abscessus* was transformed with pUX1-*katG-fmt.* The selection of bacteria

having undergone the first homologous recombination event was done by visual screening of red fluorescent colonies on 7H10^{OADC} supplemented with 250 µg/ml kanamycin. After subculturing the culture overnight in 7H9$^{T/OADC}$ in the absence of kanamycin, bacterial suspensions were serially diluted and plated onto 7H10^{OADC} with 50 µg/ml isoniazid (INH) to select for INH-resistant, Kan-sensitive and non-fluorescent colonies. The DNA junctions were subsequently PCR-sequenced to confirm the proper *Δfmt* genotype.

Complementationconstructs. Plasmidsforcomplementation were generated by PCR amplification of *fmt* in fusion with an HA tag sequence under the control of its endogenous promoter region (175 bp) using genomic DNA and the forward primer 9 (containing a KpnI site) and reverse primer 10 (containing an NcoI site as well as the HA-coding sequence) and subsequent ligation into the integrative vector pMV306 cut with KpnI and NcoI, resulting in pMV306-*fmt.* Mutating the Gly86 or/and Gly88 present in the GXG motif in Fmt in pMV306-*fmt* was achieved by site-directed mutagenesis using the primers listed in **TableS1**and the QuikChangeV Site-Directed Mutagenesis Kit (Agilent), according to manufacturer's instructions, yielding pMV306-*fmt*-G86A, pMV306-*fmt*-G88A and pMV306-*fmt*-G86A/G88A. All constructs were verified by DNA sequencing and introduced into *Δfmt*.

Westernblotting. Bacteria were harvested, resuspended in PBS, and disrupted by bead beating using 1-mm diameter glass beads. Protein concentration was assessed using the BCA Protein Assay Reagent kit (Pierce), according to the manufacturer's instructions. Equal amounts of proteins (50 µg) were separated by SDS/PAGE and transferred to a nitrocellulose membrane. For detection of Fmt-HA, Fmt-G86A-HA, Fmt-G88A-HA and Fmt-G86A/G88A-HA, membranes were probed for 1 hr with rat anti-HA (dilution 1:2000; Sigma). The KasA protein, used as an internal loading control, was revealed using rat anti-KasA antibodies (dilution 1:2000) [58]. After washing, membranes were incubated for 45 min with goat anti-rat antibody conjugated to HRP (dilution 1:5000; Abcam). The signal was revealed using the ChemiDoc MP system for imaging and analyzing gels (Bio-Rad laboratories).

AFM tip functionalization, sample preparation, parameters and analysis. Gold OMCL-TR400PB-1 AFM probes (Olympus) were functionalized by immersing them in a 1 mM ethanolic solution of 1-dodecanothiol overnight and used fresh. Bacteria were first cultured in 7H9 broth supplemented with 0.2% (w/v) glucose, 0.2% glycerol and 0.025% tyloxapol. At middle exponential phase (OD_{600}~ 1.0), bacteria were pelleted by centrifugation, re-suspended in medium lacking tyloxapol, de-aggregated via passing them 10× through a 26 GA syringe needle and finally passing them through a 5 µm PVDF syringe filter (Merck) resulting in a suspension of single bacterial cells. This suspension was seeded into a 35 mm hydrophobic (untreated) microscopy dish (iBidi) and bacteria were left to adhere to the hydrophobic surface overnight at 30°C. Just prior to AFM measurements, the bacteria were washed several times with deionized water and all AFM

experiments were performed in ultrapure deionized water. Topographic images were recorded in quantitative imaging (QI) mode using a JPK NanoWizard 4 AFM, bare MSCT probes (k = 0.02) and the following parameters: an applied force of 250 pN, 25 µm/s approach and retraction speeds, a z-length of 500 nm, a map size of 5 × 5 µm and a resolution of 128 × 128 pixels. Adhesion maps were also recorded in QI mode but using hydrophobic probes (k = 0.02) and the following parameters: 25 µm/s approach and retraction speeds, an applied force of 500 pN, a z-length of 500 nm, a map size of 250 nm × 250 nm and a resolution of 64 × 64 pixels. Owing to tip-geometry induced artefacts that occur on the steep edges of the cells, care was taken to record adhesion maps right on top of the bacteria (250 nm × 250 nm areas). The spring constants of AFM probes were determined empirically using the method of Hutter and Bechhoefer [59]. At least two different tips were used to collect data for each sample and the AFM experiment was repeated on one occasion. Force-distance curves obtained in QI mode were treated using the JPK Data Processing software (V.6.1.125).

Macrophageinfectionassays. THP-1macrophages were grown, infected with the *M. abscessus*S,*Δfmt* and *ΔfmtC*strains and processed as reported earlier [60]. At various time points (0, 1, 3- and 5-days post-infection), macrophages were washed three times with PBS and lysed with 100 µl of 1% Triton X100. Cell lysis was stopped by adding 900 µl PBS and serial dilutions were plated to monitor the intracellular bacterial counts. Colonies were counted after 5 days of incubation at 37°C. Microscopy-based infectivity assays were performed as reported previously [60]. Briefly, THP-1 cells were cultivated on coverslips in 24-well plates at a density of 10^5 cells/well. The following day, cells were infected with tdTomato expressing *M. abscessus* (MOI = 2). At various time points after infection, cells were fixed with 4% paraformaldehyde in PBS for 20 min, stained with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI) for 5 min, washed, mounted onto microscope slides using Immu-Mount (Calbiochem) and examined using a confocal microscope (63x objective) (Zeiss LSM880). Images were acquired and captured on a Zeiss Axio-imager confocal microscope equipped with a 63x oil objective and a camera and processed using Zeiss Axiovision software. Quantification and scoring of the numbers of bacilli present within macrophages were performed within focus using ImageJ. Equal parameters for the capture and scoring of images were consistently applied to all samples. For each condition, ~1000 infected macrophages were analyzed.

GPLextraction. Bacteria grown on $7H10^{OADC}$ agar plates without detergent were collected and GPL were extracted from the polar lipid fraction, first with chloroform/methanol/0.3%NaCl (9:10:3, v/v/v) and then by chloroform/methanol/0.3%NaCl (5:10:4, v/v/v). The combined solvent extracts were then mixed for 5 min with chloroform and 0.3%NaCl (1:1, v/v) and centrifuged at 3,000 *g* for 5 min to separate the lower-organic phase from the aqueous phase. The upper aqueous layer was discarded and the lower-organic

phase was evaporated under a stream of nitrogen and resuspended in chloroform/methanol (2:1, v/v). Polar lipids were then subjected to TLC analysis using Silica gel 60 F_{254} plates (Merck). GPL were separated using chloroform/methanol/water (90:10:1, v/v/v) or chloroform/methanol (9:1 or 95:5, v/v) and sprayed with orcinol/sulphuric acid vapor prior to revelation by charring.

Glycolipidpurification.1-TLC. Forpreparative TLC, 150 µl of glycolipid sample was spotted on a 150 mm band on a 60 µm silica gel plate with a glass back (20 cm x 20 cm) and migrated in a solution of chloroform/methanol/water (90:10:1, v/v/v). Plate sides were sprayed with orcinol in 20 % sulfuric acid and charred to reveal glycolipids while the plate center was reversibly colored with iodide vapor. **2-Flash chromatography.**Glycolipidsextracts were impregnated with silica gel overnight. After an equilibration step, impregnated silica was added on top of a low-pressure silica gel column 15 µm, 4 g (INTERCHIM, France) and a linear gradient started on Puriflash (INTERCHIM, France) from 100 % chloroform to reach 85 % chloroform - 15 % methanol at 20 min. This ratio was maintained for 10 min. 6 ml fractions were collected during the gradient phase and the presence of glycolipids was checked by TLC. Fractions containing GPL were pooled, dried and dissolved in chloroform/methanol (2:1, v/v).

GPLsaponificationand methanolysis.10 µl of purified GPL was dried under nitrogen, 200 µl sodium hydroxide 0.1 M in chloroform/methanol (2:1, v/v) was added and heated for 2 hrs at 37°C. After the reaction, 1 ml chloroform and 1 ml water were added. The mixture was vortexed for 1 min then centrifuged for 30 sec. The lower chloroform phase was dried under nitrogen and dissolved in 150 µl chloroform/methanol (2:1, v/v). For methanolysis, 10 µl of purified GPL were dried under nitrogen and desiccated overnight, then 200 µl of methanol/HCl 0.5 N was added and heated for overnight at 100°C. After the reaction, 1 ml chloroform and 1 ml water were added. The mixture was vortexed for 1 min, centrifuged for 30 sec and the lower chloroform phase was dried under nitrogen and dissolved in 250 µl heptane.

Itol-acetatederivatives. For hydrolysis step, 1 µg mesoinositol, 20 µl GPL fraction and 1 ml TFA 3M were mixed then heated 4 hrs at 80°C, dried and desiccated overnight. The reduction step was conducted for 4 hrs at room temperature in 500 µL $NaBH_4$ 10 mg/ml in 2M NH_4. The reaction was stopped with concentrated glacial acetic acid. Samples were dried at 55°C under a nitrogen stream by co-distillation with methanol/acetic acid three times, desiccated overnight and incubated in 500 µl anhydride acetic 4 hrs at 80°C. The reaction products were extracted several times with chloroform/water. The chloroform-rich phase was then filtered, dried and dissolved in 100 µl chloroform. Similar experiments were conducted using Rha or 6-dTal standards (CARBOSYNTH, UK).

MALDI-TOF mass spectrometry. Before spotting 5µl on the MALDI plate with a glass capillary tube, 5 µl of 10 mg/ml dihydroxybenzoïc acid (DHB) in chloroform/methanol (1:2, v/v) were mixed with 5 µl of the sample extract in chloroform/methanol (2:1, v/v). MS and MS^n spectra were acquired on an Axima Resonance (SHIMADZU, Kyoto, Japan) in reflectron mode. For MS^2 experiments, collision energy was tuned from 300 eV to 600 eV. For MS^3, ion selection occurred with standard resolution with the same collision energy.

GC-MS analysis. 1-Itol-acetates. 1µl of itol-acetate derivatives was injected in splitless mode on a Solgel 1 MS 30 m x 0.25 mm x 0.25 µm capillary column with the following gradient temperature: 120°C to 230°C at 3°C/min, then to 270°C at 10°C/min. Compounds were detected after electronic impact at 70 eV on a HP-7820 gas chromatograph coupled to a 5976 single quad (Agilent Technologies, Santa Clara, US) in full scan mode from 50 Da to 500 Da. **2-FAME and TMS derivatives.** After methanolysis, released FAME were extracted in heptane and analyzed on a TRACE 1300 gas chromatograph coupled to an ISQ single quad (THERMO FISHER SCIENTIFIC, San Jose, US). 1 µl was injected in splitless mode on a SLB-5 MS 30 m x 0.25 mm x 0.25 µm capillary column with the following gradient temperature: 50°C to 140°C at 10°C/min, then to 310°C at 15°C/min. After electronic impact at 70 eV, spectra were acquired simultaneously in selected ion monitoring at 74 and 87 Da, and in full scan mode from 50 Da to 700 Da. Trimethylsilyl-derivatives of FAMES were produced by adding 100 µl BSTFA and 100 µl pyridine. After 2 hrs at room temperature, 300 µl heptane was added and processed using the same conditions as for FAME. A methyl 3-hydroxy octadecanoate (MATREYA, US) was also submitted to TMS derivation to check the fragmentation model.

Nuclear Magnetic Resonance. Flash purified GPL were dried and dissolved in a mixture of $CDCl_3/CD_3OD$ (2:1, v/v) with 0.03 % trimethylsilane (EURISOTOP, France) three times then dissolved in a final volume of 200 µl. The sample was then introduced into a 3 mm glass tube (SHIGEMI, Allison Park, PA, US). A TBI probe was used to observe ^1H and ^{13}C nuclei at 293°K on an AVANCE II system (BRUKER BIOSPIN GmbH, Germany). Impulsion sequences used for homonuclear and heteronuclear experiments were from the manufacturer. After acquisition, phase correction and calibration on trimethylsilane signals were performed for $\delta\ ^1H$ and $\delta\ ^{13}C$. For HSQC, calibration was adjusted on the methanol signal.

Statistical analyses. Statistical analyses were performed on Prism 5.0 (Graphpad) or in R studio (for AFM experiments) and detailed for each figure legend. *p<0.05; **p<0.01; ***p<0.001.

Author contributions

WD conceived and conducted experiments, analyzed the data and wrote the manuscript. LDL, AV and JK conducted experiments. LDL, AV, YFD, YG and LK analyzed the data and participated in writing the manuscript. LK conceived the idea of the project and wrote the manuscript.

Conflictofinterest: Thefunders had no role in study design, data collection, interpretation, or the decision to submit the work for publication. The authors have no conflict of interest to declare.

SUPPORTINGINFORMATION

TableS1. List of primers used in this study.
TableS2. List of bacterial strains used in this study.
TableS3. List of plasmids used in this study.
FigureS1. Monosaccharide analysis of purified diglycosylated GPL.
FigureS2. Monosaccharide analysis of purified triglycosylated GPL
FigureS3. MS/MS sequencing of GPL-3.
FigureS4. MS/MS sequencing of saponified WT GPL-3.
FigureS5. MS^3 investigation of *m/z* 1038 fragment.
FigureS6. Analysis of *Δfmt* strains overproducing Fmt(G86A), Fmt(G88A) or Fmt(G86A/G88A).
FigureS7. MS^3 analysis of the *Δfmt* strain transformed with pMV306-*fmt*-G88A, pMV306-*fmt*-G86A or pMV306-*fmt*-G86A/G88A.
FigureS8. Atomic force microscopy with hydrophobic probes.
This information is available free of charge.

ACKNOWLEDGMENTS

We are indebted to PAGés platform (http://plateforme-pages.univ-lille1.fr) and UMS 2014 - US 41 - Plateformes Lilloises en Biologie & Santé for providing the scientific and technical environment conducive to achieving this work. YFD is Research Director at the FNRS.

This work was supported by the Fondation pour la Recherche Médicale (FRM) [grant number DEQ20150331719] to LK and the National Research Agency grant ANR-19-CE15-0012-01 (SUNLIVE) to LK and YG. We acknowledge the Infectiopôle Sud Méditerranée for funding the PhD fellowship of JK. Work at UCLouvain was supported by the Excellence of Science-EOS programme (Grant #30550343), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement n°693630), the FNRS-WELBIO (grant n°WELBIO-CR-2015A-05), the National Fund for Scientific Research (FNRS), and the Research Department of the Communauté française de Belgique (Concerted Research Action).

REFERENCES

(1) Johansen, M. D.; Herrmann, J.-L.; Kremer, L. (2020) Non-Tuberculous Mycobacteria and the Rise of *Mycobacterium Abscessus.Nat.Rev.Microbiol. 18*, 392-407.

(2) Brown-Elliott, B. A.; Nash, K. A.; Wallace, R. J. (2012) Antimicrobial Susceptibility Testing, Drug Resistance Mechanisms, and Therapy of Infections with Nontuberculous Mycobacteria. *Clin. Microbiol. Rev.25*, 545–582.

(3) Esther, C. R.; Esserman, D. A.; Gilligan, P.; Kerr, A.; Noone, P. G. (2010) Chronic *Mycobacterium abscessus* Infection and Lung Function Decline in Cystic Fibrosis. *J. Cyst. Fibros. 9*, 117–123.

(4) Sermet-Gaudelus, I.; Le Bourgeois, M.; Pierre-Audigier, C.; Offredo, C.; Guillemot, D.; Halley, S.; Akoua-Koffi, C.; Vincent, V.; Sivadon-Tardy, V.; Ferroni, A.; Berche, P.; Scheinmann, P.; Lenoir, G.; Gaillard, J.-L. (2003) *Mycobacterium abscessus* and Children with Cystic Fibrosis. *Emerging Infect. Dis. 9*, 1587–1591.

(5) Howard, S. T.; Rhoades, E.; Recht, J.; Pang, X.; Alsup, A.; Kolter, R.; Lyons, C. R.; Byrd, T. F. (2006) Spontaneous Reversion of *Mycobacterium abscessus* from a Smooth to a Rough Morphotype Is Associated with Reduced Expression of Glycopeptidolipid and Reacquisition of an Invasive Phenotype. *Microbiology (Reading, Engl.) 152*, 1581–1590.

(6) Medjahed, H.; Gaillard, J.-L.; Reyrat, J.-M. (2010) *Mycobacterium abscessus*: A New Player in the Mycobacterial Field. *Trends Microbiol. 18*, 117–123.

(7) Gutiérrez, A. V.; Viljoen, A.; Ghigo, E.; Herrmann, J.-L.; Kremer, L. (2018) Glycopeptidolipids, a Double-Edged Sword of the *Mycobacterium abscessus* Complex. *Front. Microbiol. 9*, 1145.

(8) Pawlik, A.; Garnier, G.; Orgeur, M.; Tong, P.; Lohan, A.; Le Chevalier, F.; Sapriel, G.; Roux, A. L.; Conlon, K.; Honoré, N.; Dillies, M. A.; Ma, L.; Bouchier, C.; Coppée, J. Y.; Gaillard, J. L.; Gordon, S. V.; Loftus, B.; Brosch, R.; Herrmann, J. L. (2013) Identification and Characterization of the Genetic Changes Responsible for the Characteristic Smooth-to-Rough Morphotype Alterations of Clinically Persistent *Mycobacterium abscessus*. *Mol. Microbiol. 90*, 612-629.

(9) Catherinot, E.; Roux, A.-L.; Macheras, E.; Hubert, D.; Matmar, M.; Dannhoffer, L.; Chinet, T.; Morand, P.; Poyart, C.; Heym, B.; Rottman, M.; Gaillard, J.-L.; Herrmann, J.-L. (2009) Acute Respiratory Failure Involving an R Variant of *Mycobacterium abscessus*. *J. Clin. Microbiol. 47*, 271–274.

(10) Jönsson, B. E.; Gilljam, M.; Lindblad, A.; Ridell, M.; Wold, A. E.; Welinder-Olsson, C. (2007) Molecular Epidemiology of *Mycobacterium abscessus*, with Focus on Cystic Fibrosis. *J. Clin. Microbiol. 45*, 1497–1504.

(11) Ryan, K.; Byrd, T. F. (2018) *Mycobacterium Abscessus*: Shapeshifter of the Mycobacterial World. *Front. Microbiol. 9*, 2642.

(12) Madani, A.; Ridenour, J. N.; Martin, B. P.; Paudel, R. R.; Abdul Basir, A.; Le Moigne, V.; Herrmann, J.-L.; Audebert, S.; Camoin, L.; Kremer, L.; Spilling, C. D.; Canaan, S.; Cavalier, J.-F. (2019) Cyclipostins and Cyclophostin Analogues as Multitarget Inhibitors That Impair Growth of *Mycobacterium abscessus*. *ACS Infect. Dis. 5*, 1597–1608.

(13) Bernut, A.; Viljoen, A.; Dupont, C.; Sapriel, G.; Blaise, M.; Bouchier, C.; Brosch, R.; de Chastellier, C.; Herrmann, J.-L.; Kremer, L. (2016) Insights into the Smooth-to-Rough Transitioning in *Mycobacterium bolletii* Unravels a Functional Tyr Residue Conserved in All Mycobacterial MmpL Family Members. *Mol. Microbiol. 99*, 866–883.

(14) Viljoen, A.; Gutiérrez, Ana Victoria; Dupont, C.; Ghigo, E.; Kremer, L. A. (2018) Simple and Rapid Gene Disruption Strategy in *Mycobacterium abscessus*: On the Design and Application of Glycopeptidolipid Mutants. *Front. Cell. Infect. Microbial. 8*, 69.

(15) Viljoen, A.; Viela, F.; Kremer, L.; Dufrêne, Y. F. (2020) Fast Chemical Force Microscopy Demonstrates That Glycopeptidolipids Define Nanodomains of Varying Hydrophobicity on Mycobacteria. *Nanoscale Horiz. 5*, 944-953.

(16) Bernut, A.; Herrmann, J.-L.; Kissa, K.; Dubremetz, J.-F.; Gaillard, J.-L.; Lutfalla, G.; Kremer, L. (2014) *Mycobacterium abscessus* Cording Prevents Phagocytosis and Promotes Abscess Formation. *Proc. Natl. Acad. Sci. U.S.A. 111*, E943-952.

(17) Roux, A.-L.; Viljoen, A.; Bah, A.; Simeone, R.; Bernut, A.; Laencina, L.; Deramaudt, T.; Rottman, M.; Gaillard, J.-L.; Majlessi, L.; Brosch, R.; Girard-Misguich, F.; Vergne, I.; de Chastellier, C.; Kremer, L.; Herrmann, J.-L. (2016) The Distinct Fate of Smooth and Rough *Mycobacterium abscessus* Variants inside Macrophages. *Open Biol. 6*, 160185.

(18) Whang, J.; Back, Y. W.; Lee, K.-I.; Fujiwara, N.; Paik, S.; Choi, C. H.; Park, J.-K.; Kim, H.-J. (2017) *Mycobacterium abscessus* Glycopeptidolipids Inhibit Macrophage Apoptosis and Bacterial Spreading by Targeting Mitochondrial Cyclophilin D. *Cell Death. Dis. 8*, e3012.

(19) Roux, A.-L.; Ray, A.; Pawlik, A.; Medjahed, H.; Etienne, G.; Rottman, M.; Catherinot, E.; Coppée, J.-Y.; Chaoui, K.; Monsarrat, B.; Toubert, A.; Daffé, M.; Puzo, G.; Gaillard, J.-L.; Brosch, R.; Dulphy, N.; Nigou, J.; Herrmann, J.-L. (2011) Overexpression of Proinflammatory TLR-2-Signalling Lipoproteins in Hypervirulent Mycobacterial Variants. *Cell. Microbiol. 13*, 692–704.

(20) Bernut, A.; Herrmann, J.-L.; Ordway, D.; Kremer, L. (2017) The Diverse Cellular and Animal Models to Decipher the Physiopathological Traits of *Mycobacterium abscessus* Infection. *Front. Cell. Infect. Microbiol. 7*, 100.

(21) Catherinot, E.; Clarissou, J.; Etienne, G.; Ripoll, F.; Emile, J.-F.; Daffé, M.; Perronne, C.; Soudais, C.; Gaillard, J.-L.; Rottman, M. (2007) Hypervirulence of a Rough Variant of the *Mycobacterium abscessus* TypeStrain. *Infect. Immun. 75*, 1055–1058.

(22) Bernut, A.; Dupont, C.; Sahuquet, A.; Herrmann, J.-L.; Lutfalla, G.; Kremer, L. (2015) Deciphering and Imaging Pathogenesis and Cording of *Mycobacterium abscessus* in Zebrafish Embryos. *J. Vis. Exp. 103*, e53130.

(23) Jankute, M.; Cox, J. A. G.; Harrison, J.; Besra, G. S. (2015) Assembly of the Mycobacterial Cell Wall. *Annu. Rev. Microbiol. 69*, 405–423.

(24) Brennan, P. J.; Nikaido, H. (1995) The Envelope of Mycobacteria. *Annu. Rev. Biochem. 64*, 29–63.

(25) Schorey, J. S.; Sweet, L. (2008) The Mycobacterial Glycopeptidolipids: Structure, Function, and Their Role in Pathogenesis. *Glycobiology 18*, 832–841.

(26) Ripoll, F.; Deshayes, C.; Pasek, S.; Laval, F.; Beretti, J.-L.; Biet, F.; Risler, J.-L.; Daffé, M.; Etienne, G.; Gaillard, J.-L.; Reyrat, J.-M. (2007) Genomics of Glycopeptidolipid Biosynthesis in *Mycobacterium abscessus* and M. Chelonae. *BMC Genomics 8*, 114.

(27) Burbaud, S.; Laval, F.; Lemassu, A.; Daffé, M.; Guilhot, C.; Chalut, C. (2016) Trehalose Polyphleates Are Produced by a Glycolipid Biosynthetic Pathway Conserved across Phylogenetically Distant Mycobacteria. *Cell. Chem. Biol. 23*, 278–289.

(28) Vats, A.; Singh, A. K.; Mukherjee, R.; Chopra, T.; Ravindran, M. S.; Mohanty, D.; Chatterji, D.; Reyrat, J.-M.; Gokhale, R. S. (2012) Retrobiosynthetic Approach Delineates the Biosynthetic Pathway and the Structure of the Acyl Chain of Mycobacterial Glycopeptidolipids. *J. Biol. Chem. 287*, 30677–30687.

(29) Villeneuve, C.; Etienne, G.; Abadie, V.; Montrozier, H.; Bordier, C.; Laval, F.; Daffe, M.; Maridonneau-Parini, I.; Astarie-Dequeker, C. (2003) Surface-Exposed Glycopeptidolipids of *Mycobacterium smegmatis* Specifically Inhibit the Phagocytosis of Mycobacteria by Human Macrophages. Identification of a Novel Family of Glycopeptidolipids. *J. Biol. Chem. 278*, 51291–51300.

(30) Ojha, A. K.; Varma, S.; Chatterji, D. (2002) Synthesis of an unusual polar glycopeptidolipid in glucose-limited culture of *Mycobacterium smegmatis*. *Microbiology 148*, 3039-3048.

(31) Deshayes, C.; Laval, F.; Montrozier, H.; Daffe, M.; Etienne, G.; Reyrat, J.-M. A. (2005) Glycosyltransferase Involved in Biosynthesis of Triglycosylated Glycopeptidolipids in *Mycobacterium smegmatis*: Impact on Surface Properties. *J. Bacteriol. 187*, 7283–7291.

(32) Wiersma, C.; Belardinelli, J. M.; Avanzi, C.; Angala, S. K.; Everall, I.; Angala, B.; Kendall, E.; Moura, V.; Verma, D.; Benoit, J.; Brown, K.; Jones, V.; Malcom, K.; Strong, M.; Nick, J.; Floto, R. A.; Parkhill, J.; Ordway, D.; Davidson, R.; McNeil, M. R.; Jackson, M. C. (2020) Cell Surface Remodeling of *Mycobacterium abscessus* under Cystic Fibrosis Airway Growth Conditions. *ACS Infect. Dis. 6*, 2143-2154.

(33) Sweet, L.; Zhang, W.; Torres-Fewell, H.; Serianni, A.; Boggess, W.; Schorey, J. (2008) *Mycobacterium avium* Glycopeptidolipids Require Specific Acetylation and Methylation Patterns for Signaling through Toll-like Receptor 2. *J. Biol. Chem. 283*, 33221–33231.

(34) Billman-Jacobe, H.; McConville, M. J.; Haites, R. E.; Kovacevic, S.; Coppel, R. L. (1999) Identification of a Peptide Synthetase Involved in the Biosynthesis of Glycopeptidolipids of *Mycobacterium smegmatis*. *Mol. Microbiol. 33*, 1244–1253.

(35) Sondén, B.; Kocíncová, D.; Deshayes, C.; Euphrasie, D.; Rhayat, L.; Laval, F.; Frehel, C.; Daffé, M.; Etienne, G.; Reyrat, J.-M. (2005) Gap, a Mycobacterial Specific Integral Membrane Protein, Is Required for Glycolipid Transport to the Cell Surface. *Mol. Microbiol. 58*, 426–440.

(36) Jeevarajah, D.; Patterson, J. H.; Taig, E.; Sargeant, T.; McConville, M. J.; Billman-Jacobe, H. (2004) Methylation of GPLs in *Mycobacterium smegmatis* and *Mycobacterium avium*. *J. Bacteriol. 186*, 6792–6799.

(37) Richard, M.; Gutiérrez, A. V.; Viljoen, A.; Rodriguez-Rincon, D.; Roquet-Baneres, F.; Blaise, M.; Everall, I.; Parkhill, J.; Floto, R. A.; Kremer, L. (2019) Mutations in the MAB_2299c TetR Regulator Confer Cross-Resistance to Clofazimine and Bedaquiline in *Mycobacterium abscessus*. *Antimicrob. Agents Chemother. 63*, e01316-18.

(38) Lee, M. H.; Pascopella, L.; Jacobs, W. R.; Hatfull, G. F. (1991) Site-Specific Integration of Mycobacteriophage L5: Integration-Proficient Vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and Bacille Calmette-Guérin. *Proc. Natl. Acad. Sci. U.S.A. 88*, 3111–3115.

(39) Gao, J.; Sampson, N. S. (2014) A GMC Oxidoreductase Homologue Is Required for Acetylation of Glycopeptidolipid in *Mycobacterium smegmatis*. *Biochemistry 53*, 611–613.

(40) Lopez-Marin, L. M.; Gautier, N.; Laneelle, M.-A.; Silve, G.; Daffe, M. (1994) Structures of the Glycopeptidolipid Antigens of *Mycobacterium abscessus* and *Mycobacterium chelonae* and Possible Chemical Basis of the Serological Cross-Reactions in the *Mycobacterium fortuitum* Complex. *Microbiology 140*, 1109–1118.

(41) López Marin, L. M.; Lanéelle, M. A.; Promé, D.; Daffé, M.; Lanéelle, G.; Promé, J. C. (1991) Glycopeptidolipids from *Mycobacterium fortuitum*: A Variant in the Structure of C-Mycoside. *Biochemistry 30* , 10536–10542.

(42) Keller, J. P.; Smith, P. M.; Benach, J.; Christendat, D.; deTitta, G. T.; Hunt, J. F. (2002) The Crystal Structure of MT0146/CbiT Suggests That the Putative Precorrin-8w Decarboxylase Is a Methyltransferase. *Structure 10*, 1475–1487.

(43) Rainczuk, A. K.; Klatt, S.; Yamaryo-Botté, Y.; Brammananth, R.; McConville, M. J.; Coppel, R. L.; Crellin, P. K. MtrP, (2020) A Putative Methyltransferase in *Corynebacteria*, Is Required for Optimal Membrane Transport of Trehalose Mycolates. *J. Biol. Chem. 295*, 6108–6119.

(44) Pérez, E.; Constant, P.; Laval, F.; Lemassu, A.; Lanéelle, M.-A.; Daffé, M.; Guilhot, C. (2004) Molecular Dissection of the Role of Two Methyltransferases in the Biosynthesis of Phenolglycolipids and Phthiocerol Dimycoserosate in the *Mycobacterium tuberculosis* Complex. *J. Biol. Chem. 279*, 42584–42592.

(45) Jeevarajah, D.; Patterson, J. H.; McConville, M. J.; Billman-Jacobe, H. (2002) Modification of Glycopeptidolipids by an *O*-Methyltransferase of *Mycobacterium smegmatis*. *Microbiology (Reading, Engl.) 148*, 3079–3087.

(46) Jankute, M.; Nataraj, V.; Lee, O. Y.-C.; Wu, H. H. T.; Ridell, M.; Garton, N. J.; Barer, M. R.; Minnikin, D. E.; Bhatt, A.; Besra, G. S. (2017) The Role of Hydrophobicity in Tuberculosis Evolution and Pathogenicity. *Sci. Rep. 7*, 1315.

(47) Alsteens, D.; Dague, E.; Rouxhet, P. G.; Baulard, A. R.; Dufrêne, Y. F. (2007) Direct Measurement of Hydrophobic Forces on Cell Surfaces Using AFM. *Langmuir. 23*, 11977–11979.

(48) Doyle, R. J. (2000) Contribution of the Hydrophobic Effect to Microbial Infection. *Microbes Infect. 2*, 391–400.

(49) Minnikin, D. E.; Lee, O. Y.-C.; Wu, H. H. T.; Besra, G. S.; Bhatt, A.; Nataraj, V.; Rothschild, B. M.; Spigelman, M.; Donoghue, H. D. (2015) Ancient Mycobacterial Lipids: Key Reference Biomarkers in Charting the Evolution of Tuberculosis. *Tuberculosis (Edinb) 95*, S133-139.

(50) Johansen, M. D.; Kremer, L. (2020) A Zebrafish Model of *Mycobacterium kansasii* Infection Reveals Large Extracellular Cord Formation. *J. Infect. Dis. 222*, 1046-1050.

(51) Alsteens, D.; Verbelen, C.; Dague, E.; Raze, D.; Baulard, A. R.; Dufrêne, Y. F. (2008) Organization of the Mycobacterial Cell Wall: A Nanoscale View. *Pflugers Arch. - Eur. J. Physiol. 456*, 117–125.

(52) Dague, E.; Alsteens, D.; Latgé, J.-P.; Verbelen, C.; Raze, D.; Baulard, A. R.; Dufrêne, Y. F. (2007) Chemical Force Microscopy of Single Live Cells. *Nano. Lett. 7*, 3026–3030.

(53) Arbues, A.; Lugo-Villarino, G.; Neyrolles, O.; Guilhot, C.; Astarie-Dequeker, C. (2014) Playing Hide-and-Seek with Host Macrophages through the Use of Mycobacterial Cell Envelope Phthiocerol Dimycocerosates and Phenolic Glycolipids. *Front. Cell. Infect. Microbiol. 4*, 173.

(54) Etienne, G.; Villeneuve, C.; Billman-Jacobe, H.; Astarie-Dequeker, C.; Dupont, M.-A.; Daffé, M. (2002) The Impact of the Absence of Glycopeptidolipids on the Ultrastructure, Cell Surface and Cell Wall Properties, and Phagocytosis of *Mycobacterium smegmatis*. *Microbiology 148*, 3089–3100.

(55) Bartley, S. N.; Tzeng, Y.-L.; Heel, K.; Lee, C. W.; Mowlaboccus, S.; Seemann, T.; Lu, W.; Lin, Y.-H.; Ryan, C. S.; Peacock, C.; Stephens, D. S.; Davies, J. K.; Kahler, C. M. (2013) Attachment and Invasion of *Neisseriameningitidis* to Host Cells Is Related to Surface Hydrophobicity, Bacterial Cell Size and Capsule. *PLoSONE8*, e55798.

(56) Araújo, A. M. M.; Oliveira, I. C. M. de; Mattos, M. C. de; Benchetrit, L. C. (2008) Cell Surface Hydrophobicity and Adherence of a Strain of Group B *Streptococci* during the Post-Antibiotic Effect of Penicillin. *Rev. Inst. Med. Trop. SaoPaulo 50*, 203–207.

(57) Horstmann, J. A.; Lunelli, M.; Cazzola, H.; Heidemann, J.; Kühne, C.; Steffen, P.; Szefs, S.; Rossi, C.; Lokareddy, R. K.; Wang, C.; Lemaire, L.; Hughes, K. T.; Uetrecht, C.; Schlüter, H.; Grassl, G. A.; Stradal, T. E. B.; Rossez, Y.; Kolbe, M.; Erhardt, M. (2020) Methylation of *Salmonellatyphimurium* Flagella Promotes Bacterial Adhesion and Host Cell Invasion. *Nat. Commun. 11*, 2013.

(58) Viljoen, A.; Blaise, M.; de Chastellier, C.; Kremer, L. (2016) MAB_3551c Encodes the Primary Triacylglycerol Synthase Involved in Lipid Accumulation in *Mycobacteriumabscessus*. *Mol. Microbiol. 102*, 611–627.

(59) Hutter, J. L.; Bechhoefer, J. (1993) Calibration of Atomic-force Microscope Tips. *Reviewof Scientific Instruments 64*, 1868–1873.

(60) Raynaud, C.; Daher, W.; Johansen, M.; Roquet-Baneres, F.; Blaise, M.; Onajole, O. K.; Kozikowski, A.; Herrmann, J.-L.; Dziadek, J.; Gobis, K.; Kremer, L. (2020) Active Benzimidazole Derivatives Targeting the MmpL3 Transporter in *Mycobacterium abscessus*. *ACSInfectiousDiseases 6*, 324-337.

SUPPORTINGINFORMATION

O-Methylation of the glycopeptidolipid acyl chain defines surface hydrophobicity of *Mycobacterium abscessus* and macrophage invasion

Wassim Daher[1,4,&], Louis-David Leclercq[2,&], Albertus Viljoen[1,3], Jona Karam[1], Yves F. Dufrêne[3], Yann Guérardel[2], and Laurent Kremer[1,4,#]

[1]Centre National de la Recherche Scientifique UMR 9004, Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France.

[2]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.

[3]Louvain Institute of Biomolecular Science and Technology, Université Catholique de Louvain, Croix du Sud, 4-5, bte L7.07.07., B-1348, Louvain-la-Neuve, Belgium.

[4]INSERM, IRIM, 34293 Montpellier, France.

[#]To whom correspondence should be addressed: Tel: (+33) 4 34 35 94 47;

E-mail: laurent.kremer@irim.cnrs.fr

[&]These authors contributed equally to that work.

TableS1. Primers used in this study. Restriction sites are bold and underlined.

Number	5'→3' sequence	Restriction site	F (sense), R (antisense)
Cloning in pUX1-*katG*			
1	gaga**TTAATTAA**AAGCAGCAACGTCCTGGTCGG	PacI	To clone the left arm in pUX1-*katG* (F)
2	gaga**CAATTG**GATGAATTTCTGAACCCGCTGCGTGA	MfeI	To clone the left arm in pUX1-*katG* (R)
3	gaga**CAATTG**GTCGAATACCGGATGTACCGCTTTGT	MfeI	To clone the right arm in pUX1-*katG*(F)
4	gaga**GCTAGC**GAGGACACTGCGGTATCGAGAAG	NheI	To clone the right arm in pUX1-*katG*(R)
Primers to verify double homologous recombination			
5	GAATGGTGATCTCGCTGAAATCGGTAC	-	To check 5' integration *Δfmt* (F)
6	GTTCTATCGACTCGACGACGTTCATC	-	To check 5' integration *Δfmt* (R)
7	CAAACTTCATACTGTCCCTGGCCC	-	To check 3' integration *Δfmt* (F)
8	CAAGTGCCACAGCAAAGTACTCAACTAG	-	To check 3' integration *Δfmt* (R)
Cloning in pMV306-KanR			
9	Gaga**GGTACC**ACAAATCCGCGTCTTCGGCCAG	KpnI	To clone the *fmt-HA* ORF in pMV306 (F)
10	gaga**CCATGG**CTAAGCGTAATCTGGAACATCGT ATGGGTAGGCCTTGACAAAGCGGTACATCC	NcoI	To clone the *fmt-HA* ORF in pMV306 (R)
Site-directed mutagenesis			
11	CGGGTGTTGGAGGTCGCTTGTGGTCACGGGGGTGGTGCC	-	To introduce the G86A mutation in pMV306-*fmt* (F)
12	GGCACCACCCCCGTGACCACAAGCGACCTCCAACACCCG	-	To introduce the G86A mutation in pMV306-*fmt* (R)
13	CGGGTGTTGGAGGTCGGTTGTGCTCACGGGGGTGGTGCC	-	To introduce the G88A mutation in pMV306-*fmt* (F)
14	GGCACCACCCCCGTGAGCACAACCGACCTCCAACACCCG	-	To introduce the G88A mutation in pMV306-*fmt* (R)
15	CGGGTGTTGGAGGTCGCTTGTGCTCACGGGGGTGGTGCC	-	To introduce the G86A/G88A mutations in pMV306-*fmt* (F)
16	GGCACCACCCCCGTGAGCACAAGCGACCTCCAACACCCG	-	To introduce the G86A/G88A mutations in pMV306-*fmt* (R)

TableS2. Bacterial strains used in this study.

Name	Description/genotype	Marker	Reference
Strains			
M. abscessus Smooth (S)	*M. abscessussensustricto*, strain CIP104536^T, S morphotype	-	Laboratoire de Référence des Mycobactéries
M. abscessus Rough (R)	*M. abscessussensustricto*, strain CIP104536^T, R morphotype	-	Laboratoire de Référence des Mycobactéries
Δfmt	Unmarked deletion of *fmt* in *M. abscessus*S	-	This study
Δfmt + pMV306-*fmt*	*Δfmt* carrying pMV306-*fmt*	Kan	This study
Δfmt + pMV306-*fmt*-G86A	*Δfmt* carrying pMV306-*fmt*-G86A	Kan	This study
Δfmt + pMV306-*fmt*-G88A	*Δfmt* carrying pMV306-*fmt*-G88A	Kan	This study
Δfmt + pMV306-*fmt*-G86A/G88A	*Δfmt* + pMV306-*fmt*-G86A/G88A	Kan	This study
E.coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17supE44relA1 lac* [F´ *proABlacI*qZΔ*M15* Tn*10* (Tetr)].	Tet	Stratagene

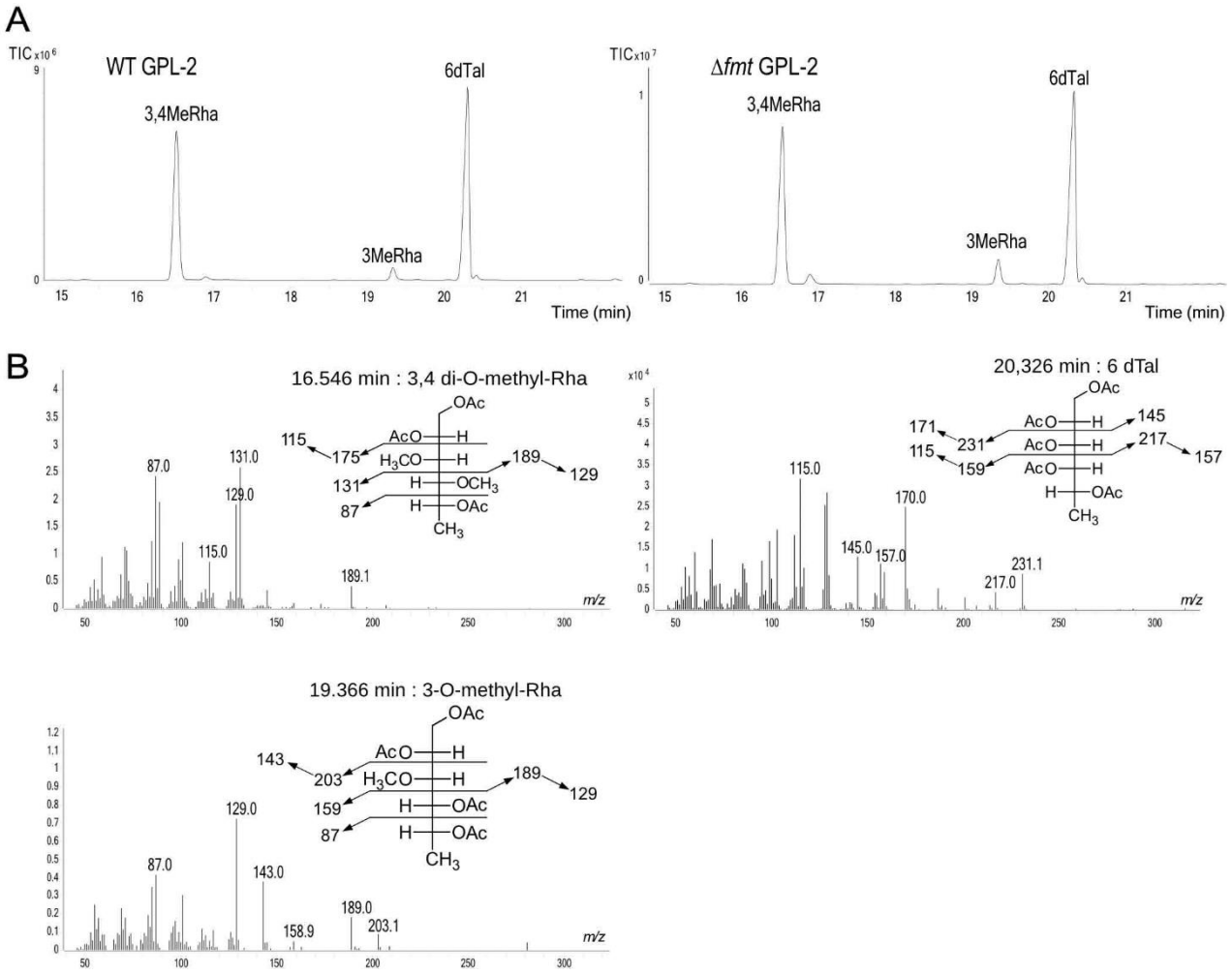
TableS3. Plasmids used in this study.

Plasmids			
pTEC27	Multicopy *E. coli*/mycobacterial shuttle vector expressing *tdTomato* under the control of a strong mycobacterial promoter.	Hyg	Addgene (plasmid 30182)
pUX1-*katG*	Variant of *pUX1* containing the *katG* gene of *M. tuberculosis* as a counter-selectable marker in the presence of INH and used for the generation of unmarked chromosomal alterations.	Kan, Hyg	(1)
pUX1-*katG-fmt*	pUX1-*katG* containing the upstream and downstream sequences surrounding *fmt.*	Kan	This study
pMV306	*E. coli*/mycobacterial integrative vector	Kan	(2)
pMV306-*fmt*	pMV306 allowing expression of Fmt-HA in mycobacteria	Kan	This study
pMV306-*fmt*-G86A	pMV306 allowing expression of Fmt-HA-G86A in mycobacteria	Kan	This study
pMV306-*fmt*-G88A	pMV306 allowing expression of Fmt-HA-G88A in mycobacteria	Kan	This study
pMV306-*fmt*-G86A/G88A	pMV306 allowing expression of Fmt-HA-G86A/G88A in mycobacteria	Kan	This study

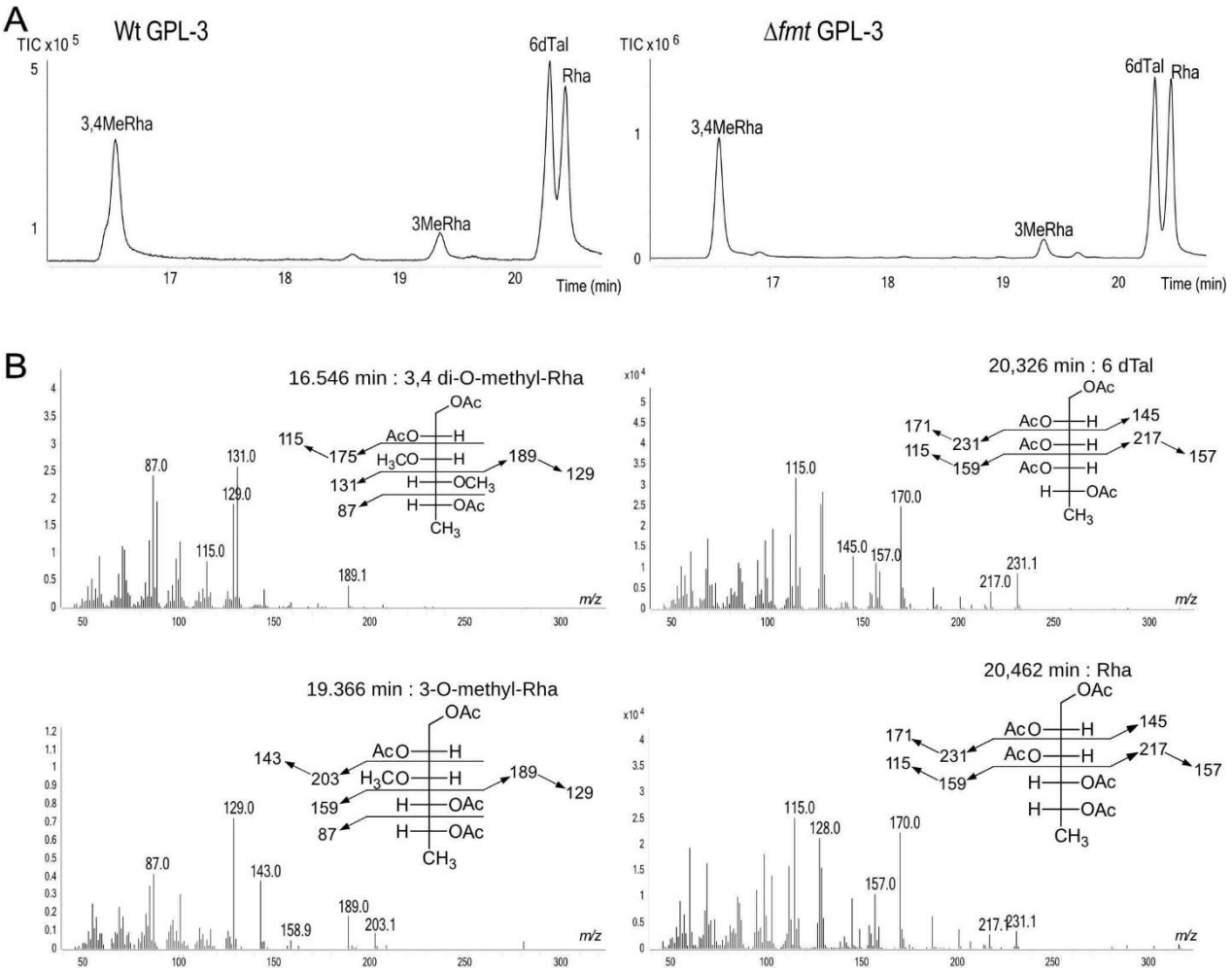
Hyg, hygromycin; Kan, kanamycin.

1. Richard, M.; Gutiérrez, A. V.; Viljoen, A.; Rodriguez-Rincon, D.; Roquet-Baneres, F.; Blaise, M.; Everall, I.; Parkhill, J.; Floto, R. A.; Kremer, L. (2019) Mutations in the MAB_2299c TetR Regulator Confer Cross-Resistance to Clofazimine and Bedaquiline in *Mycobacterium abscessus*. *Antimicrob. Agents Chemother. 63*, e01316-18.

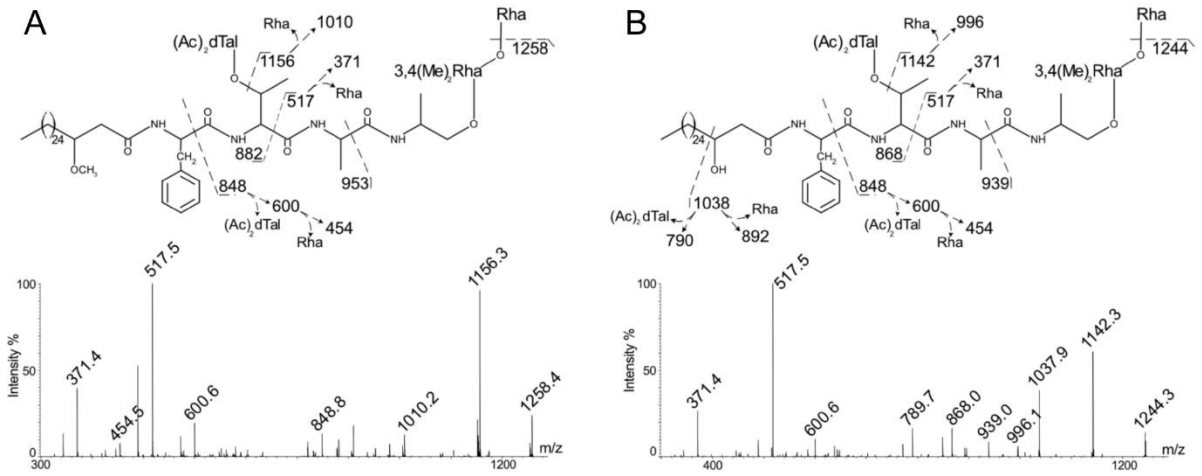
2. Stover, C.K.; de la Cruz V.F.; Fuerst, T.R.; Burlein, J.E.; Benson, L.A.; Bennett, L.T.; Bansal, G.P.; Young, J.F.; Lee, M.H.; Hatfull, G.F. (1991) New use of BCG for recombinant vaccines. *Nature 351*, 456–460.



FigureS1. Monosaccharideanalysisofpurified diglycosylatedGPL.(A)GC/MS chromatograms of partially methylated itol acetate derivatives from GC/MS shows two identical major signals for WT GPL-2 and *Δfmt* GPL-2. **(B)** EI-MS identified the major signals as 3,4 di-*O*-methyl-Rha (16,546 min) and 6-dTal (20.360 min). 3-*O*-methyl-Rha (19.366 min) appeared as a minor component.



FigureS2. Monosaccharideanalysisofpurified triglycosylatedGPL.(A)GC/MS chromatograms of partially methylated itol acetate derivatives show identical major signals for WT GPL-2 and *Δfmt* GPL-2. **(B)** EI-MS identified major signals as 3,4 di-*O*-methyl-Rha(16.5 min), Rha (20.5 min) and 6-dTal (20.3 min). 3-*O*-methyl-Rha (19.4 min) appeared as a minor component. 6-dTal and Rha were discriminated using chromatographic mobilities of authentic standards.



FigureS3.MS/MS sequencingofGPL-3.MS^2 fragmentation spectra of MS signals at **(A)** *m/z* 1404 of WT-GPL-3; MS^2 fragmentation spectra of MS signals at **(B)** *m/z* 1390 of *Δfmt* GPL-3.

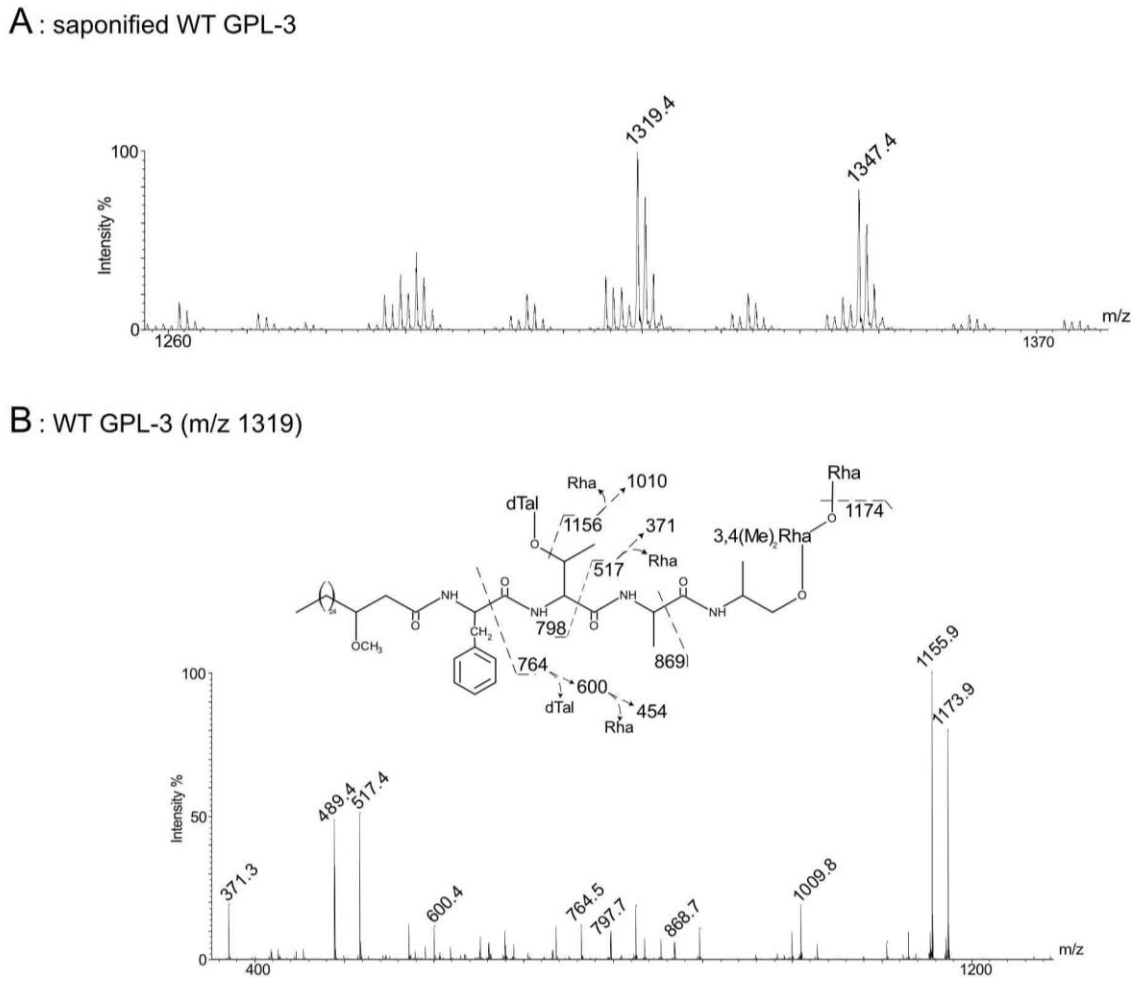
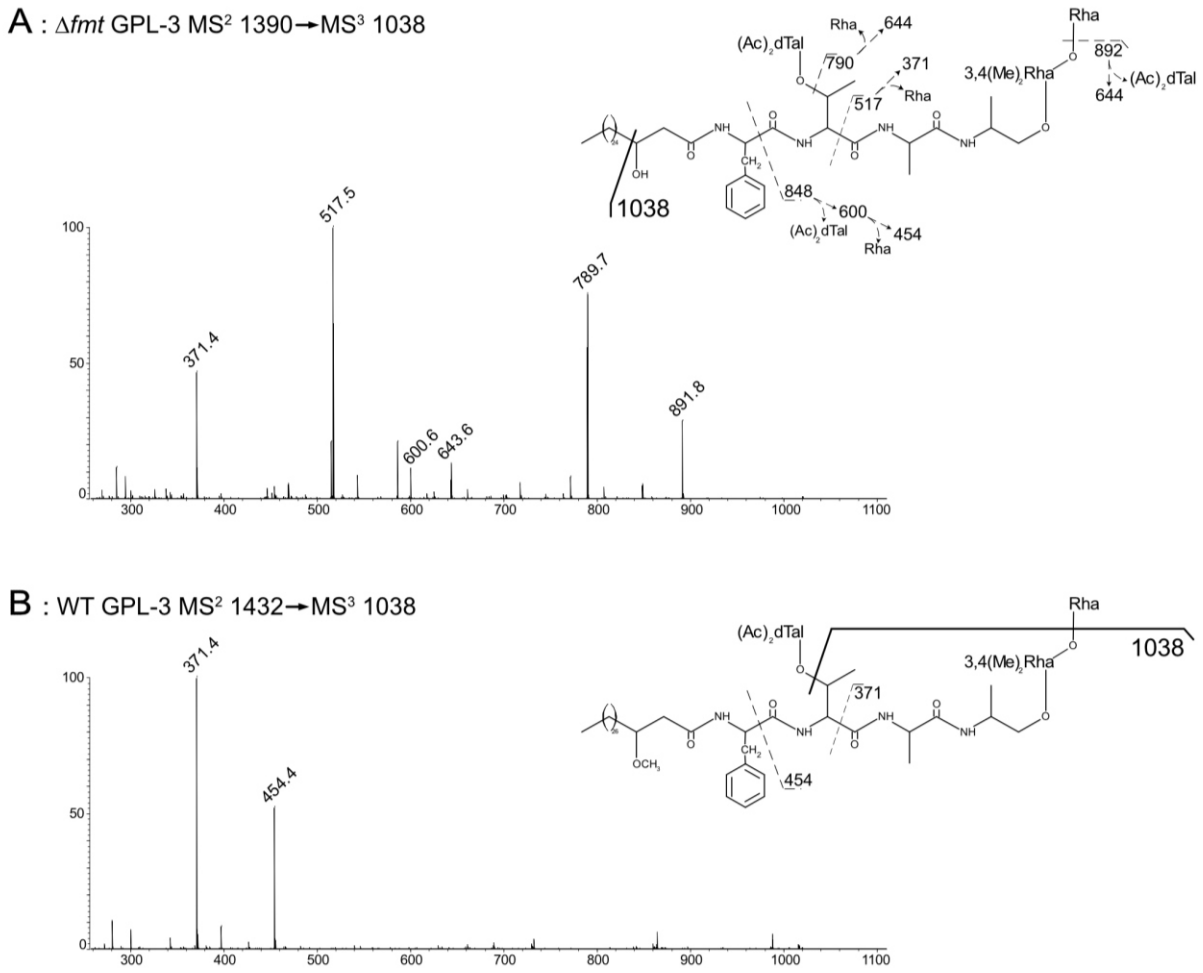


Figure S4. MS/MS sequencing of saponified WT GPL-3.(A) MS spectrum of saponified WT GPL-3 shows two major signals at *m/z* 1319 and 1347 with a -84 m.u. discrepancy compared to native WT GPL-3 **(Fig. 3C)**. This 84 m.u. decrease was observed for all saponified fractions (data not shown). **(B)** MS^2 spectra of the ion at *m/z* 1319 in comparison with corresponding MS^2 spectra from the native fraction **(Fig S3A)** established that both acetyl groups substitute 6-dTal.



FigureS5.MS³investigationof*m/z* 1038fragment.(A) MS³ analysis of the MS² fragment ion at *m/z* 1038 from *Δfmt* GPL-3 establishes the presence of an OH group at the C3 position of the lipid moiety owing to fragment ions at *m/z* 892, 790/644 and 517/371. **(B)** Corresponding MS³ spectrum from the WT GPL-3 fraction does not establish the presence of such a structural feature.

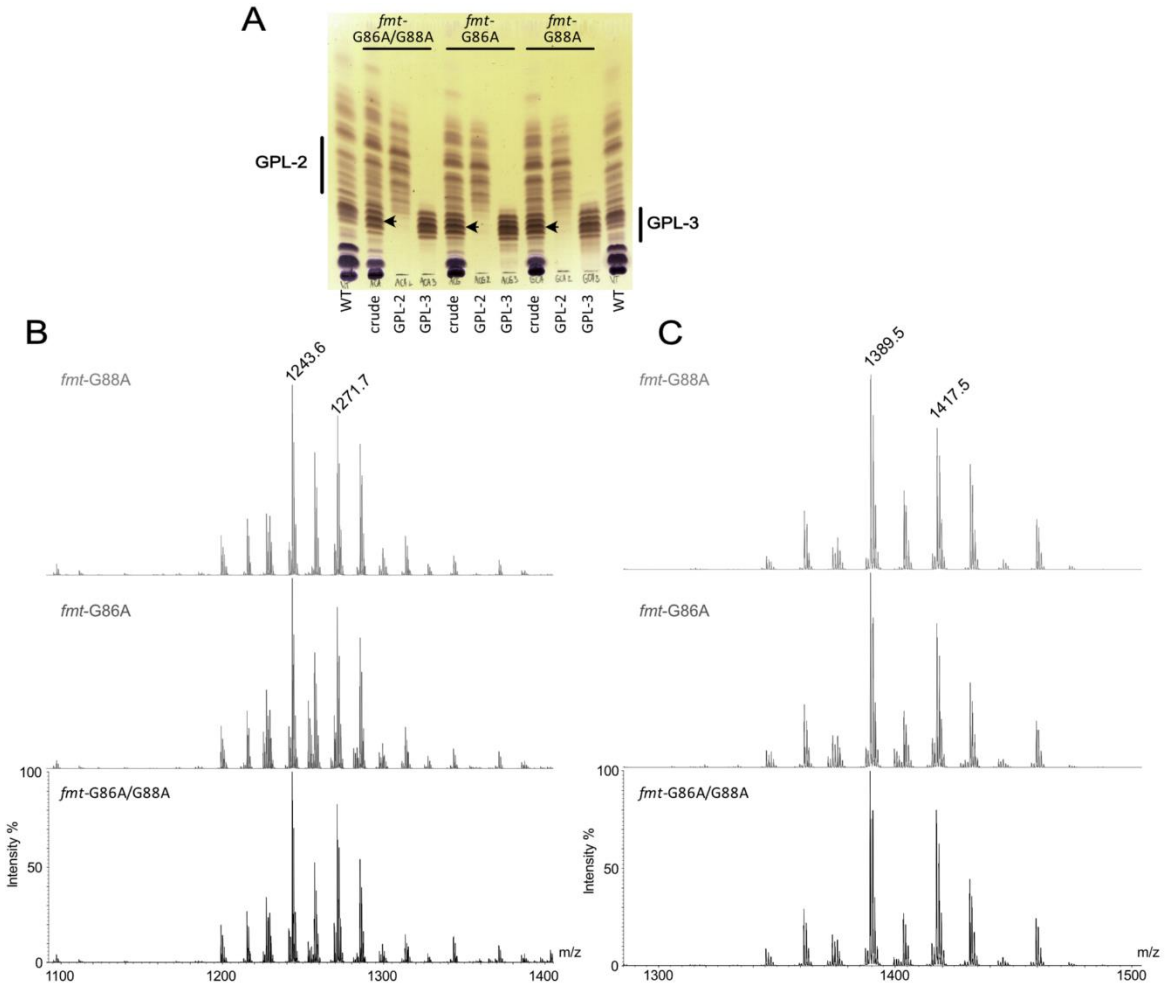


Figure S6. Analysis of *Δfmt* strains overproducing Fmt(G86A), Fmt(G88A) or Fmt(G86A/G88A). (A) TLC analyses of Δ*fmt* mutants complemented either with pMV306-*fmt*-G86A/G88A, pMV306-*fmt*-G86A or pMV306-*fmt*-G88A before or after purification of GPL-2 and GPL-3. Black arrows indicate a set of intense bands corresponding to more polar GPL-3 compared to the WT S progenitor (first and last lane). **(B)** Overlay of MS spectra obtained for GPL-2 fractions from the three transformed Δ*fmt* strains shows ions at *m/z* 1244 and 1272. **(C)** Overlay of MS spectra obtained for GPL-3 fractions from the three transformed Δ*fmt* strains shows ions at *m/z* 1390 and 1418.

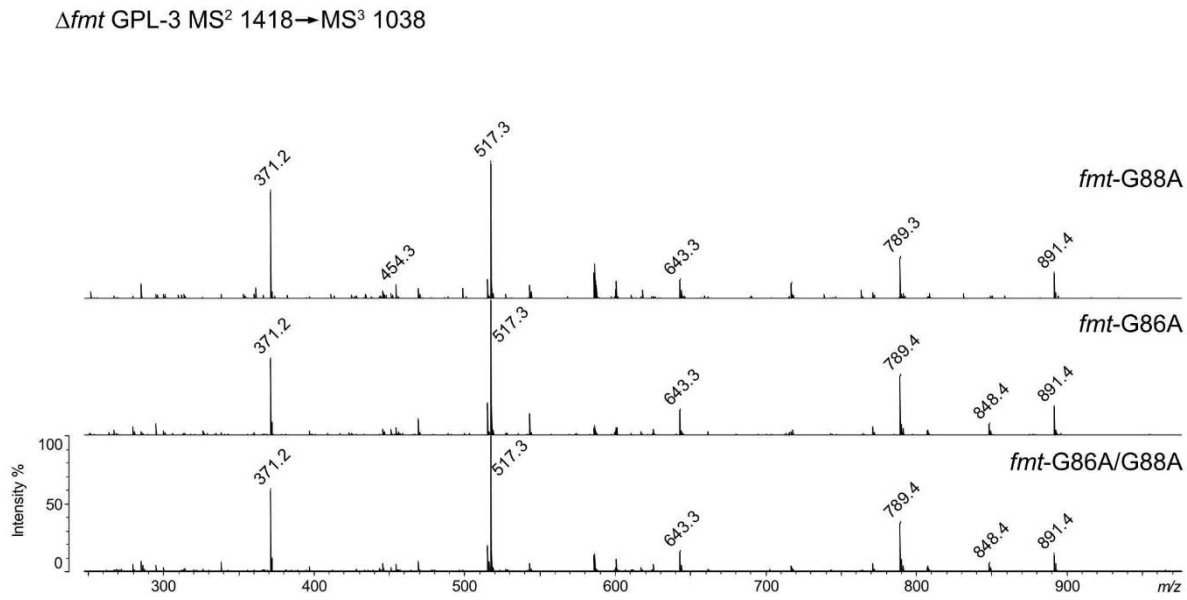


Figure S7. MS^3 analysis of the *Δfmt* strain transformed with pMV306-*fmt*-G88A (upper panel), pMV306-*fmt*-G86A (middle panel) or pMV306-*fmt*-G86A/G88A (lower panel). MS^3 spectra were derived from selection of the GPL-3 MS ion at *m/z* 1418 and the MS^2 ion at *m/z* 1038. The scheme is given in **Fig.S5A**.

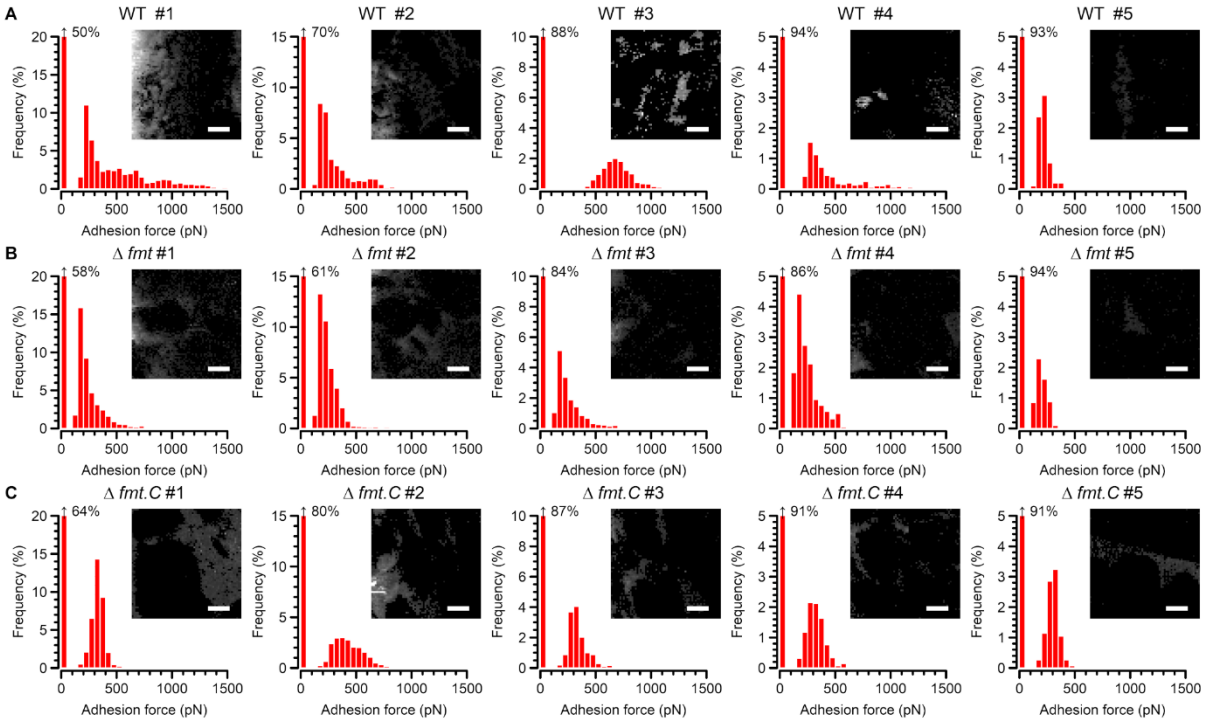


Figure S8. Atomic force microscopy with hydrophobic probes reveals a shift in the magnitudes of hydrophobic adhesive forces on the surface of *Δfmt* cells. (A) Histogram plots showing binding probability *versus* adhesion force and their corresponding force maps (insets, scalebar = 50 nm, z-range depicted in equispaced-interval grayscale from black, 0 pN, to white, 1500 pN) for five wild type (WT) *M. abscessus* S variant cells. **(B)** Data sets for five *Δfmt* cells **(C)** and five *ΔfmtC* cells.

III.E.5.b Etude de la glycosylation des GPLs via l'analyse structurale des GPLs des mutants *Δgtfs* et impact in-vitro sur la primo-infection (2ème article)

Ces travaux ont donné lieu à la publication d'un article en 2022 dans le journal *Cellular and Chemical Biology* dont je suis co-premier auteur[391]. Les résultats seront présentés dans la version libre de droits de l'article et commentés en faisant référence aux figures originales de l'article. Des figures supplémentaires qui n'ont pas été publiées seront également présentées pour illustrer ma contribution.

Dans la suite de l'étude de la biosynthèse des GPLs, j'ai poursuivi la caractérisation de nouvelles formes s'accumulant chez les mutants simples, doubles et triples des trois glycosyltransférases *gtf1*, *gtf2* et *gtf3*. Parmi les mutants simples analysés en CCM, *Δgtf1* continuait de produire des GPLs avec un profil très différent de la souche sauvage, *Δgtf2* ne produisait plus de GPLs et les bandes correspondantes au GPL-3 avait disparu du profil de *Δgtf3* (Fig. 1B). D'après la composition en monosaccharides des extraits polaires totaux, le 6dTal était absent dans le mutant *Δgtf1* et le rhamnose dans le mutant *Δgtf3* alors qu'aucun des déoxyhexoses propres aux GPLs n'avaient pu être détectés dans le mutant *Δgtf2* (Fig. 1C).

La purification des GPLs à partir du mutant *Δgtf3* a révélé des fractions positives à l'orcinol-sulfurique plus polaires que les GPL-2a, et dont le spectre de masse produit un ion intense *m/z* 1357. Le spectre MS^2 de cet ion affiche un fragment M-274 u.m.a. qui lui-même donne un ion M-248 u.m.a. sur le spectre MS^3. C'est pourquoi, j'ai supposé que ce composé était la forme succinylée du GPL-2a perdant d'abord le 2-*O*- succinyl 3,4 di-*O*-méthyl rhamnose puis le 3,4 di-*O*-acétyl 6 déoxy talose mais ces données n'ont pas été publiées (Figure 51A & B). Cette même structure a été décrite par Villeneuve et collaborateurs chez *Msmeg*[345].

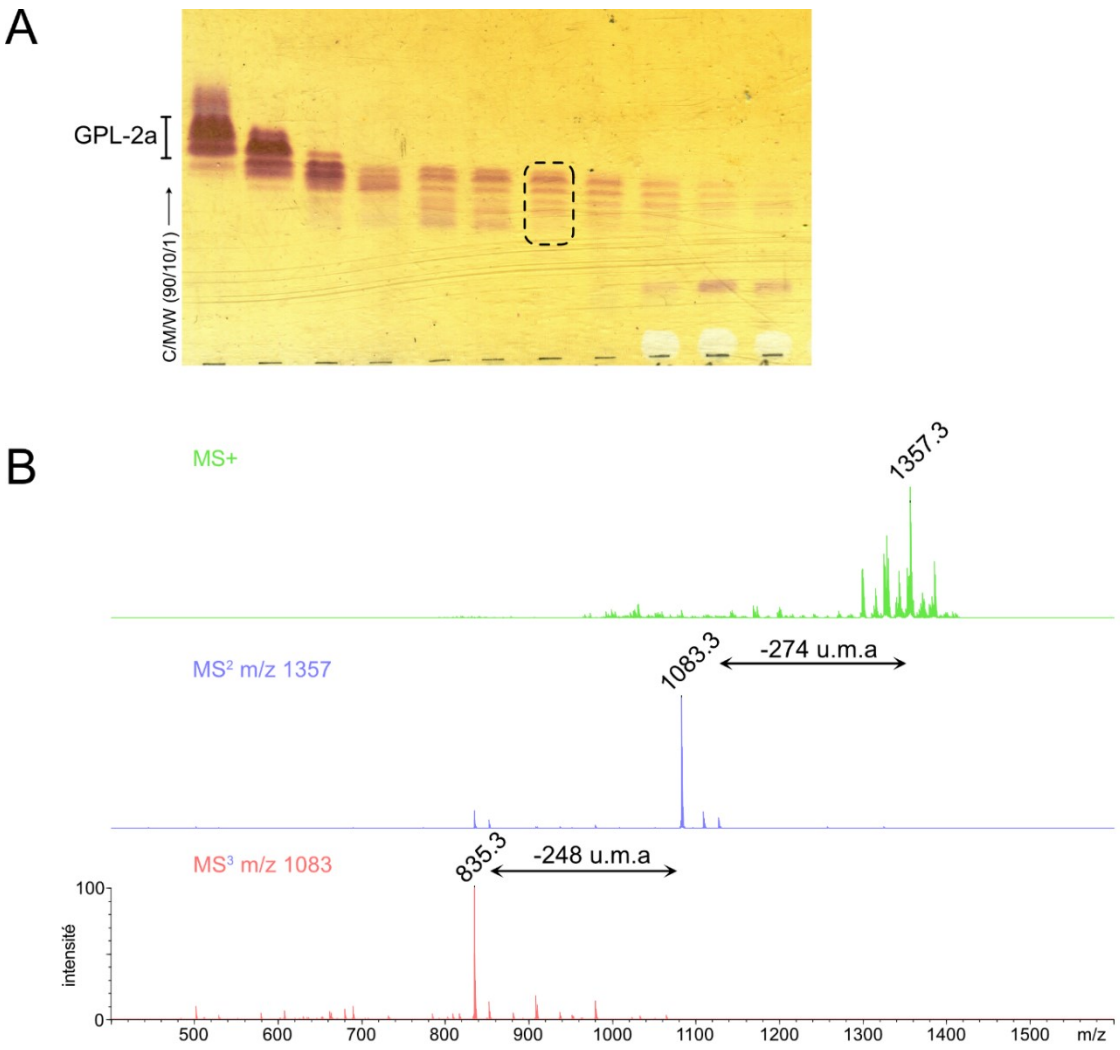


Figure 51 : Caractérisation du GPL-2a succinylé dans *Δgtf3*. La purification de la fraction polaire total du mutant *Mabs* S *Δgtf3* (A) a conduit à la détection par l'orcinol d'une fraction plus polaire que le GPL-2a. Son spectre de masse (vert) révèle un ion majeur à *m/z* 1357 dont la fragmentation MS^2 (bleu) produit un ion à *m/z* 1083 très intense M-274 u.m.a.. Le spectre MS^3 (rouge) de ce dernier génère un ion à *m/z* 835, M-248 u.m.a. correspondant à la perte du 3,4 di-*O*-acétyl 6 déoxy talose.

Chez les mutants doubles, j'ai caractérisé pour le mutant *Δgtf1,2* deux bandes CCM révélées uniquement par l'acide phosphomolybidique (PMA) qui se sont révélées être le lipopeptide non glycosylé, contenant soit un résidu d'alaninol soit un résidu de valinol. En effet, après purification de deux fractions F1 et F2 par CCM préparative (Figure 52A), les spectres de masse montrent respectivement 2 ions majoritaires *m/z* 837 et 865 puis 865 et 893 (Figure 52B). En parallèle, les corrélations des protons de l'alaninol observées pour la fraction F1 (Figure 52C) disparaissent sur le spectre $^1H/^1H$ COSY de la fraction F2 (Figure 52D). J'en ai déduit que la différence constatée de masse de 28 u.m.a. était moins lié à la variation de longueur de chaine du lipide qu'à une modification de l'alaninol en valinol.

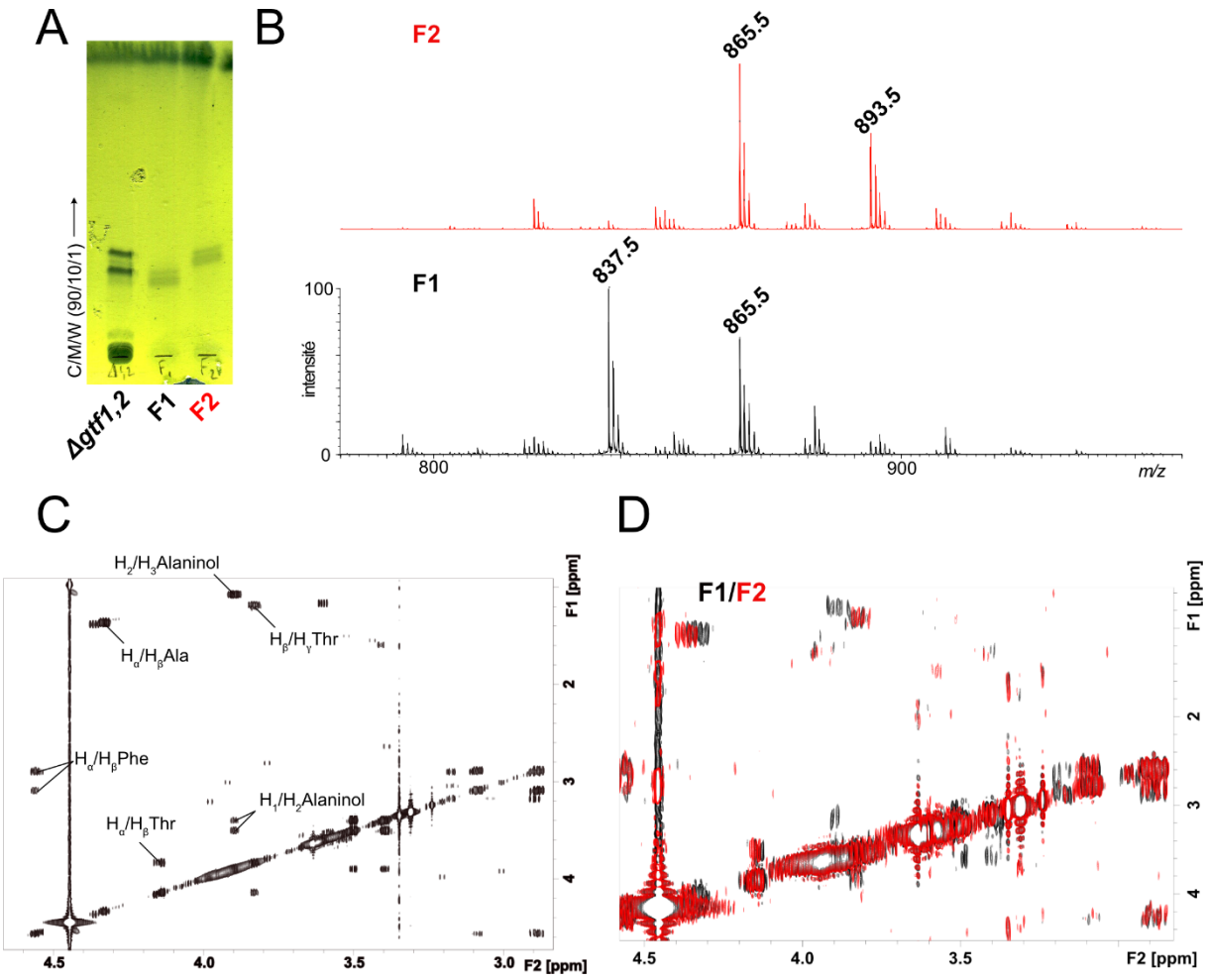


Figure 52 : Caractérisation du lipopeptide dans *Δgtf1,2*. La purification de la fraction polaire total du mutant *Mabs* S *Δgtf1,2* (A) a conduit à la détection par le PMA de 2 fractions F1 et F2. Leurs spectres de masse (B) révèlent chacun deux ions majeurs à *m/z* 837 et 865 ainsi que 865 et 893. Le spectre $^{1}H/^{1}H$ COSY de la fraction F1 (C) présentent les corrélations des acides aminés et de l'alaninol. La superposition avec le spectre $^{1}H/^{1}H$ COSY de la fraction F2 (D) montre que les corrélations de l'alaninol disparaissent.

Dans le cas du mutant *Δgtf1,3*, une analyse RMN complète a confirmé la présence d'un nouveau GPL-1b monoglycosylé ainsi que de valinol (Fig. S3C). En outre, ce composé a été mis en évidence par la CCM et la spectrométrie de masse dans le mutant *Δgtf1.* Une forme diglycosylée appelée GPL-2b a finalement était distinguée dans ce mutant (Fig. S3A). Pour appuyer les hypothèses publiées, j'ai notamment séparé les GPLs s'accumulant de *Δgtf1* (Figure 53) et *Δgtf1,3* (Figure 54) en trois fractions à partir desquelles j'ai déterminé la composition en monosaccharide et réalisé une analyse MALDI-TOF. Par exemple, le GPL-2b a été mis en évidence dans la fraction F1 la plus polaire obtenue à partir de *Δgtf1*. Cette fraction présentait une part équivalente de rhamnose et de 3,4 di-*O*-méthyl rhamnose ainsi qu'une paire d'ions *m/z* 1159 et 1187 (Figure 53A-C), ce dernier révélant des

fragments M-146 u.m.a. sur le spectre MS^2, témoin de la perte de déoxyhexose (Fig. S3A). De plus l'hydrolyse alcaline ne modifiait pas le rapport *m/z* des ions détectés pour les GPL-1b et GPL-2b.

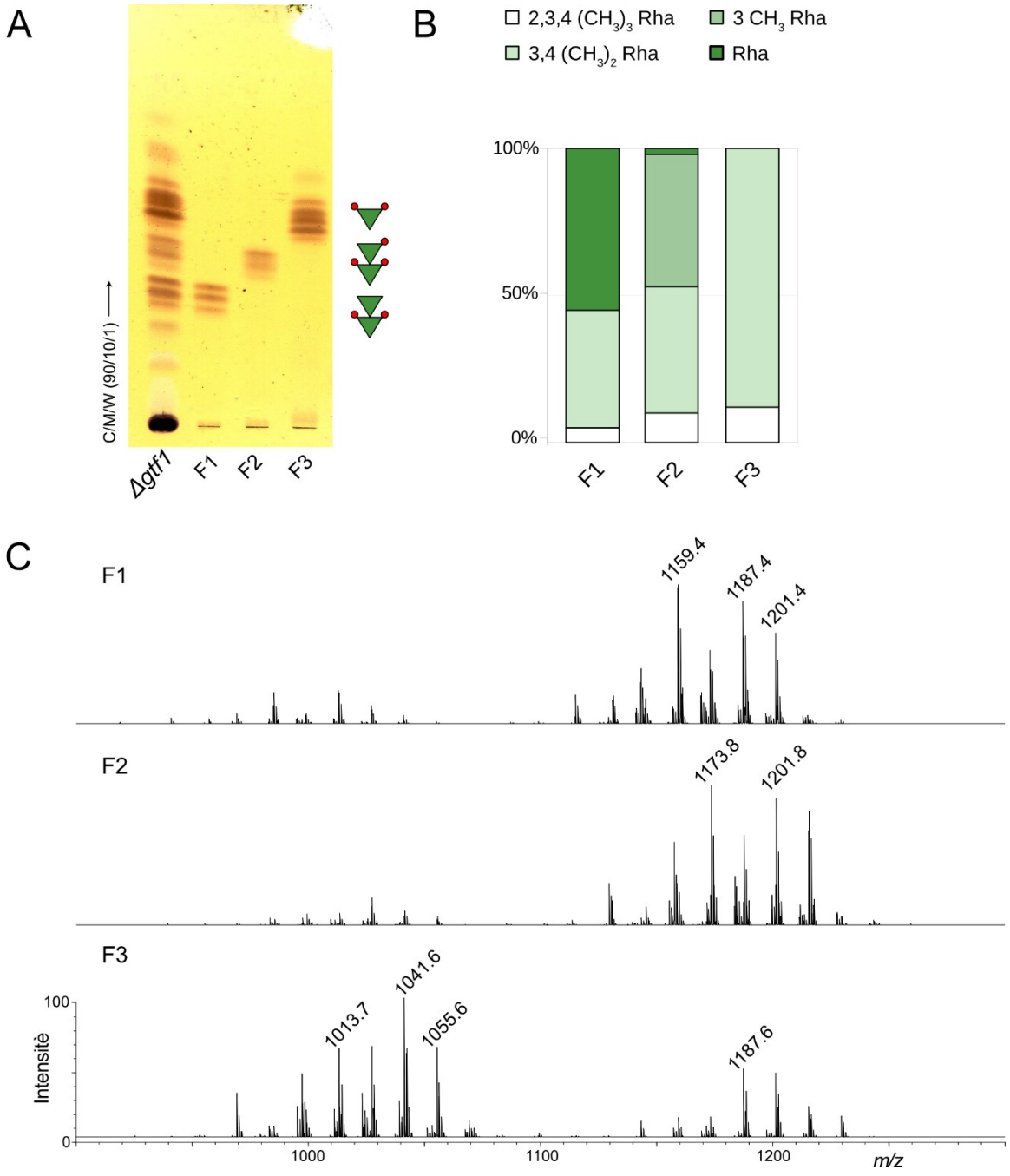


Figure 53 : Analyse des GPLs présents dans *Δgtf1*. Après purification de 3 fractions par CCM préparative (A) la composition en monosaccharides (B) confirme la présence de résidus de rhamnose plus ou moins méthylés influençant la polarité des GPLs. Les spectres de masse des 3 fractions (C) suggèrent que les formes dirhamnosylées sont majoritaires dans F1 et F2 mais que la forme monorhamnosylée est prépondérante dans F3.

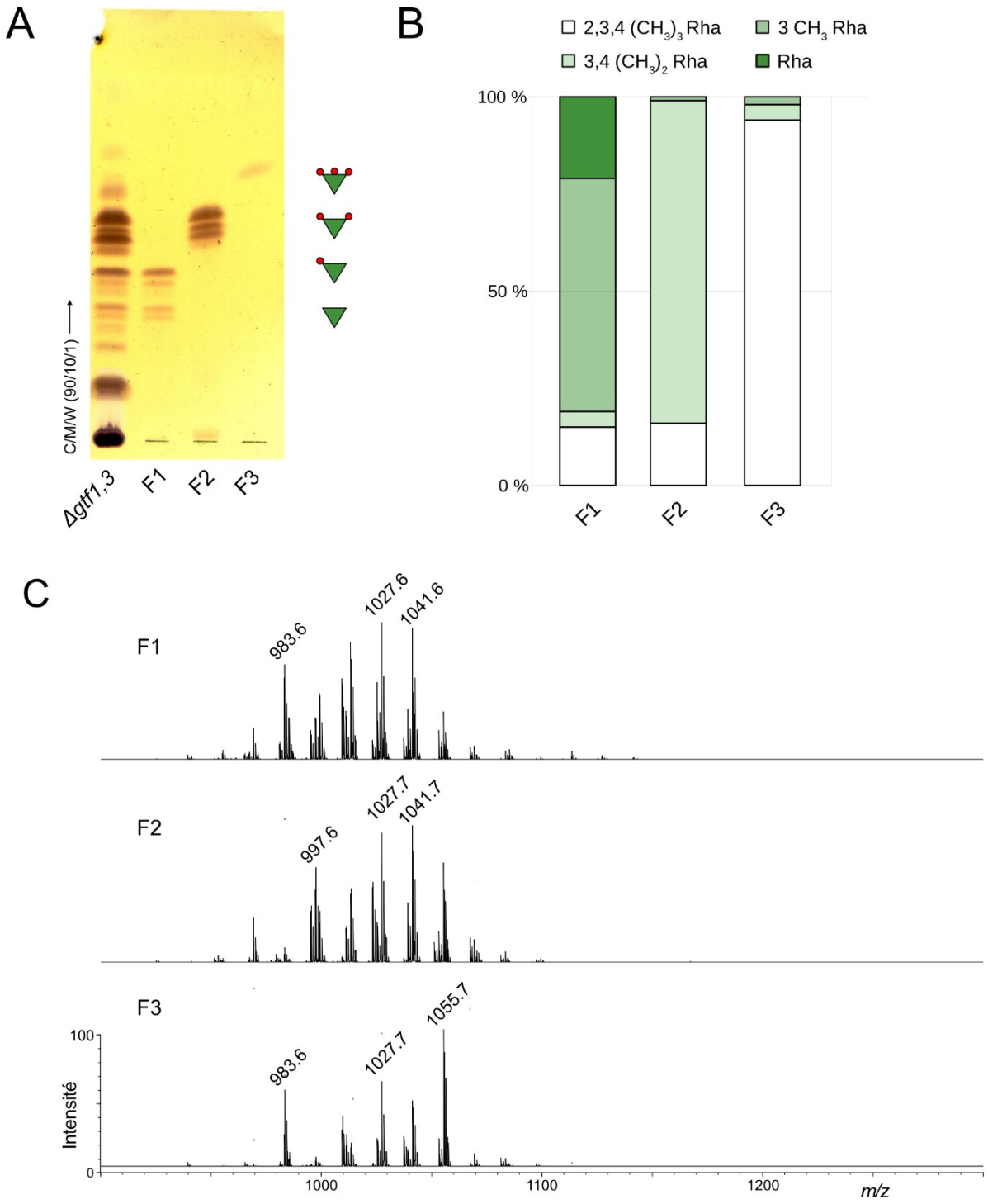


Figure 54 : Analyse des GPLs présents dans *Δgtf1,3*. Après purification de 3 fractions par CCM préparative (A) la composition en monosaccharides (B) confirme la présence de résidus de rhamnose plus ou moins méthylés influençant la polarité des GPLs. Les spectres de masse des 3 fractions (C) suggèrent que toutes les formes sont monorhamnosylées.

La détection simultanée par la CCM et la spectrométrie de masse a ainsi révélé la présence de cinq glycoformes de GPLs parmi les six possibles (Fig. 2A). De façon surprenante, la forme hypothétique GPL-1a monoglycosylée par le 6dTal sur le résidu de thréonine n'a pas pu être isolée même dans le mutant *Δgtf2,3*. Nous en avons déduit que le produit du gène *gtf1* n'est pas capable d'agir sur le lipopeptide, et que c'est le produit du gène *gtf2* qui initie la glycosylation (Fig. 2B) comme proposée pour son orthologue *GtfB* chez *Mav*[392]. D'autre part, j'ai également montré que le produit du gène *gtf3* ne convertit pas l'intégralité des GPLs en la forme triglycosylée GPL-3. En outre, le niveau de détection du lipopeptide par le PMA ne fut pas équivalent parmi les souches qui ne contenaient pas de GPLs glycosylés.

Dans ce projet, nous avons aussi évalué la contribution potentielle des GPL-1b, GPL-2b, GPL-2a et GPL-3 envers la phagocytose par les macrophages (Fig. 7). Pour cela, j'ai développé une méthode de couplage des GPLs à des micro-billes fluorescentes qui ont été utilisées par nos collaborateurs dans des tests de phagocytose. J'ai tout d'abord normalisé le dosage des GPLs avec le 3,4 di-*O*-méthyl rhamnose commun aux quatre glycoformes, afin d'utiliser une quantité identique de composés pour recouvrir des billes polymériques et fluorescentes. Ces dernières ont ensuite été introduites dans des cultures cellulaires de macrophages pour comptabiliser le nombre d'objets phagocytés. Les résultats ont montré que le GPL-2a et encore plus le GPL-3 favorisaient la phagocytose des billes contrairement aux GPL-1b et GPL-2b. D'autre part, l'utilisation d'anticorps dirigés contre le récepteur à mannose ainsi que le récepteur du complément CR3 inhibe significativement la phagocytose. Le même effet a été observé pour l'$AcPIM_2$ choisi comme contrôle positif.

Finalement, mes travaux analytiques sur l'étude de la glycosylation des GPLs de l'ensemble des mutants de glycosyltransférases a confirmé l'implication de chacune des trois glycosyltransférases *gtf1*, *gtf2* et *gtf3* dans l'addition du 6dTal sur la thréonine, puis des deux résidus de rhamnose sur l'alaninol, comme décrit par Myiamoto et ses collaborateurs chez *Msmeg*[335]. Alors que le premier résidu de rhamnose est nécessaire à la poursuite de la glycosylation, le rhamnose terminal n'est pas systématiquement présent. Néanmoins, sa présence favorise la phagocytose *in vitro*. En outre, mes travaux ont aussi montré que l'hétérogénéité des ions détectés en spectrométrie de masse n'était pas uniquement liée à une distribution des lipides mais à un possible remplacement de l'alaninol par le valinol. S'agissant de la morphologie, seul le mutant *Δgtf3* étant resté lisse, la présence de GPLs mais pas du 6dTal chez les mutants *Δgtf1* et *Δgtf1,3* nous a conduit à questionner l'implication de ce monosaccharide dans la morphologie devenue rugueuse et dans la capacité à former des cordes.

Manuscript Click here to view linked References

Di-rhamnosylated glycopeptidolipids promote mannose receptor-dependent internalization of *Mycobacterium abscessus*

Wassim Daher[1,5,&], Louis-David Leclercq[2,&], Matt D. Johansen[1,3], Jona Karam[1], Xavier Trivelli[6], Jérôme Nigou[4], Yann Guérardel[2,7,#], and Laurent Kremer[1,5,#,*]

[1]Centre National de la Recherche Scientifique UMR 9004, Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France.

[2]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.

[3]Centre for Inflammation, Centenary Institute and University of Technology Sydney, Faculty of Science, School of Life Sciences, Sydney, NSW, Australia.

[4]Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, CNRS, Université Paul Sabatier, Toulouse, France.

[5]INSERM, IRIM, 34293 Montpellier, France.

[6]Univ. Lille, CNRS, INRAE, Centrale Lille, Univ. Artois, FR 2638 - IMEC - Institut Michel-Eugène Chevreul, F-59000 Lille, France.

[7]Institute for Glyco-core Research (iGCORE), Gifu University, Gifu, Japan.

[#]For correspondence: E-mails yann.guerardel@univ-lille.fr, Tel: (+33) 3 20 33 63 47; laurent.kremer@irim.cnrs.fr, Tel: (+33) 4 34 35 94 47.

[*]Lead contact

[&]These authors contributed equally to this work.

HIGHLIGHTS

- Gtf1, Gtf2 and Gtf3 transfer deoxy-talose, the internal and extra rhamnose on GPL
- Gtf3 is required for efficient uptake of *M. abscessus via* the mannose receptor
- Deletion of Gtf1 or Gtf2 leads to rough corded bacilli and increased virulence
- GPL glycosylation controls *M. abscessus* surface properties and pathogenicity

SUMMARY

Mycobacterium abscessus is an emerging and difficult-to-manage mycobacterial species that exhibits smooth (S) or rough (R) morphotypes. Disruption of glycopeptidolipid (GPL) production results in transition from S to R and severe lung disease. A structure-activity relationship study was undertaken to decipher the role of GPL glycosylation in morphotype transition and pathogenesis. Deletion of *gtf3* uncovered the prominent role of the extra rhamnose in enhancing mannose receptor-mediated internalization of *M. abscessus* by macrophages. In contrast, the absence of the 6-deoxy-talose and the first rhamnose in mutants lacking *gtf1* and *gtf2*, respectively, only slightly affected *M. abscessus* phagocytosis but resulted in the S-to-R transition. Strikingly, *gtf1* and *gtf2* mutants displayed a strong propensity to form cords and abscesses in zebrafish, leading to robust and lethal infection. Together, these results underscore the importance and differential contribution of GPL monosaccharides in promoting virulence and infection outcomes.

KEYWORDS: *Mycobacterium abscessus*, morphotype, glycopeptidolipid, rhamnose, glycosyltransferase, macrophage, internalization, mannose receptor, zebrafish, virulence.

INTRODUCTION

Mycobacterium abscessus is a non-tuberculous mycobacteria (NTM), recognized as a multidrug-resistant species that has recently emerged as a major threat to individuals with underlying lung diseases such as chronic obstructive pulmonary disease (COPD) or cystic fibrosis (CF) (Floto et al., 2016; Johansen et al., 2020a). Infections with this fast-growing mycobacteria result in accelerated inflammatory lung damage (Esther et al., 2010), which may require surgical intervention (Ryan and Byrd, 2018). *M. abscessus* lung disease is difficult to treat due to intrinsic antimicrobial resistance, often leading to treatment failure despite prolonged multi-drug chemotherapy (Jarand et al., 2011; Johansen et al., 2020a; Kwak et al., 2019). Although originally thought to be acquired independently of the environment, the majority of individuals are infected with one of several dominant clones in circulation, indicating the presence of global transmission networks of *M. abscessus* and highlighting the possibility of human-to-human transmission (Bryant et al., 2016).

Distinguishable *in vitro* and *in vivo* properties have been assigned to smooth (S) or rough (R) variants of *M. abscessus* and include changes in sliding motility, biofilm formation, drug susceptibility (Bernut et al., 2016; Gutiérrez et al., 2018; Howard et al., 2006; Johansen et al., 2020b; Madani et al., 2019; Ryan and Byrd, 2018), bacterial surface hydrophobicity (Daher et al., 2020; Viljoen et al., 2018, 2020), aggregation, and cord formation (Bernut et al., 2014a, 2016; Howard et al., 2006; Viljoen et al., 2018). Other distinctive phenotypes influencing clinical disease outcome are related to intracellular interactions and trafficking in macrophages (Roux et al., 2016; Whang et al., 2017) or induction of potent pro-inflammatory responses (Roux et al., 2011). Numerous studies have highlighted the increased pathogenesis of the R variant as compared to the S variant (Bernut et al., 2014a, 2016, 2017; Catherinot et al., 2007). In this context, the zebrafish infection model has received considerable attention and represents a genetically tractable host-pathogen pairing for dissecting *M. abscessus* interactions with host immune cells (Bernut et al., 2015). Zebrafish are particularly susceptible to *M. abscessus* R infection, as evidenced by hyperinflammatory responses, increased bacterial burden, formation of massive serpentine cords and abscesses, resulting in enhanced larval mortality (Bernut et al., 2014a). These observations are consistent with epidemiological surveys, which showcase the prominence of the R variant in patients with severe lung infections and chronic airway colonization in CF patients (Catherinot et al., 2009; Jönsson et al., 2007). These findings underscore the clinical relevance of morphological distinctions between S and R morphotypes, and the need to identify the molecular mechanisms responsible for S-to- transition.

The transition from S to R relies on the production of high or low/absent levels of surface-associated glycopeptidolipids (GPL), respectively (Bernut et al., 2016; Gutiérrez et al., 2018; Howard et al., 2006; Medjahed et al., 2010). Genomic and transcriptomic studies have unraveled the presence of DNA insertions or deletions occurring in R forms, mainly in *mps1* and *mps2* encoding non-ribosomal peptide synthases responsible for the biosynthesis of the GPL peptide backbone, and *mmpL4a*/*mmpL4b* encoding membrane proteins involved in GPL transport (Bernut et al., 2016; Dedrick et al., 2021; Park et al., 2015; Pawlik et al.,

2013). Structurally, *M. abscessus* GPL comprises a peptide core consisting of an L-alaninol-linked D-phenylalanine-D-*allo*threonine-D-alanine tripeptide assembled by Mps1 and Mps2, which is acylated with a C_{24}-C_{33} 3-hydroxy/methoxy fatty acid. This lipopeptide is further decorated by a glycosylation arrangement that is built from *O*-methylated and *O*-acetylated deoxy hexoses (Daher et al., 2020; Ripoll et al., 2007; Whang et al., 2017). The lipopeptide is glycosylated with *allo*-Thr linked to a 6-deoxy-α-L-talose (6-dTal) while the terminal alaninol is linked to an α-L-rhamnose (Rha), producing the less-polar diglycosylated GPL species (GPL-2). In addition to GPL-2 that contain a 3,4-di-*O*-acetylated 6-dTal and a 3,4-di-*O*-methylated Rha, *M. abscessus* synthesizes more polar GPL species (GPL-3) *via* the addition of a 2,3,4-tri-hydroxylated Rha to the alaninol-linked 3,4-di-*O*-methylated Rha (Daher et al., 2020; Ripoll et al., 2007; Whang et al., 2017). GPL are heterogenous in structure and vary in length of the fatty acyl chain and the degree of hydroxylation or *O*-methylation of the glycosidic moieties through the action of *O*-methyltransferases, *O*-acetyltransferases and *O*-glycosyltransferases. However, the exact functions of these enzymes with respect to the GPL structure and their possible contribution to the pathogenicity of *M. abscessus* have not been addressed experimentally.

Many important interactions between bacterial pathogens and their host cells are highly specific receptor interactions that involve carbohydrates. Glycan interactions can mediate adhesion, invasion and immune evasion and be used by the pathogen to establish a successful infection (Poole et al., 2018). The question of whether GPL could take part in the adhesion and internalization of mycobacteria within cells has been proposed in previous studies. Pre-incubation of cells with blocking antibodies to complement receptor 3 (CR3) and mannose receptor (MR) partially abrogated the uptake of GPL-coated beads as well as phagocytosis of *Mycobacterium smegmatis* and *Mycobacterium avium*, which are GPL-producing NTM (Villeneuve et al., 2003, 2005). However, the structural elements of GPL that participate in recognition by CR3 and MR remain unknown. Furthermore, no information is available regarding the individual or collective contribution of the different GPL monosaccharides in the direct interaction with cell receptors during the early steps of interaction with host receptors, particularly within the context of *M. abscessus* pathogenesis.

Herein, we investigated the glyco-interactions of GPL with host cells by combining genetic, structural and biological studies. We performed a stepwise dissection of the three *gtf* genes that encode putative GPL glycosyltransferases in *M. abscessus.* This led to a panel of mutants that attempted to address the requirements and contribution of each monosaccharide in bacterial morphotypes, bacterial surface properties, adhesion and internalization of bacilli by human macrophages as well as pathogenicity in the zebrafish infection model. Synthesis of a GPL variant lacking 6-dTal and/or the first Rha in *gtf1* and *gtf2* mutants, respectively, was associated with an S-to-R morphotype transition showing hypercording and enhanced virulence and pathogenicity in zebrafish. We also provide evidence that the terminal Rha in GPL-3 enhances *M. abscessus* internalization by macrophages in a MR-dependent manner. Overall, this in-depth dissection of the GPL glycome provides new insights into the biological function and importance of the

different monosaccharides in the establishment of infection and disease pathology in this emerging human pathogen.

RESULTS

gtf1, *gtf2* and *gtf3* encode specific glycosyltransferases decorating GPL in *M. abscessus*.

The *gpl* locus encompasses three genes encoding putative glycosyltransferases, designated *gtf1-gtf3* **(Figure S1A)**. To investigate the contribution and biological functions of the corresponding glycosyltransferases in the GPL structure, *gtf1-gtf3* were deleted either individually or simultaneously in the GPL-producing S variant of *M. abscessus,* using a recently developed unmarked deletion system that allows for multiple gene deletion in *M. abscessus* (Richard et al., 2019) **(Figure S1A)**. PCR/sequencing performed on the parental and mutant strains using the primers in **Table S1** confirmed the proper genotype of the various *gtf* mutants **(Figure S1B)**. Genetic complementation of Δ1, Δ2 and Δ3 was done through specific integration at the *attB* chromosomal site (Stover et al., 1991) of a copy of *gtf1*, *gtf2* or *gtf3* with an HA tag at the 3'-end and placed under the control of the *hsp60* promoter (for *gtf1*) or their endogenous promoters (for *gtf2* and *gtf3*) **(Figure S1C)**. Probing the crude lysates from the various strains using anti-HA antibodies revealed a single band of the expected size, corresponding to Gtf1-HA, Gtf2-HA or Gtf3-HA proteins, validating complementation of the single mutants **(Figure S1D)**.

M. abscessus S synthesizes both diglycosylated (GPL-2a) and triglycosylated (GPL-3) GPL, containing one and two partially *O*-methylated Rha residues, respectively, in addition to the diacetyl-*O*-6-dTal residue **(Figure 1A)** (Daher et al., 2020). Thin layer chromatography (TLC) analysis of the polar fractions revealed a complex pattern of polar glycolipids, tentatively assigned to GPL-2a and GPL-3, all absent in *M. abscessus* R (Daher et al., 2020; Ripoll et al., 2007) **(Figure 1B)**. Monosaccharide compositional analysis confirmed the presence of the deoxyhexoses 6-dTal, Rha, and 3,4-di-*O*-methyl Rha in *M. abscessus* with a low proportion of 3-*O*-methyl Rha, suggesting the presence of minor isomers of GLP-2a and GPL-3, in which 3,4-di-*O*-methyl Rha is replaced by 3-*O*-methyl Rha **(Figure 1C)**. Mass spectrometry (MS) analysis showed two major signal patterns dominated by ion pairs at *m/z* 1258/1286 and 1404/1432, corresponding to GPL-2a and GPL-3, respectively (**Figure 1D**) (Daher et al., 2020). The presence of both acetyl groups on 6-dTal was confirmed by MS and MS/MS of polar glycolipids treated with mild alkali, resulting in a decrease of 84 m.u. at *m/z* 1174-1202 for GPL-2a and at *m/z* 1320-1348 for GPL-3 on the MS spectrum **(Figure S2A).** Comparison of the MS/MS fragmentation patterns of native and alkali-treated glycolipids confirmed the structure of GPL-3 and GPL-2a and the localization of acetyl groups on 6-dTal **(Figures 1E, 1F, S2B** and **S2C)**. Finally, MS analysis of individual bands (data not shown) confirmed the migration pattern of GPL on a TLC plate **(Figure 1B)**.

Deletion of *gtf3* leads to accumulation of GPL-2a whereas deletion of *gtf1* is associated with new GPL-1b and GPL-2b glycoforms.

TLC analysis showed that Δ3 accumulates a set of glycolipids, tentatively related to GPL-2a **(Figure 1B)**. MS analysis showed a pattern dominated by ion pairs at *m/z* 1258/1286 in the native fraction **(Figure 1D)** and *m/z* 1174-1202 on a saponified fraction **(Figure S2A)** that were identified as GPL-2a based on the MS/MS fragmentation pattern (data not shown). The absence of Rha in the polar glycolipid fraction confirmed that Δ3 fails to synthesize GPL-3 **(Figure 1C)**.

Δ1 and Δ1,3 synthesize a set of glycolipids different from the WT strain, characterized by the absence of GPL-2a and GPL-3 **(Figure 1B, 1C *and* 1D)**. MS analysis of Δ1 revealed two clusters of signals with a difference of 146 m.u centered at *m/z* 1188/1202 and 1014/1042 **(Figure 1D)**. MS/MS analysis of the characteristic parent ion at *m/z* 1188 established the presence of a terminal Rha-$(Me)_2$Rha-disaccharide through the observation of fragment ions M-146 at *m/z* 1042 and secondary ions at *m/z* 517/371 574/428 763/617 without loss of $(Ac)_2$dTal **(Figure S3A)**. MS/MS analysis of the characteristic parent ion at *m/z* 1042 established the presence of a terminal $(Me)_2$Rha residue owing to the observation of the fragment ion M-174 at *m/z* 867 and the absence of Rha and $(Ac)_2$dTal **(Figure S3B)**. MS analysis of saponified GPL did not show an 84 m.u. decrease, further demonstrating the absence of $(Ac)_2$dTal **(Figure S2A)**. Altogether, these analyses indicate that Δ1 synthesizes two new GPL isomers devoid of $(Ac)_2$dTal, designated GPL-2b and GPL-1b, substituted or not by a terminal Rha on the $(Me)_2$Rha. MS analysis of Δ1,3 showed a single cluster of signals with similar *m/z* values to GPL-1b **(Figure 1D)**. The MS/MS fragmentation pattern of the signal at *m/z* 1042 confirmed the identification of GPL-1b (**Figure S3C**). The presence of a single Rha was confirmed by homo- and heteronuclear NMR spectroscopy, as exemplified by the assignment of glycan and peptide moieties on the ^{1}H-^{13}C HSQC spectrum **(Figure S3D)**. In agreement with the compositional analysis (**Figure 1C**), two isomers of mono-glycosylated GPL-1b substituted with Rha or MeRha instead of $(Me)_2$Rha were also identified by MS/MS in the lipid polar fraction isolated from Δ1,3 (**data not shown** and **Figure 1C**).

Non-glycosylated lipopeptide intermediates accumulate in the *gtf2* mutant.

The absence of orcinol-sensitive bands on the TLC plate and the lack of GPL signal on the MS spectra implies that deletion of *gft2* abrogates GPL synthesis (**Figure 1B** and **1C**). However, staining with phospho-molybdic acid showed the accumulation of two compounds in Δ2 and Δ1,2, tentatively assigned to non-glycosylated GPL biosynthetic intermediates **(Figure S3E)**. The MALDI-TOF spectrum of the lower R_f polar lipid in Δ1,2 showed two intense ions at *m/z* 838 and 866 that can be attributed to the unmethoxylated C28 and C30 tripeptide bound to an alaninol residue **(Figures 1D** and **S3F)**. The ^{1}H-^{1}H COSY NMR spectrum confirmed the presence of three amino acids Ala, Thr and Phe as well as alaninol **(Figure S3G)**. Similar analyses of the polar

lipid with a higher R_f of Δ1,2 indicated the presence of the unmethoxylated C28 and C30 tripeptide linked to a valinol residue (data not shown).

Overall, structural analyses of the various *gtf* mutants identified at least 5 distinct GPL glycoforms **(Figure 2A)**. Similar to previous studies (Howard et al., 2006; Lee et al., 2017; Ripoll et al., 2007), we show here that GPL-3 and GPL-2a are predominantly synthesized in the WT S strain but absent in the R variant. While only GPL-2a is found in Δ3, the newly described GPL-1b and GPL-2b accumulated in Δ1. Δ2 and Δ1,2 do not produce GPL but accumulate non-glycosylated lipopeptide intermediates. Importantly, only GPL-1b was detected in Δ1,3 which implies that Gft1 could not transfer $(Ac)_2$dTal onto the unglycosylated lipopeptide and that the transfer of $(Me)_2$Rha by Gft2 precedes Gft1 activity, similar to *M. avium* (Eckstein et al., 2003) **(Figure 2B)**. In these two NTM species, Gtf2 in *M. abscessus* and its ortholog in *M. avium*, first catalyze the linkage between Rha and the branched-chain amino-alcohol, alaninol. Next, the addition of the second Rha to the first Rha or of 6-dTal to the threonine residue are accomplished by Gtf1 or Gtf3.

Deletion of *gtf1* or *gtf2* affects morphological and physiological properties of *M. abscessus in vitro*.

While it is well established that mutations in genes involved in peptide core production (*mps1* and *mps2*) or in transport (*mmpL4a* and *mmpl4b*) of GPL are responsible for the morphological S-to-R transition (Johansen et al., 2020a), the role of the different monosaccharides decorating the GPL backbone in this process remains unknown. We, therefore, appraised whether *gtf* deletions had an impact on colony morphology. Observations of individual colonies on Tryptic Soy agar show that, in contrast to the parental WT S progenitor, all mutants lacking either *gtf1* or *gtf2* (Δ1, Δ2, Δ1,2, Δ1,3, Δ2,3, and Δ1,2,3) formed rough, corded colonies **(Figure 3A)**. While functional complementation of Δ1 restored a smooth morphotype (Δ1_C), this was not the case for Δ2, despite apparent recovery of GPL production in Δ2_C **(Figure 3A)**. In contrast, deletion of *gtf3* did not affect the morphology of the strain, which remained smooth **(Figure 3A)**.

Partitioning of mycobacterial pellets between hexadecane and aqueous buffer has been used as a quantitative marker of cell surface hydrophobicity in mycobacteria to associate low pathogenicity with reduced hydrophobicity (Jankute et al., 2017). To determine whether the GPL glycan moieties affect the surface hydrophobicity, aqueous hexadecane-buffer partitioning was applied to the S, R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3 strains. The Δ*fmt* mutant with a smooth phenotype, lacking the *O*-methyltransferase of the GPL lipid moiety, was included as a highly hydrophilic control strain (Daher et al., 2020). The complemented Δ*fmt* strain (Δ*fmt*_C) displaying parental hydrophobicity levels, was also added in this assay. Consistent with their colony morphotype, S, Δ3_C, and Δ*fmt*_C were more hydrophilic than the R, Δ1, Δ2, Δ1,2, Δ1,3, Δ2,3, and Δ1,2,3 strains, while Δ1_C and Δ2_C showed partial phenotype restoration **(Figure 3B)**. Only Δ3 and Δ*fmt* were significantly more hydrophilic than the parental S strain **(Figure 3B)**, suggesting that removal of 6-dTal and/or the internal Rha increases the surface hydrophobicity of *M.*

abscessus. This, however, does not affect the *in vitro* growth of the strains, indicating that deletion of the *gtf* genes does not impact the replication rate of *M. abscessus* in planktonic culture **(Figure S4A)**.

Previous studies demonstrated that smooth strains of *M. abscessus* display a higher capacity to slide on agar and a lower propensity to aggregate compared to rough strains (Deshayes et al., 2005; Martínez et al., 1999). We thus measured the sliding distances of the *gtf* mutants on 7H9 supplemented by 0.3% agar. All rough strains (except Δ2_C) showed significantly reduced motility compared to smooth strains **(Figure 3C)**. Consistently, all cord-forming mutants (rough strains) produced intense bacterial aggregates at the bottom of the glass tubes, while smooth strain cultures remained homogenous and turbid **(Figure S4B)**. This indicates that the *gtf1*- and *gtf2*-deletion mutants were associated with low sliding and high aggregation capacities. Whether this affects their ability to form biofilms was investigated using the Crystal Violet (CV) assay. The CV absorbance at 570_{nm} was ~2 fold higher for rough strains (except for Δ2_C) compared to smooth strains **(Figure 3D)**. In contrast, the CFU counts in the biomass formed at the bottom of the glass tubes were much higher in the smooth strains **(Figures 3E** and **S4D)**. Examination of biofilm formation in 6-well plates revealed that they were visible as pellicles at the air-liquid interface. The pellicles of the smooth variants appeared more filamentous and reticulated while those of rough variants appeared thinner **(Figure S4C)**.

We next investigated whether the glycan moieties affect the drug susceptibility profile of *M. abscessus* towards a panel of clinically used antibiotics. Although S and R strains shared comparable minimal inhibitory concentrations (MIC) to cefoxitin, imipenem, bedaquiline and amikacin, all R strains exhibited a 4-fold increase in susceptibility to rifabutin (MIC = 12.5 µg/mL) compared to S strains (MIC = 50 µg/mL) **(Figure S4E)**. These results highlight the importance of GPL glycosylation in resistance to rifabutin.

Deletion of *gft1* and *gtf2* leads to large extracellular cords and abscesses in zebrafish embryos.

The zebrafish embryo has recently emerged as an excellent model for studying the *in vivo* pathophysiology and drug susceptibility traits of NTM species, and in particular *M. abscessus* (Bernut et al., 2014b, 2014a; Johansen and Kremer, 2020). Embryos are optically transparent and display functional innate immunity from 24 hrs post-fertilization, allowing visualization of host-pathogen interactions in real-time using fluorescent bacterial reporters (Cronan and Tobin, 2014). Furthermore, exploitation of the zebrafish model uncovered the dissection of S and R variant virulence traits, with the R variant forming large serpentine bacterial cords that lead to abscess formation and tissue destruction, while the S variant leads to a chronic and comparatively avirulent disease phenotype (Bernut et al., 2014a). Thus, in this study, we sought to determine whether Δ1 and Δ2 maintain their rough phenotype in an *in vivo* setting, and whether they similarly display the pathophysiological signs of the R variant of *M. abscessus* infection. Infection of embryos with ~250-300 CFU of *M. abscessus* resulted in significant embryonic mortality in the WT R-infected, Δ1 or Δ2 groups compared to all other sets **(Figure 4A)**. Importantly, embryos infected with the complemented Δ1 or Δ2 strains, exhibited similar virulence to WT S strain. Furthermore, there was a significant increase in mortality in Δ2-

infected embryos compared with WT R, highlighting that this mutant displays greater virulence in the zebrafish embryo model. To further dissect the pathophysiological traits of *gtf* mutants compared to their WT counterpart, we examined the kinetics of extracellular cord and abscess formation in real-time using a fluorescent microscope. When examining extracellular cord formation, we observed that there was a significantly greater proportion of embryos with cords in wild-type R, Δ1 or Δ2 mutants compared with the parental WT S and corresponding complemented strains. This phenotype was apparent as early as 2 days post-infection (dpi), and was even more pronounced at 4 dpi **(Figure 4B)**. We observed significantly lower proportions of embryos with cords in the Δ1_C or Δ2_C complemented strains compared to their corresponding parental mutant, demonstrating that complementation abrogates cord formation in these mutants and further emphasizing that cord formation is a critical driver of increased virulence in these mutants. When we examined the total number of cords per embryo in each group, we identified a similar number of cords per embryo in the wild-type R, Δ1 or Δ2 groups, which differed significantly from the WT S and the complemented strains **(Figure 4C)**. Abscess formation is strongly correlated with cord formation, often arising due to unrestrained extracellular cord expansion leading to substantial tissue destruction at infection sites (Bernut et al., 2014a). As such, we further examined whether abscess formation differed in Δ1 or Δ2 compared with WT strains. We observed significantly increased abscess formation in the Δ2 group compared with the parental WT S and the complemented Δ2 groups as early as 2 dpi, highlighting a plausible mechanism for the increased virulence of this strain compared with the WT R **(Figure 4D** and **4F)**. At 4 and 6 dpi, we identified significantly greater proportions of embryos with abscesses in the wild-type R, Δ1 or Δ2 groups compared with the parental WT S and corresponding complemented strains. When we further stratified abscess formation in embryos, we identified that WT R, Δ1 or Δ2 groups displayed an increased number of abscesses per embryo compared with wild-type S and complemented strains. This confirms that abscess formation strongly correlates with the presence of extracellular cord formation in the zebrafish embryo **(Figure 4B-4F)**. Overall, using the zebrafish embryo, we can conclude that deletion of *gtf1* or *gtf2* leads to a stable rough phenotype that is preserved in an *in vivo* setting and displays the pathophysiological hallmarks of WT R infection, such as the large extracellular cord and abscess formation.

Strains lacking GPL production form cords that represent a typical marker and trait associated with acute infection and disease severity (Bernut et al., 2014c; Gutiérrez et al., 2018; Howard et al., 2006). The question remains whether smaller modifications, such as glycosylation, in the GPL structure induce cord formation, both within and outside the macrophage. To address this question, human THP-1 macrophages were individually infected with all strains of *M. abscessus* and the fate of the bacilli was examined at 3 dpi. Intracellular and extracellular giant cords were observed with all strains lacking either *gtf1* and/or *gtf2* **(Figure S5)**. Complementation of Δ1 inhibits cord formation in this strain, whereas complementation of Δ2 retains the ability to generate these bacterial structures **(Figure S5)**. These observations suggest a direct link between the absence of 6-dTal or the first Rha in cord formation.

6-dTal and di-Rha are required for *M. abscessus* uptake by macrophages.

Glycans displayed at the surface of microbial pathogens play an important role in mediating early interactions between microbes and the host, facilitating the first steps of infection. We investigated the impact of GPL glycosylation on the adhesion and invasion of *M. abscessus* by human THP-1 macrophages. Cells were infected with *M. abscessus* S, R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Δ1,2, Δ1,3, Δ2,3 or Δ1,2,3 expressing tdTomato for 3 hrs at a multiplicity of infection (MOI) of 2:1, prior to assessing phagocytosis and intracellular growth. After infection and treatment with 250 µg/mL amikacin to prevent extracellular bacterial growth, macrophages were lysed at 3, 24, and 72 hrs post-infection (hpi) and plated to determine intracellular bacterial burden. At day 0 (3 hpi), the invasion rate of single, double and triple mutants was reduced by ~ 0.5 log compared with the parental WT S, and this effect was rescued in all complemented strains (Δ1_C, Δ2_C and Δ3_C) **(Figure 5A)**. This decrease in bacterial burden was maintained at 24 and 72 hpi **(Figure 5A)**. In parallel, at 3, 24, and 72 hpi, macrophages were stained with anti-CD43 and DAPI. Microscopic observations and quantification of infected cells revealed a marked reduction in the number of THP-1 cells infected with single, double and triple mutants immediately after phagocytosis. This effect was also maintained over time compared with cells infected with WT progenitors or complemented strains **(Figure 5B** and **5C)**. To assess the consequences of the invasion defect on intracellular bacterial loads, infected macrophages were classified into three categories at 3 hpi based on their bacterial content: weakly infected (<5 bacilli), moderately infected (5−10 bacilli) and heavily infected (>10 bacilli), as illustrated in **Figure 5D**. This quantitative analysis indicates that deletion of the *gtf* genes in single, double and triple mutants profoundly reduced the percentage of moderately and heavily infected THP-1 cells while increasing the proportion of the weakly infected category **(Figure 5C** and **5D)**. Together, this indicates that i) GPL are involved in *M. abscessus* internalization by macrophages, and ii) deoxy-talose and extra rhamnose are both required for optimal internalization.

As a first step to characterize the putative receptor(s) involved in *M. abscessus* entry into macrophages, we performed binding competition experiments. Since GPL monosaccharides are involved in the uptake of *M. abscessus* by macrophages and *M. smegmatis* GPL have previously been reported to interact with the macrophage Mannose Receptor (MR) and the lectin domain of Complement Receptor 3 (CR3) (Villeneuve et al., 2005), we investigated the possible involvement of three lectins, MR, CR3, and Dectin-1, a β-glucan-binding lectin (Brown and Gordon, 2001) like CR3. Therefore, cells were pre-treated 1 hr prior to infection with 100 µg/mL of soluble ligands of each of the three lectins: mannan for MR (Shepherd et al., 1983), laminarin for Dectin-1 and CR3 (Gantner et al., 2005; Xia and Ross, 1999), or methyl-α-D-glucopyranoside for CR3 (Lin et al., 2010). At 3 hpi, intracellular bacterial loads were assessed by determination of CFU but no significant impact of the various pre-treatments was observed on the ability of the different strains, S, Δ1, Δ2, Δ3 or Δ1,3, to invade THP-1 cells under the conditions tested **(Figure S6)**. These competition assays do

not establish the participation of MR, CR3, and Dectin-1 in the internalization of *M. abscessus* S by macrophages. This could be due to the heterogenous composition of GPL, consisting of both diglycosylated (GPL-2) and triglycosylated (GPL-3) forms of GPL, which mitigates the potential involvement of GPL-3 in bacterial uptake (Daher et al., 2020).

Di-rhamnosylation of GPL enhances the uptake of *M. abscessus* by macrophages.

To assess the contribution of GPL-3 in *M. abscessus* uptake, we tipped the GPL-2/GPL-3 balance in favor of GPL-3. To do so, we cloned the *gtf3* gene in fusion with an HA tag in pMV261 (pMV261-*gtf3-HA*), which was used to generate the Gtf3_ov strain. Overproduction of Gtf3-HA was confirmed by Western blotting using anti-HA antibodies. In Gtf3_ov, the level of Gtf3 production was greatly increased compared with the level in Δ3_C carrying the integrative pMV306-*gtf3-HA*, unlike KasA **(Figure 6A)**. To test whether Gtf3 overexpression increased the proportion of the GPL-3/GPL-2 ratio, whole polar lipids were extracted and analyzed by TLC. The GPL profile of Gtf3_ov was shifted to bands migrating closer to the origin, as would be expected for the accumulation of GPL-3 subspecies with a concomitant reduction in GPL-2 subspecies **(Figure 6B)**. This indicates that in Gtf3_ov, most of the GPL-2 forms were converted to the corresponding GPL-3 forms, making this strain particularly useful for further investigate the role of the extra Rha in functional assays.

Gtf3_ov remained smooth **(Figure 3A),** grew similarly to the parental S strain in planktonic cultures **(Figure S4A)**, was more hydrophobic than the parental CIP104536^{T} (S) strain **(Figures 3B** and **S4B)** while exhibiting reduced sliding motility **(Figure 3C)** and produced high biomass of bacteria **(Figure 3E)**. Next, we interrogated whether increasing the GPL-3/GPL-2 ratio influences the early interactions between *M. abscessus* and macrophages. THP-1 macrophages were infected with either the parental S strain producing both GPL-2 and GPL-3 glycoforms or with Gtf3_ov producing mainly GPL-3. The uptake of Gtf3_ov was significantly increased (0.8 log at 3 hpi and >1 log at 24 and 72 hpi) compared with parental *M. abscessus* **(Figure 6C)**. Similar results were obtained using primary human monocyte-derived macrophages at 3 hpi **(Figure 6D)**. Moreover, the percentage of infected macrophages was increased 2-fold with Gtf3_ov compared to *M. abscessus* S **(Figure 6E** and **6F)**. Furthermore, the category of infected macrophages containing >10 bacilli increased significantly at 3, 24 and 72 hpi **(Figure 6E** and **6G)**. Importantly, internalization of Gtf3_ov by macrophages pre-treated with mannan, laminarin or methyl-α-D-glucopyranoside was inhibited by 35%, 45% and 50%, respectively **(Figure 6H)**, suggesting that MR, Dectin-1, and/or CR3 may be involved in the recognition of the recombinant strain and in the interaction with GPL-3. Collectively, these results suggest that early interactions of Gtf3_ov with and phagocytosis by macrophages are conditioned by the extent of GPL rhamnosylation and are dependent on the lectin receptor(s).

Enhanced phagocytosis of GPL-3-coupled beads is mediated by C-type lectin and β-glucan receptors.

The importance of GPL-3 in the early interaction of *M. abscessus* with host cells prompted us to conduct an in-depth analysis to evaluate the contribution of each GPL monosaccharide in phagocytosis. To eliminate interference from other bacterial surface molecules, four truncated forms of GPL produced in various *gtf* mutants were purified and coupled to green fluorescent beads (1 µm): GPL-1b substituted by a single Rha, GPL-2a substituted by 6-dTal and a single Rha, GPL-2b characterized by di-rhamnoside chain only and full GPL-3 comprising 6-dTal and di-Rhamnoside **(Figure 2A)**. Uncoated beads were included as negative control. The ability of the lipid-coated beads to enter macrophages was analyzed by pre-incubating THP-1 cells (~10 beads/cell) for 4 hrs, followed by immunofluorescence imaging. Quantification of the percentage of bead-containing cells clearly demonstrates that macrophages internalized significantly more GPL-2a-coaded beads (~ 2-fold), GPL-3-coaded beads (~ 5-fold) than GPL-1b- or GPL-2b-coaded beads **(Figure 7A** and **7B)**. Furthermore, THP-1 macrophages that phagocytosed GPL-3-coated beads contained an average of 4 beads/cell, whereas the average remains around 2 beads/cell for all other lipid-coated and uncoated beads **(Figure 7A** and **7C)**. High-resolution confocal 3D imaging confirmed the intracellular localization of the counted beads **(Figure S7A, Movies 1** and **2)**. Structure-function comparison of GPL-1b-, GPL-2a- and GPL-2b-conjugated beads indicates that the lack of 6-dTal or the extra Rha impairs the efficient uptake of GPL-coated beads **(Figure 7B)**, in agreement with results obtained with live bacteria.

Interestingly, uptake of GPL-2a- and GPL-3-beads was partially inhibited by mannan, laminarin or methyl-α-D glucopyranoside **(Figure 7D)**. Pre-incubation of cells with these compounds also significantly reduced the average number of GPL-3 beads/macrophage **(Figure S7B)**. To determine more specifically which lectin receptor(s) are involved in GPL binding, we performed antibody blocking experiments, using specific antibodies directed against MR, CR3 or Dectin-1. We found that phagocytosis of GPL2a- and GPL3-coupled beads was primarily dependent on MR **(Figures 7E, 7F** and **S7C)**. Phagocytosis of GPL3-coated beads was also dependent on CR3, although to a weaker extent **(Figure 7E** and **7F)**. No differences were observed for the uptake of GPL-1b or GPL-2b beads with any of the antibodies used **(Figure 7F)**. Quantitative measurements showed that the number of GPL-2a- or GPL-3-beads per macrophage is significantly reduced in the presence of MR and/or CR3 blocking antibodies **(Figure 7G)**. Overall, these results support a primarily role of MR in the recognition of GPL carbohydrates.

DISCUSSION

Most studies have reported the effects of near-total loss of GPL in *M. abscessus*, portraying an incomplete picture of the functions of these lipids in the pathogenesis of this emerging human pathogen while the structural elements related to these functions remain elusive. This study provides an in-depth overview of the structure-function relationship of GPL glycosylation and glyco-interactions with host receptors. We show

that in *M. abscessus*, Gtf1 transfers a 6-dTal to the threonine of the peptide core, while Gtf2 transfers Rha to alaninol and Gtf3 adds an extra Rha to the alaninol-bound Rha, similar to *M. smegmatis* (Miyamoto et al., 2006). Importantly, major and unexpected *in vitro* and *in vivo* phenotypes were uncovered using a set of genetically-defined *gft* mutants. Results indicate that GPL glycosylation 1) influences the hydrophilicity of the mycobacterial cell surface; 2) sustains the S morphotype; 3) facilitates the internalization of bacilli by macrophages; and 4) attenuates the virulence of *M. abscessus* in macrophages and in zebrafish embryos.

Low GPL production in *M. abscessus* results in a rough, hydrophobic strain, promoting bacterial aggregation and a reduced ability to form biofilms and slide (Daher et al., 2020; Viljoen et al., 2018). Lack of *O*-methylation of the GPL fatty acid chain, caused by *fmt* inactivation, increases the hydrophilicity of *M. abscessus* (Daher et al., 2020). Here, we demonstrate that loss of the 6-dTal or the internal Rha generates highly hydrophobic variants, with phenotypes similar to the WT R strain. Atomic force microscopy revealed that the surface of smooth bacilli comprises both hydrophilic and hydrophobic cell surface nanodomains (Viljoen et al., 2020). Therefore, it might be speculated that glycosylation of GPL present in hydrophilic nanodomains may mask important components required for the S-to-R transition, whereas specific loss of 6-dTal or total glycan moiety may expose these components, thereby affecting the cell surface composition and hydrophilic properties characterizing the S variant. Contrarily, the absence of terminal Rha residue does not affect drastically the hydrophilic balance, suggesting a tight regulation of the surface of bacilli. Thus, in addition to the chemical changes affected in the lipid portion of the GPL (Daher et al., 2020), these data point to an essential role of GPL glycosylation in bacterial surface properties.

Comparative studies between S and R variants highlighted the positive correlation between low GPL production and increased hydrophobicity with virulence in animal models (Gutiérrez et al., 2018; Johansen et al., 2020a). These features are indicative of the critical role that surface hydrophobicity has played in the evolution of *M. tuberculosis* toward greater pathogenicity and the assumption that increased hydrophobicity enhances aerosol transmission capacity and pathogenicity (Jankute et al., 2017). The fact that hydrophobic Δ1 and Δ2 mutants exhibit increased virulence in macrophages and in zebrafish supports this hypothesis. Strikingly, infection with both mutants was associated with large extracellular cord formation at comparable levels with the WT R, consistent with the fact that cording represents a hallmark of disease severity in this animal model (Bernut et al., 2014c). Due to their excessive size, *M. abscessus* cords cannot be ingested by professional phagocytes, leading to uncontrolled extracellular growth and abscess formation that rapidly leads to larval killing, as observed with Δ1 and Δ2. This clearly indicates that, in the presence of the sole innate immune system, the simple loss of 6-dTal and/or the internal Rha conveys a virulent phenotype and lethal infection in zebrafish embryos.

The outermost carbohydrate-rich bacterial layer represents the interaction interface between bacteria and host cells and is known to participate in adhesion and modulation of the immune response (Poole et al., 2018). L-Rha is an important component of mycobacterial cell walls (Mistou et al., 2016). Our results

emphasize the importance of GPL-associated Rha residues in controlling virulence and infectivity. L-Rha (6-deoxy-L-mannose) is a monosaccharide belonging to the deoxy-hexose family, commonly encountered in bacteria and plants (Wagstaff et al., 2021). Rha is ubiquitously incorporated into the surface carbohydrates of Gram-negative and Gram-positive bacteria (Wagstaff et al., 2021) and is a major constituent of the surface carbohydrates of pathogenic bacteria, such as *Shigella flexneri, Pseudomonas aeruginosa* or *Streptococcus pyogenes* (Wagstaff et al., 2021). In addition, previous work showed the crucial role of Rha-containing Lancefield antigens from group A *Streptococcus* and group B *Streptococcus* in bacterial growth and pathogenesis (Caliot et al., 2012). Therefore, in addition to their role in bacterial virulence and growth, our study indicates that di-Rha plays a direct role in the uptake of *M. abscessus* by macrophages. Since the internal Rha is mono- or di-methylated, additional studies are required to address whether methylation imparts on the internalization of *M. abscessus* by host cells. Similarly, *P. aeruginosa* mutants lacking the capacity to produce large D-Rha-rich polysaccharide antigens adhere less efficiently to human bronchial epithelial cells (Wagstaff et al., 2021). 6-dTal, which is the C2 epimer of 6-dGal, exerts important antimicrobial and anti-inflammatory properties, however data focused on the contribution to bacterial pathogenicity are scarce (Van Overtveldt et al., 2018). Our results incriminate 6-dTal, which is associated with attenuated virulence of *M. abscessus* in fish, as a determinant promoting bacterial cell entry. However, as 6-dTal is di-*O*-acetylated, it is unclear whether the phenotypes associated with $\Delta 1$ are directly related to the loss of the monosaccharide and/or caused by the absence of the two acetyl groups. The *M. abscessus gpl* locus possesses two genes, *atf1* and *atf2*, whose translation products transfer acetyl groups to 6-dTal sequentially (Ripoll et al., 2007). Future studies on *atf1* and *atf2* mutants are required to explore the potential involvement of GPL acetylation in *M. abscessus* morphotype, virulence and internalization by host cells.

Carbohydrates are known to interact with peptides through interactions involving hydrogen bonding with one or more hydroxyl acceptor groups of their constituting monosaccharide residues (Camiruaga et al., 2017; Roy et al., 2016). Considering GPL-3, the di-*O*-acetylated 6-dTal and the external Rha exhibit free hydroxyl groups (4 in total) that are likely to interact with their cognate receptors. In contrast, GPL-2a (lacking the extra Rha) possesses only 2 hydroxyl groups (one on the acetylated 6-dTal and one on the internal Rha), which may reduce hydrogen bonding and receptor affinity. Consistent with this hypothesis, phagocytosis assays indicate that GPL-2a and GPL-3 interact primarily with MR and to a lesser extent with CR3, and that receptor binding is augmented in the presence of the 6-dTal and the extra Rha in GPL-3. Accordingly, phagocytosis of the Gtf3_ov strain was inhibited by mannan. However, mannan pre-incubation had no impact on WT S uptake, most likely due to a heterogenous GPL composition comprising both GPL-2a and GPL-3 that may mask the contribution of di-rhamnosylated GPL in MR recognition. This suggests that other receptors are likely to be involved in the GPL-dependent uptake of *M. abscessus* S.

Rha, as indicated above, is structurally related to mannose and with its hydroxyl groups in equatorial positions at the *O*-3 and *O*-4 exhibits an optimal configuration for binding to the carbohydrate-recognition

domains of the mannose- and fucose-binding lectins, such as the MR (Feinberg et al., 2021; Zelensky and Gready, 2005). The mycobacterial envelopes are rich in mannose-containing glycoconjugates and several other ligands of the MR have been previously identified, such as the mannose-capped lipoarabinomannan (ManLAM) of *M. tuberculosis* or higher-order PIMs ($PIM_{5–6}$), which associate with the MR *via* their mannosylated α(1→2) motifs located at their non-reducing ends (Turner and Torrelles, 2018). The MR recognizes the terminal mannosyl unit of these motifs in Ca^{2+}-fashion *via* its carbohydrate-recognition domain 4, but supplementary interactions with the penultimate mannose residue allow an enhanced affinity for disaccharide motifs (Feinberg et al., 2021). In addition to MR, our data indicate binding of GPL3 to CR3, a heterodimer of α (CD11b) and β (CD18) transmembrane glycoproteins. It has been previously shown to recognize a mixture of di- and tri-glycosylated forms of *M. smegmatis* (Villeneuve et al., 2005). Its carbohydrate specificity is broader than initially appreciated, allowing it to react with a wide panel of polysaccharides containing mannose, glucose, as well as galactose (Poole et al., 2020; Thornton et al., 1996). CR3 has been involved in *M. tuberculosis* phagocytosis (Fenton et al., 2005) and mannosylated cell wall components, such as PIM_2, can bind directly to the lectin domain of CR3 (Villeneuve et al., 2005), arbitrating *M. tuberculosis* uptake by macrophages (Turner and Torrelles, 2018). Tal is structurally related to galactose. Although acetylated at positions 3 and 4 in GPL, the acetyl groups may allow hydrogen bonding to the receptors binding domain.

SIGNIFICANCE

Our results indicate that the monosaccharide composition of GPL from *M. abscessus* affects virulence and disease severity. It remains to be established whether GPL glycosylation is regulated in macrophages or during infection. Polar species of GPL in *M. smegmatis* are produced only under carbon starvation and induce smooth-colony formation (Mukherjee et al., 2005; Ojha et al., 2002), opening up the likelihood that the structure of GPL in *M. abscessus* is modulated in response to changing environments and during intracellular survival or infection. Thus, future studies should identify factors regulating the expression of *gtf1, gtf2* and *gtf3*, which would further enlighten our understanding of cell wall assembly and adaptation of *M. abscessus* to its host.

ACKNOWLEDGMENTS

This study was supported the French National Research Agency ANR-19-CE15-0012-01 (SUNLIVE). MDJ received a post-doctoral fellowship granted by Labex EpiGenMed, an «Investissements d'avenir» program (ANR-10-LABX-12-01). We acknowledge the Infectiopoĺe Sud Međiterraneé for funding the PhD fellowship of JK. We are grateful to the Montpellier RIO Imaging facilities and the CRBM zebrafish facility (Montpellier), P. Richard and M. Plays for zebrafish husbandry, Fabien Blanchet for providing primary human macrophages. We are grateful to the PAGes core facility (http://plateforme- pages.univ-lille1.fr) and UMS 2014 - US 41 -

Plateformes Lilloises en Biologie & Santé for providing the scientific and technical environment conducive to achieving this work.

AUTHORS CONTRIBUTIONS

WD conceived and conducted experiments, analyzed the data, and wrote the manuscript. LDL, MDJ, JK and XT conducted experiments. LDL, MDJ, JN and YG analyzed the data and participated in writing the manuscript. LK conceived the idea of the project, analyzed the data and wrote the manuscript.

DECLARATION OF INTEREST

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental Information includes 7 figures, 1 Table and 2 movies.

REFERENCES

Bernut, A., Herrmann, J.-L., Kissa, K., Dubremetz, J.-F., Gaillard, J.-L., Lutfalla, G., and Kremer, L. (2014a). *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess formation. Proc. Natl. Acad. Sci. U.S.A. *111*, E943-952.

Bernut, A., Le Moigne, V., Lesne, T., Lutfalla, G., Herrmann, J.-L., and Kremer, L. (2014b). *In vivo* assessment of drug efficacy against *Mycobacterium abscessus* using the embryonic zebrafish test system. Antimicrob. Agents Chemother. *58*, 4054–4063.

Bernut, A., Herrmann, J.-L., Kissa, K., Dubremetz, J.-F., Gaillard, J.-L., Lutfalla, G., and Kremer, L. (2014c). *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess formation. Proc. Natl. Acad. Sci. U.S.A. *111*, E943-952.

Bernut, A., Dupont, C., Sahuquet, A., Herrmann, J.-L., Lutfalla, G., and Kremer, L. (2015). Deciphering and imaging pathogenesis and cording of *Mycobacterium abscessus* in zebrafish embryos. J. Vis. Exp. *103*, e53130.

Bernut, A., Viljoen, A., Dupont, C., Sapriel, G., Blaise, M., Bouchier, C., Brosch, R., de Chastellier, C., Herrmann, J.-L., and Kremer, L. (2016). Insights into the smooth-to-rough transitioning in *Mycobacterium bolletii* unravels a functional Tyr residue conserved in all mycobacterial MmpL family members. Mol. Microbiol. *99*, 866–883.

Bernut, A., Herrmann, J.-L., Ordway, D., and Kremer, L. (2017). The diverse cellular and animal models to decipher the physiopathological traits of *Mycobacterium abscessus* infection. Front Cell Infect Microbiol *7*, 100.

Blanchet, F.P., Moris, A., Nikolic, D.S., Lehmann, M., Cardinaud, S., Stalder, R., Garcia, E., Dinkins, C., Leuba, F., Wu, L., et al. (2010). Human immunodeficiency virus-1 inhibition of immunoamphisomes in dendritic cells impairs early innate and adaptive immune responses. Immunity *32*, 654–669.

Brown, G.D., and Gordon, S. (2001). A new receptor for β-glucans. Nature *413*, 36–37.

Bryant, J.M., Grogono, D.M., Rodriguez-Rincon, D., Everall, I., Brown, K.P., Moreno, P., Verma, D., Hill, E., Drijkoningen, J., Gilligan, P., et al. (2016). Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. Science *354*, 751–757.

Caliot, É., Dramsi, S., Chapot-Chartier, M.-P., Courtin, P., Kulakauskas, S., Péchoux, C., Trieu-Cuot, P., and Mistou, M.-Y. (2012). Role of the Group B antigen of *Streptococcus agalactiae*: a peptidoglycan-anchored polysaccharide involved in cell wall biogenesis. PLoS Pathog *8*, e1002756.

Camiruaga, A., Usabiaga, I., Insausti, A., León, I., and Fernández, J.A. (2017). Sugar–peptidic bond interactions: spectroscopic characterization of a model system. Phys. Chem. Chem. Phys. *19*, 12013–12021.

Catherinot, E., Clarissou, J., Etienne, G., Ripoll, F., Emile, J.-F., Daffé, M., Perronne, C., Soudais, C., Gaillard, J.-L., and Rottman, M. (2007). Hypervirulence of a rough variant of the *Mycobacterium abscessus* type strain. Infect. Immun. *75*, 1055–1058.

Catherinot, E., Roux, A.-L., Macheras, E., Hubert, D., Matmar, M., Dannhoffer, L., Chinet, T., Morand, P., Poyart, C., Heym, B., et al. (2009). Acute respiratory failure involving an R variant of *Mycobacterium abscessus*. J. Clin. Microbiol. *47*, 271–274.

Cronan, M.R., and Tobin, D.M. (2014). Fit for consumption: zebrafish as a model for tuberculosis. Dis Model Mech *7*, 777–784.

Daher, W., Leclercq, L.-D., Viljoen, A., Karam, J., Dufrêne, Y.F., Guérardel, Y., and Kremer, L. (2020). O-Methylation of the Glycopeptidolipid Acyl Chain Defines Surface Hydrophobicity of *Mycobacterium abscessus* and Macrophage Invasion. ACS Infect Dis *6*, 2756–2770.

Dedrick, R.M., Smith, B.E., Garlena, R.A., Russell, D.A., Aull, H.G., Mahalingam, V., Divens, A.M., Guerrero-Bustamante, C.A., Zack, K.M., Abad, L., et al. (2021). *Mycobacterium abscessus* Strain Morphotype Determines Phage Susceptibility, the Repertoire of Therapeutically Useful Phages, and Phage Resistance. MBio *12*.

Deshayes, C., Laval, F., Montrozier, H., Daffé, M., Etienne, G., and Reyrat, J.-M. (2005). A glycosyltransferase involved in biosynthesis of triglycosylated glycopeptidolipids in *Mycobacterium smegmatis*: impact on surface properties. J Bacteriol *187*, 7283–7291.

Eckstein, T.M., Belisle, J.T., and Inamine, J.M. (2003). Proposed pathway for the biosynthesis of serovar-specific glycopeptidolipids in *Mycobacterium avium* serovar 2. Microbiology (Reading) *149*, 2797–2807.

Esther, C.R., Esserman, D.A., Gilligan, P., Kerr, A., and Noone, P.G. (2010). Chronic *Mycobacterium abscessus* infection and lung function decline in cystic fibrosis. J. Cyst. Fibros. *9*, 117–123.

Feinberg, H., Jégouzo, S.A.F., Lasanajak, Y., Smith, D.F., Drickamer, K., Weis, W.I., and Taylor, M.E. (2021). Structural analysis of carbohydrate binding by the macrophage mannose receptor CD206. J Biol Chem 100368.

Fenton, M.J., Riley, L.W., and Schlesinger, L.S. (2005). Receptor-Mediated Recognition of *Mycobacterium tuberculosis* by Host Cells. Tuberculosis and the Tubercle Bacillus 405–426.

Floto, R.A., Olivier, K.N., Saiman, L., Daley, C.L., Herrmann, J.-L., Nick, J.A., Noone, P.G., Bilton, D., Corris, P., Gibson, R.L., et al. (2016). US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis: executive summary. Thorax *71*, 88–90.

Gantner, B.N., Simmons, R.M., and Underhill, D.M. (2005). Dectin-1 mediates macrophage recognition of *Candida albicans* yeast but not filaments. EMBO J *24*, 1277–1286.

Gutiérrez, A.V., Viljoen, A., Ghigo, E., Herrmann, J.-L., and Kremer, L. (2018). Glycopeptidolipids, a Double-Edged Sword of the *Mycobacterium abscessus* Complex. Front Microbiol *9*, 1145.

Howard, S.T., Rhoades, E., Recht, J., Pang, X., Alsup, A., Kolter, R., Lyons, C.R., and Byrd, T.F. (2006). Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. Microbiology (Reading, Engl.) *152*, 1581–1590.

Jankute, M., Nataraj, V., Lee, O.Y.-C., Wu, H.H.T., Ridell, M., Garton, N.J., Barer, M.R., Minnikin, D.E., Bhatt, A., and Besra, G.S. (2017). The role of hydrophobicity in tuberculosis evolution and pathogenicity. Sci Rep *7*, 1315.

Jarand, J., Levin, A., Zhang, L., Huitt, G., Mitchell, J.D., and Daley, C.L. (2011). Clinical and microbiologic outcomes in patients receiving treatment for *Mycobacterium abscessus* pulmonary disease. Clin. Infect. Dis. *52*, 565–571.

Johansen, M.D., and Kremer, L. (2020). Large Extracellular Cord Formation in a Zebrafish Model of *Mycobacterium kansasii* Infection. The Journal of Infectious Diseases *222*, 1046–1050.

Johansen, M.D., Herrmann, J.-L., and Kremer, L. (2020a). Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. Nat Rev Microbiol *18*, 392–407.

Johansen, M.D., Daher, W., Roquet-Banères, F., Raynaud, C., Alcaraz, M., Maurer, F.P., and Kremer, L. (2020b). Rifabutin Is Bactericidal against Intracellular and Extracellular Forms of *Mycobacterium abscessus*. Antimicrob Agents Chemother *64*, e00363-20, /aac/64/11/AAC.00363-20.atom.

Jönsson, B.E., Gilljam, M., Lindblad, A., Ridell, M., Wold, A.E., and Welinder-Olsson, C. (2007). Molecular epidemiology of *Mycobacterium abscessus*, with focus on cystic fibrosis. J. Clin. Microbiol. *45*, 1497–1504.

Kwak, N., Dalcolmo, M.P., Daley, C.L., Eather, G., Gayoso, R., Hasegawa, N., Jhun, B.W., Koh, W.-J., Namkoong, H., Park, J., et al. (2019). *Mycobacterium abscessus* pulmonary disease: individual patient data meta-analysis. Eur. Respir. J. *54*.

Lamason, R.L., Mohideen, M.-A.P.K., Mest, J.R., Wong, A.C., Norton, H.L., Aros, M.C., Jurynec, M.J., Mao, X., Humphreville, V.R., Humbert, J.E., et al. (2005). SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science *310*, 1782–1786.

Lee, S.-Y., Kim, H.-Y., Kim, B.-J., Kim, H., Seok, S.-H., Kim, B.-J., and Kook, Y.-H. (2017). Effect of amikacin on cell wall glycopeptidolipid synthesis in *Mycobacterium abscessus*. J. Microbiol. *55*, 640–647.

Lin, J.-S., Huang, J.-H., Hung, L.-Y., Wu, S.-Y., and Wu-Hsieh, B.A. (2010). Distinct roles of complement receptor 3, Dectin-1, and sialic acids in murine macrophage interaction with Histoplasma yeast. Journal of Leukocyte Biology *88*, 95–106.

Madani, A., Ridenour, J.N., Martin, B.P., Paudel, R.R., Abdul Basir, A., Le Moigne, V., Herrmann, J.-L., Audebert, S., Camoin, L., Kremer, L., et al. (2019). Cyclipostins and Cyclophostin Analogues as Multitarget Inhibitors That Impair Growth of *Mycobacterium abscessus*. ACS Infect Dis *5*, 1597–1608.

Martínez, A., Torello, S., and Kolter, R. (1999). Sliding motility in mycobacteria. J Bacteriol *181*, 7331–7338.

Medjahed, H., Gaillard, J.-L., and Reyrat, J.-M. (2010). *Mycobacterium abscessus*: a new player in the mycobacterial field. Trends Microbiol. *18*, 117–123.

Mistou, M.-Y., Sutcliffe, I.C., and van Sorge, N.M. (2016). Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. FEMS Microbiol Rev *40*, 464–479.

Miyamoto, Y., Mukai, T., Nakata, N., Maeda, Y., Kai, M., Naka, T., Yano, I., and Makino, M. (2006). Identification and characterization of the genes involved in glycosylation pathways of mycobacterial glycopeptidolipid biosynthesis. J. Bacteriol. *188*, 86–95.

Mukherjee, R., Gomez, M., Jayaraman, N., Smith, I., and Chatterji, D. (2005). Hyperglycosylation of glycopeptidolipid of *Mycobacterium smegmatis* under nutrient starvation: structural studies. Microbiology *151*, 2385–2392.

Neelamegham, S., Aoki-Kinoshita, K., Bolton, E., Frank, M., Lisacek, F., Lütteke, T., O'Boyle, N., Packer, N.H., Stanley, P., Toukach, P., et al. (2019). Updates to the Symbol Nomenclature for Glycans guidelines. Glycobiology *29*, 620–624.

Ojha, A.K., Varma, S., and Chatterji, D. (2002). Synthesis of an unusual polar glycopeptidolipid in glucose-limited culture of *Mycobacterium smegmatis*. Microbiology (Reading, Engl.) *148*, 3039–3048.

Park, I.K., Hsu, A.P., Tettelin, H., Shallom, S.J., Drake, S.K., Ding, L., Wu, U.-I., Adamo, N., Prevots, D.R., Olivier, K.N., et al. (2015). Clonal Diversification and Changes in Lipid Traits and Colony Morphology in *Mycobacterium abscessus* Clinical Isolates. J. Clin. Microbiol. *53*, 3438–3447.

Pawlik, A., Garnier, G., Orgeur, M., Tong, P., Lohan, A., Le Chevalier, F., Sapriel, G., Roux, A.-L., Conlon, K., Honoré, N., et al. (2013). Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent *Mycobacterium abscessus*. Mol. Microbiol. *90*, 612–629.

Poole, J., Day, C.J., von Itzstein, M., Paton, J.C., and Jennings, M.P. (2018). Glycointeractions in bacterial pathogenesis. Nat Rev Microbiol *16*, 440–452.

Poole, J., Day, C.J., Haselhorst, T., Jen, F.E.-C., Torres, V.J., Edwards, J.L., and Jennings, M.P. (2020). Repurposed Drugs That Block the Gonococcus-Complement Receptor 3 Interaction Can Prevent and Cure Gonococcal Infection of Primary Human Cervical Epithelial Cells. MBio *11*.

Richard, M., Gutiérrez, A.V., Viljoen, A., Rodriguez-Rincon, D., Roquet-Baneres, F., Blaise, M., Everall, I., Parkhill, J., Floto, R.A., and Kremer, L. (2019). Mutations in the MAB_2299c TetR Regulator Confer Cross-Resistance to Clofazimine and Bedaquiline in *Mycobacterium abscessus*. Antimicrob. Agents Chemother. *63*.

Ripoll, F., Deshayes, C., Pasek, S., Laval, F., Beretti, J.-L., Biet, F., Risler, J.-L., Daffé, M., Etienne, G., Gaillard, J.-L., et al. (2007). Genomics of glycopeptidolipid biosynthesis in *Mycobacterium abscessus* and *M. chelonae*. BMC Genomics *8*, 114.

Roux, A.-L., Ray, A., Pawlik, A., Medjahed, H., Etienne, G., Rottman, M., Catherinot, E., Coppée, J.-Y., Chaoui, K., Monsarrat, B., et al. (2011). Overexpression of proinflammatory TLR-2-signalling lipoproteins in hypervirulent mycobacterial variants. Cell. Microbiol. *13*, 692–704.

Roux, A.-L., Viljoen, A., Bah, A., Simeone, R., Bernut, A., Laencina, L., Deramaudt, T., Rottman, M., Gaillard, J.-L., Majlessi, L., et al. (2016). The distinct fate of smooth and rough *Mycobacterium abscessus* variants inside macrophages. Open Biol *6*.

Roy, R., Murphy, P.V., and Gabius, H.-J. (2016). Multivalent Carbohydrate-Lectin Interactions: How Synthetic Chemistry Enables Insights into Nanometric Recognition. Molecules *21*.

Ryan, K., and Byrd, T.F. (2018). *Mycobacterium abscessus*: Shapeshifter of the Mycobacterial World. Front Microbiol *9*.

Schlesinger, L.S., Hull, S.R., and Kaufman, T.M. (1994). Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of *Mycobacterium tuberculosis* to human macrophages. J Immunol *152*, 4070–4079.

Shepherd, V.L., Stahl, P.D., Bernd, P., and Rabinovitch, M. (1983). Receptor-mediated entry of beta-glucuronidase into the parasitophorous vacuoles of macrophages infected with *Leishmania mexicana* amazonensis. Journal of Experimental Medicine *157*, 1471–1482.

Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., Bansal, G.P., Young, J.F., Lee, M.H., and Hatfull, G.F. (1991). New use of BCG for recombinant vaccines. Nature *351*, 456–460.

Thornton, B.P., Větvicka, V., Pitman, M., Goldman, R.C., and Ross, G.D. (1996). Analysis of the sugar specificity and molecular location of the beta-glucan-binding lectin site of complement receptor type 3 (CD11b/CD18). J Immunol *156*, 1235–1246.

Turner, J., and Torrelles, J.B. (2018). Mannose-capped lipoarabinomannan in *Mycobacterium tuberculosis* pathogenesis. Pathog Dis *76*.

Van Overtveldt, S., Gevaert, O., Cherlet, M., Beerens, K., and Desmet, T. (2018). Converting Galactose into the Rare Sugar Talose with Cellobiose 2-Epimerase as Biocatalyst. Molecules *23*.

Viljoen, A., Blaise, M., de Chastellier, C., and Kremer, L. (2016). MAB_3551c encodes the primary triacylglycerol synthase involved in lipid accumulation in *Mycobacterium abscessus*. Mol Microbiol *102*, 611–627.

Viljoen, A., Gutiérrez, Ana Victoria, Dupont, C., Ghigo, E., and Kremer, L. (2018). A simple and rapid gene disruption strategy in *Mycobacterium abscessus*: on the design and application of glycopeptidolipid mutants. Front. Cell. Infect. Microbial. *8:69*.

Viljoen, A., Viela, F., Kremer, L., and Dufrêne, Y.F. (2020). Fast chemical force microscopy demonstrates that glycopeptidolipids define nanodomains of varying hydrophobicity on mycobacteria. Nanoscale Horiz *5*, 944–953.

Villeneuve, C., Etienne, G., Abadie, V., Montrozier, H., Bordier, C., Laval, F., Daffe, M., Maridonneau-Parini, I., and Astarie-Dequeker, C. (2003). Surface-exposed glycopeptidolipids of *Mycobacterium smegmatis* specifically inhibit the phagocytosis of mycobacteria by human macrophages. Identification of a novel family of glycopeptidolipids. J Biol Chem *278*, 51291–51300.

Villeneuve, C., Gilleron, M., Maridonneau-Parini, I., Daffé, M., Astarie-Dequeker, C., and Etienne, G. (2005). Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. J Lipid Res *46*, 475–483.

Wagstaff, B.A., Zorzoli, A., and Dorfmueller, H.C. (2021). NDP-rhamnose biosynthesis and rhamnosyltransferases: building diverse glycoconjugates in nature. Biochemical Journal *478*, 685–701.

Whang, J., Back, Y.W., Lee, K.-I., Fujiwara, N., Paik, S., Choi, C.H., Park, J.-K., and Kim, H.-J. (2017). *Mycobacterium abscessus* glycopeptidolipids inhibit macrophage apoptosis and bacterial spreading by targeting mitochondrial cyclophilin D. Cell Death Dis *8*, e3012.

Woods, G.L., Brown-Elliott, B.A., Conville, P.S., Desmond, E.P., Hall, G.S., Lin, G., Pfyffer, G.E., Ridderhof, J.C., Siddiqi, S.H., Wallace, R.J., et al. (2011). Susceptibility Testing of Mycobacteria, Nocardiae, and Other Aerobic Actinomycetes (Wayne (PA): Clinical and Laboratory Standards Institute).

Xia, Y., and Ross, G.D. (1999). Generation of recombinant fragments of CD11b expressing the functional beta-glucan-binding lectin site of CR3 (CD11b/CD18). J Immunol *162*, 7285–7293.

Zelensky, A.N., and Gready, J.E. (2005). The C-type lectin-like domain superfamily. FEBS J *272*, 6179–6217.

FIGURE LEGENDS

Figure 1. GPL composition characterizing the *gtf* mutants.

(A) Structure of methoxylated GPL-3. The Symbol Nomenclature for Glycans (SFNG) is used throughout the figure to symbolize monosaccharides (Neelamegham et al., 2019). X represents the branched-chain amino-alcohol, alaninol or valinol. The lipid moiety is made of a combination of saturated and unsaturated β-methoxylated fatty acids.

(B) TLC profiles of GPL isolated from WT *M. abscessus* S and R strains, the single deletion mutants *gtf1* (Δ1), *gtf2* (Δ2) and *gtf3* (Δ3) as well as their complemented counterparts (Δ1_C, Δ2_C and Δ3_C) and multiple *gtf* deletion mutants (Δ1,2; Δ1,3; Δ2,3; Δ1,2,3). WT profiles were restored in all complemented strains. TLC plates were developed in $CHCl_3/CH_3OH/H_2O$ (90/10/1, v/v/v), stained with orcinol reagent and charred.

(C) Monosaccharide composition (expressed as the proportion of each monosaccharide) shows the absence of deoxy hexoses in *M. abscessus* R while deletion of *gtf1* or *gft3* induces the disappearance of 3(CH_3)Rha and 6-dTal or Rha, respectively. No deoxy hexose was observed in Δ2, similarly to *M. abscessus* R.

(D) MALDI-TOF-MS spectra of GPL from *M. abscessus* strains. The R variant is dominated by Ac_1PIM_2 signal whereas the S variant shows two heterogenous clusters of major signals at *m/z* 1258/1286 and 1404/1432, tentatively identified as $[M+Na]^+$ adducts of diglycoylsated (GPL-2a) and triglycosylated (GPL-3) GPL, respectively. Single *gtf* deletions induced drastic changes in the GPL profiles, as shown in the spectra.

(E) MALDI-TOF-MS^2 spectrum of WT GPL-3 parent ion at *m/z* 1432, confirms the respective positions of monosaccharides on the peptide moiety.

(F) MALDI-TOF-MS^2 spectrum of wild-type GPL-2a parent ion at *m/z* 1286 shows the absence of the external Rha.

Figure 2. Proposed pathway for *M. abscessus* GPL glycosylation.

(A) The presence of different GPL glycoforms found in WT and *gtf* mutants was evaluated by TLC and confirmed by MS. While GPL-1a has not been confirmed in any of the *gtf* mutants, all other glycoforms (GPL-1a, GPL-1b, GPL-2a, GPL-2b and GPL-3) were detected at specific *m/z* as native or saponified GPL and assigned in the different mutants.

(B) The glycosylation biosynthetic pathway of GPL in *M. abscessus* was modeled based on the distribution of individual GPL in the *gtf* mutants. First, Gtf2 adds the internal Rha to the branched chain amino-alcohol of the peptidolipid core, resulting in GPL-1b. Second, Gtf1 and Gtf3 add the 6-dTal and the external Rha residue, resulting in GPL-2b and GPL-2a, respectively. The combined action of all three enzymes generates GPL-3.

Figure 3. Deletion of *gtf* genes alters colonial morphology, hydrophobicity, sliding motility and biofilm formation.

(A) Knock-out of *gtf1* or *gtf2,* but not *gtf3,* is associated with rough and corded colonies, similar to the WT R reference strain. The various *M. abscessus* strains were grown on Tryptic Soy agar plates and incubated at 37° C for 5 days.

(B) Gtf1 and Gtf2 maintain a normal hydrophilicity index of the smooth morphotype. *M. abscessus* R, Δ1, Δ2, double and triple mutants and Gtf3 overexpressing strain are more hydrophobic than S, while the Δ3 and Δ*fmt* mutants hydrophilicity is significantly higher than S, as assessed by hexadecane partitioning and shown as the hydrophobicity index (in percentage of aqueous phase OD prior to partitioning). Histograms and error bars are means ± SD of four independent experiments (each time in triplicate) (n=12). Differences between means were analyzed for significance using a two-tailed Student's t-test employing Welch's correction for unequal variances. ns, non-significant; ***p < 0.001.

(C) Disruption of *gtf1* and/or *gtf2* genes reduces the sliding motility distance traveled by *M. abscessus* strains. Smooth morphotype strains slide more than R variants. Quantification of sliding ability. The sliding distances were measured in mm and plotted. Histograms and error bars are means ± SD of six independent experiments (each time in triplicate) (n=18). Data were analyzed with one-way ANOVA and Tukey's multiple comparison. ns, non-significant; **p < 0.01; ***p < 0.001.

(D) The absence of Gtf1 and Gtf2 promotes biomass formation at the air-liquid interface in *M. abscessus*. The rough variants of *M. abscessus* develop more significant biofilm structures at the air-liquid interface than the smooth variants. Biomass (expressed as CV absorbance) was greater with the rough strains than with the smooth variants. Histograms and error bars are means ± SD of three independent experiments (each time in sextuplicate) (n=24). Data were analyzed with one-way ANOVA and Tukey's multiple comparison. ns, non-significant; **p < 0.01; ***p < 0.001. See also Figure S4.

(E) The knock-out of *gtf1* and/or *gtf2* genes impairs biomass formation at the solid/liquid interface in *M. abscessus*. Smooth variants of *M. abscessus* evolve more important biofilm structures at the solid/liquid interface than rough morphotype strains. To assess the biomass development of each strain at the solid/liquid interface, 6-day-old *M. abscessus* biofilms were enumerated for CFUs. Data are mean values ± SD for three independent experiments (each time in quadruplicate) (n=12). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001. See also Figure S4.

Figure 4. *Δgtf1* and *Δgtf2* mutants lead to large extracellular cord and abscess formation in a zebrafish embryo model.

(A) Embryos at 30 hours post-fertilization were infected with ~250-300 CFU of *M. abscessus via* caudal vein injection and survival was monitored daily over a 12 day-period. Statistical analysis was performed using the log-rank (Mantel-Cox) statistical test for survival analysis.

(B) The presence of cords in each embryo was quantified via fluorescent microscopy at 2- and 4-day post-infection.

(C) The number of cords per embryo was further quantified using fluorescent microscopy at 2- and 4-days post-infection.

(D) The proportion of embryos with abscesses was identified using fluorescent microscopy at 2-, 4- and 6-days post-infection.

(E) The number of abscesses per embryo was quantified with fluorescent microscopy at 2-, 4- and 6-days post-infection. Statistical analysis for cord and abscess quantification was performed using unpaired students t-tests.

(F) Representative images of each *M. abscessus* strain expressing the pTEC27 fluorescent plasmid in zebrafish embryos at 4 days post-infection. Values for all data displayed are the mean ± standard deviation with the merge of three independent experiments. $P \leq 0.05$, ** ≤ 0.01, *** $P \leq 0.001$.

Figure 5. Internalization of *M. abscessus* by macrophages requires glycosylated GPL.

(A) Intracellular bacterial load is reduced in macrophages infected with single, double or triple *gtf* mutants. CFU were determined at hours 3, 24, and 72 post-infection. Data are mean values ± SD for ten independent experiments (each time in triplicate) (n=30). One-tailed Mann−Whitney test: ns, nonsignificant; ***p < 0.001.

(B) The percentage of infected macrophages decreased when cells were infected with single, double or triple *gtf* mutants. Percentage of infected THP-1 macrophages at hours 3, 24, and 72 post-infection. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=180). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001.

(C) Twelve immunofluorescent fields were taken after 24 hours post-infection at a 40X magnification (using a confocal microscope), showing the macrophages infected with the various strains (in red). The nuclei are shown in blue, the CD43 protein associated with the plasma membrane and present on the surface of macrophages is stained in green. White arrows indicate mycobacteria within the macrophages. Scale bars represent 20 µm.

(D) The number of bacilli per infected macrophage declines when cells are infected with single, double or triple *gtf* mutants. Analysis of the percentage of macrophage categories infected with different numbers of bacilli. The first image illustrates a macrophage containing fewer than 5 bacilli, the second encompasses 5 to 10 bacilli, and the third comprises >10 bacilli (right panels). The surface of macrophages was detected using anti-CD43 antibodies (green). The nuclei were stained with DAPI (blue). The macrophage categories were counted at 3 hpi (left panel). The percentages of macrophages containing different numbers of bacilli (0−5, 5−10, and >10) are represented on the y axis. Values are means ± SD for three independent experiments performed each time in triplicate (n=900 infected macrophages). Data were analyzed using the one-tailed nonpaired t test. ns, non-significant; ***p < 0.001.

Figure 6. Triglycosylated GPL favors colonization of macrophages by *M. abscessus.*

(A) Overexpression of Gtf3-HA in *M. abscessus*. The *gtf3* gene was fused to HA and cloned into the integrative vector pMV306 (lane 3) or episomal pMV261 (lane 4) allowing its expression under the control of its native promoter (in Δ3 strain) or hsp60 (in the wild-type smooth strain) respectively. Western blot analysis of Gtf3 expression in the complemented Δ3 strain or its overexpression in the wild-type S morphotype strain using anti-HA antibodies (upper panel). The KasA protein (probed with anti-KasA antibodies) was used as a loading control (lower panel).

(B) GPL-2 and GPL-3 accumulated in Δ3, and Gtf3_ov strains respectively. TLC analysis of the crude lipid fractions of the smooth, Δ3 mutant, Δ3 complemented and Gtf3 overexpressing strains. GPL-2 and GPL-3 were separated using a solvent composed of CHCl3/MeOH (9:1, v/v) and revealed with orcinol staining after one round of migration.

(C and D) Overexpression of Gtf3 leading to accumulation of GPL-3 enhances entry of bacilli into THP-1 and primary human macrophages. CFU were determined at hours 3 (C and D), 24 (C), and 72 (C) post-infection. Data are mean values ± SD for three independent experiments (each time in pentaplicate (C) or quadruplicate (D)) (n=15 (C) or 12 (D)). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001.

(E) Two immunofluorescent fields were taken after 24 hours post-infection at a 40X magnification (using a confocal microscope), showing the macrophages infected with the various strains (in red). The nuclei are shown in blue and the CD43 protein is stained in green. White arrows indicate mycobacteria within the macrophages. Scale bars represent 20 µm.

(F) The percentage of infected macrophages augmented when cells were infected with Gtf3_ov strain. Percentage of infected THP-1 macrophages at hours 3, 24, and 72 post-infection. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=60 fields). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001.

(G) The number of bacilli per infected macrophage upsurges when cells are infected with Gtf3_ov strain. Analysis of the percentage of macrophage categories infected with different numbers of bacilli. The macrophage categories were counted at hours 3, 24, and 72 post-infection. The percentages of macrophages containing different numbers of bacilli (0−5, 5−10, and >10) are represented on the y axis. Values are means ± SD for three independent experiments performed each time in triplicate (n=900 infected macrophages). Data were analyzed using the one-tailed nonpaired t test. ns, non-significant; **p < 0.01.

(H) Increased invasion of the strain that overexpresses Gtf3 is reduced in macrophages pretreated with mannan, laminarin and methyl-α-D-glucopyranoside. Intracellular bacterial load is significantly reduced in macrophages pretreated with the 3 compounds mentioned above and infected specifically with the strain that overexpresses Gtf3. CFU were determined at 3 hpi. Data are mean values ± SD for four independent experiments (each time in triplicate) (n=12). One-tailed Mann−Whitney test: ***p < 0.001.

Figure 7. Triglycosylated GPL-conjugated beads are efficiently phagocytosed by macrophages and internalization depends on MR and CR3 receptors.

(A) Representative fluorescent images showing the efficiency of internalization of GPL-3-coupled beads by macrophages. CD43 (red), DAPI (blue), fluorescent beads (green). White arrows indicate fluorescent beads within the macrophages. Scale bars represent 20 µm. See also Figure S7.

(B) Beads ingestion is optimal only when they are coupled to the triglycosylated form of GPL. Quantification of the percentage of THP-1 containing beads was determined after 4 hours. The beads alone were used as a negative control. The ratio of beads per macrophage was 10:1. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=60 fields). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001.

(C) On average, macrophages ingest twice as many beads coupled to GPL-3. Quantification of the average number of beads per macrophage was set on after 4 hours of phagocytosis. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=900 macrophages containing beads). One-tailed Mann−Whitney test: ns, non-significant; ***p < 0.001.

(D) Blocking both MR and CR3 receptors significantly reduces cell colonization by beads coupled to both GPL-2a and GPL-3. No significant inhibition of phagocytosis was observed when the pre-treated macrophages were incubated in presence of beads alone or coupled to GPL-1b and GPL-2b. Quantification of the percentage of THP-1 containing beads was determined after 4 hours. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=60 fields). One-tailed Mann−Whitney test: ns, nonsignificant; ***p < 0.001. See also Figure S7.

(E) Representative fluorescent images showing the impact of neutralizing anti-MR and/or anti-CR3 antibodies on the internalization of GPL-2a and/or GPL-3 beads by macrophages. CD43 (red), DAPI (blue), fluorescent beads (green). White arrows indicate the presence of fluorescent beads inside macrophages. Scale bars represent 20 µm. See also Figure S7.

(F) Blocking MR or CR3 receptors with neutralizing antibodies significantly reduced cell colonization by beads coupled to GPL-2a and GPL-3. Quantification of the percentage of THP-1 containing beads was determined after 4 hours. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=180 fields). One-tailed Mann-Whitney test: ***p < 0.001. See also Figure S7.

(G) Macrophages pretreated with neutralizing antibodies ingested less beads coupled to GPL-2a (anti-MR) or GPL-3 (anti-MR and anti-CR3) than the non-treated control beads. Quantification of the number of beads per macrophage was performed after 4 hours of phagocytosis. The red bar in each violin plot represents median. The colored area of a violin plot corresponds to the probability density of the data and contains first and third quartiles. Data are mean values ± standard deviation from three independent experiments (each time in triplicate) (n=900 macrophages containing beads per condition). One-tailed Mann-Whitney test: *p < 0.05; ***p < 0.001.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Model: Bacterial Strains		
Mycobacterium abscessus sensu stricto, strain CIP104536^T, smooth	Laboratoire de Référence des Mycobactéries (IP, France)	ATCC19977^T
Mycobacterium abscessus sensu stricto, strain CIP104536^T, rough	Laboratoire de Référence des Mycobactéries (IP, France)	ATCC19977^T
Δ1: Unmarked deletion of *gtf1* in *M. abscessus* S	This study	N/A
Δ1 + pMV306-*gtf1 :* Δ1 carrying pMV306-*gtf1*	This study	N/A
Δ2: Unmarked deletion of *gtf2* in *M. abscessus* S	This study	N/A
Δ2 + pMV306-*gtf2 :* Δ2 carrying pMV306-*gtf2*	This study	N/A
Δ3: Unmarked deletion of *gtf3* in *M. abscessus* S	This study	N/A
Δ3 + pMV306-*gtf3 :* Δ3 carrying pMV306-*gtf3*	This study	N/A
S + pMV261-*gtf3 :* S carrying pMV261-*gtf3*	This study	N/A
Δ1,2: Unmarked deletions of *gtf1* and *gtf2* in *M. abscessus* S	This study	N/A
Δ1,3: Unmarked deletions of *gtf1* and *gtf3* in *M. abscessus* S	This study	N/A
Δ2,3: Unmarked deletions of *gtf2* and *gtf3* in *M. abscessus* S	This study	N/A
Δ1,2,3: Unmarked deletions of *gtf1* and *gtf2* and *gtf3* in *M. abscessus* S	This study	N/A
E. coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*q*ZΔM15* Tn*10* (Tetr)]	Stratagene
Oligonucleotides		
Primers used in this study are listed in Table S1		
Zebrafish line		
golden mutant	(Lamason et al., 2005)	N/A
Chemicals, Enzymes and kits		
Hygromycin B	Sigma-Aldrich	Cat#H3274
Kanamycin	Euromedex	Cat#EU0420-B
Difco Middlebrook 7H9 Broth	Thermo Fisher Scientific	Cat#DF0713-17-9
Middlebrook OADC Growth Supplement	Sigma-Aldrich	Cat#M0678
Tween-80	Sigma-Aldrich	Cat#P1754
Q5 DNA Polymerase	New England Biolabs	Cat#M0491L
T4 DNA Ligase	New England Biolabs	Cat#M0202S
NucleoSpin Plasmid Kit	Machery-Nagel	Cat#740588.50
NucleoSpin Gel and PCR Clean-up	Machery-Nagel	Cat#740609.50

GenElute HP Plasmid Midiprep Kit	Sigma-Aldrich	Cat#NA0200-1KT
Laminarin	Sigma-Aldrich	Cat#L9634-100MG
Mannan	Sigma-Aldrich	Cat#M7504-250MG
Methyl α-D-GLUCOPYRANOSIDE	Sigma-Aldrich	Cat#M9376-100G
Super Signal West Femto Maximum Sens	Life-Technologies	Cat#34095
Plasmids		
pTEC27 (Multicopy *E. coli*/mycobacterial shuttle vector expressing *tdTomato* under the control of a strong mycobacteri promoter)	Addgene	Cat#30182
pUX1-*katG* (Variant of *pUX1* containing the *katG* gene of *M. tuberculosis* as a counter-selectable marker in the presence of INH and used for the generation of unmarked chromosomal alterations)	(Richard et al., 2019)	N/A
pMV306 (*E. coli*/mycobacterial integrative vector)	(Stover et al., 1991)	N/A
pMV261 (Multi-copy *E. coli*/mycobacterial shuttle vector. Cloned genes are under the control of the constitutive *hsp60* promoter)	(Stover et al., 1991)	N/A
Software and Algorithms		
Prism 9.0	Graphpad	https://www.graphpad.com
Fiji (ImageJ)	NIH	https://fiji.sc
Zen (Blue edition)	Zeiss	https://www.zeiss.com/microscopy/int/products/microscope-software/zen.htmL
Imaris	Oxford	https://imaris.oxinst.com
Biological samples and Cell Lines		
Primary human monocyte-derived macrophages from human blood (Buffy coats)	Etablissement Français du Sang (EFS), Montpellier, France	N/A
THP-1 macrophages	This paper	ATCC® TIB-202™
Antibodies		
Rat anti-HA (3F10)	Merck	Cat#11867423001
Rat anti-KasA	(Viljoen et al., 2016)	N/A
Mouse anti-HU CD11B	Life Technologies	Cat#4-0118-82
Mouse anti-HU CD206	Becton Dickinson	Cat#555953
Mouse anti-HU CD43	Becton Dickinson	Cat#555474
Mouse IgG1	Becton Dickinson	Cat#555746
Mouse anti-HU Dectin-1	InvivoGen	Cat#mabg-hdect
Alexa Fluor 488 anti-Mouse IgG (Goat)	Life Technologies	Cat#A11017
Alexa Fluor 594 anti-Mouse IgG (Goat)	Life Technologies	Cat#A11032
Anti-Rat IgG H&L (HRP) (Goat)	Abcam	Cat#ab97057
CCM analysis		
TLC plate	Macherey Nagel	Cat#818333
Chloroform	Carlo	Cat#412652
Methanol	VWR	Cat#20864.320
Orcinol	Sigma	Cat#O1875
Phosphomolybdic acid	Sigma	Cat#78560
GC analysis		
6-deoxy L-Talose	Carbosynth	Cat#MD04657

L-Rhamnose monohydrate	Carbosynth	Cat#MR00035
SOLGEL 1-MS	SGE	Cat#054795
MALDI TOF analysis		
Super DHB	Sigma	Cat#50862
RMN analysis		
CDCl3	Eurisotop	Cat#D007H
CD3OD	Eurisotop	Cat#D048B
Purification		
Silica column	Interchim	Cat#FF-15SIHP-F0004
TLC plate	Macherey Nagel	Cat#809061
Coating		
Fluospheres	Thermo Fisher Scientific	Cat#F8852
Carbonate	Sigma	Cat#S7795
Bicarbonate	Sigma	Cat#S5761
Butanol	VWR	Cat#83633.290
Phosphate Buffer Saline	Euromedex	Cat#ET330
Bovine Serum Albumin	Sigma	Cat#A0281

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Laurent Kremer, (laurent.kremer@irim.cnrs.fr).

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Mycobacterial strains, growth conditions and reagents

All bacterial strains are listed in key resources table. Rough (R) and smooth (S) variants of *M. abscessus* CIP104536^T were typically grown in Middlebrook 7H9 broth (BD Difco) supplemented with 0.05% Tween 80 and 10% oleic acid, albumin, dextrose, catalase (OADC enrichment; BD Difco) (7H9$^{T/OADC}$) at 37°C in the presence of antibiotics, when required. Electrocompetent mycobacteria were transformed using a Bio-Rad Gene pulser (25 µF, 2500 V, 800 Ohms). For bacterial selection, media were supplemented either with 1 mg/mL hygromycin for strains carrying pTEC27 (Addgene, plasmid 30182), allowing tdTomato expression, or with 250 µg/mL kanamycin when harboring the pMV306 or pMV261 derivatives. On plates, colonies were selected either on Middlebrook 7H10 agar (BD Difco) supplemented with 10% OADC enrichment (7H10^{OADC}) or on LB agar or on Tryptic Soy Broth agar (Sigma). Antibiotics were purchased from Sigma-Aldrich.

In vitro growth, sliding motility and sedimentation assays

Growth was inspected by inoculating the mid-log phase cultures into fresh 7H9 at an OD_{600} of 0.05. Cultures were incubated at 37° C with shaking and OD600 was monitored for 144 hrs using a Synergy H1 hybrid reader (BioTek). To visualize colony morphology, bacteria from log phase cultures (OD_{600} = 1) were resuspended in phosphate buffered saline (PBS) and 2 µL of the cell suspension was spotted onto Tryptic Soy Broth agar. Plates were then incubated for 5 days at 37° C and colonies were imaged using a Zeiss microscope equipped with a Zeiss Plan Neo Fluor Z1×/0.25 FWD objective. Images were acquired with an Axiocam503 monochrome

(Zeiss) camera and processed using ZEN 2 (blue edition). Sliding motility was examined as previously described (Bernut et al., 2016). *M. abscessus* were inoculated on 7H9 with 0.3% agar and 6 days of incubation at 37° C, the sliding distance was measured using ImageJ. For sedimentation experiments, cultures were diluted to an OD_{600} of 1 and 2 mL of culture were transferred to a glass tube. Images were taken after 5 min.

Biofilm assays

Pellicle formation was assessed by inoculating 10 µL of mid-log phase cultures onto M63 medium supplemented with 10% glucose, 1 mM $CaCl_2$ and 1 mM $MgSO_4$, and the samples were incubated for 6 days at 30°C without agitation. To quantify biofilm production at the liquid/air and solid/liquid interfaces, 10 µL of mid-log phase cultures were deposited on the surface of M63 broth (6 mL) in glass tubes. After 6 days of incubation at 37° C, biofilms were visible at the liquid/air interface but also at the bottom of the glass tubes. Biomasses at the bottom of the tubes were collected with M63 medium and after centrifugation and removal of the supernatant, bacterial pellets were suspended in 1 mL PBS, diluted and finally spread on LB agar medium. After 5 days of incubation at 37° C, the CFU were counted. The biofilms at the liquid/air interface were stained with 1 mL of 0.1% (w/v) crystal violet in 20% (v/v) ethanol/water for 30 min. The excess stain was washed with four consecutive PBS washes and the remaining crystal violet was solubilized in 30% acetic acid and absorbance measured at 570nm.

Hexadecane partitioning

Exponentially-growing bacteria were washed twice in PUM buffer (100 mM K_2HPO_4, 54 mM KH_2PO_4, 30 mM urea, 0.8 mM $MgCl_2$) and suspended to an OD_{600} of 0.7. Aliquots (3 mL) were transferred to glass tubes and hexadecane (2.4 mL) was added. After a brief mixing, samples were incubated for 8 min at 37° C and phase separation was allowed to occur for 15 min at 22° C. The hydrophobicity index was defined as aqueous phase at OD_{600}, expressed as a percentage of that of the bacterial suspension in PUM buffer alone (Jankute et al., 2017).

Deletion of *gtf* genes in *M. abscessus*

The suicide vector pUX1-*katG* was used to generate unmarked single, double and triple deletion mutants in *M. abscessus* CIP104536^T (S). Briefly, the left and right arms (LA and RA, respectively) were PCR-amplified using genomic DNA and Q5 polymerase (New England Biolabs) as well as primers 1/2, 5/6, and 9/10 (LA) and primers 3/4, 7/8, and 11/12 (RA) **(Table S1)**. The purified LA and RA amplicons were restricted with PacI/MfeI and MfeI/NheI, respectively, and ligated to the PacI/NheI-linearized pUX1-*katG*, yielding pUX1-*katG-gtf1*, pUX1-*katG-gtf2* and pUX1-*katG-gtf3,* designed to delete 1182 bp (93%), 1104 bp (88%), and 1113 bp (84%) of the *gtf1* or *gtf2* or *gtf3* open reading frames, respectively. Electrocompetent *M. abscessus* was transformed with pUX1-*katG-gtf1*, pUX1-*katG-gtf2* and pUX1-*katG-gtf3* to generate the single mutants. Δ3

was subsequently used to generate the Δ1,3 and Δ2,3 double mutants. The Δ1,2 and Δ1,2,3 strains were derived from Δ1 and Δ1,3, respectively. The selection of bacteria having undergone the first homologous recombination event was done by visual screening of red fluorescent colonies on $7H10^{OADC}$ supplemented with 250 µg/mL kanamycin. After subculturing the culture overnight in $7H9^{T/OADC}$ in the absence of kanamycin, bacterial suspensions were serially diluted and plated onto $7H10^{OADC}$ with 50 µg/mL INH to select for INH-resistant, Kan-sensitive and non-fluorescent colonies. The DNA junctions were PCR sequenced to confirm the genotypes of the appropriate mutants using primers listed in **Table S1**.

Complementation and overexpression of Gtf3 in *M. abscessus.*

Complementation plasmids were generated by PCR amplifying the *gtf1*, *gtf2* and *gtf3* genes in fusion with an HA-tagging sequence under the control of the *hsp60* promoter (for *gtf1*) or of their endogenous promoters (175 bp for *gtf2* and 500 bp for *gtf3*) using genomic DNA and the forward 19/21/23 and reverse primers 20/22/24 (containing the HA-coding sequence) and ligation into the integrative pMV306 cut with the appropriate restriction enzymes **(Table S1)**, yielding pMV306-*gtf1*, pMV306-*gtf2* or pMV306-*gtf3*. All constructs were verified by DNA sequencing and introduced into Δ1, Δ2 or Δ3. Overexpression was achieved by PCR amplification of *gtf3* in fusion with an HA tag using genomic DNA and the forward primer (25; MscI) and reverse primer (26; EcoRI). The amplicon was digested with MscI/EcoRI and ligated into the MscI/EcoRI-restricted multicopy pMV261 to generate pMV261-*gtf3*-HA. The construct was sequenced and introduced in WT S.

Western blotting

Bacteria were harvested, resuspended in PBS, and disrupted by bead beating using 1 mm diameter glass beads. Protein concentration was assessed using the BCA Protein Assay Reagent kit (Pierce), according to the manufacturer's instructions. Equal amounts of proteins (50 µg) were separated by SDS/PAGE and transferred to a nitrocellulose membrane. For detection of Gtf1-HA, Gtf2-HA, Gtf3-HA and KasA (loading control), membranes were probed for 1 hr with either rat anti-HA (Sigma) or rat anti-KasA antibodies (Viljoen et al., 2016) (dilution 1:2000). After washing, membranes were incubated for 45 min with goat anti-rat antibody conjugated to HRP (dilution 1:5000; Abcam) and bands revealed using the ChemiDoc MP system (Bio-Rad laboratories).

Drug susceptibility testing

The MICs were determined according to the CLSI guidelines (Woods et al., 2011). The broth microdilution method was used in CaMHB with an inoculum of $5x10^6$ CFU/mL in the exponential growth phase. The bacterial suspension was seeded in 100 µL volumes in all of the wells of a 96-well plate, except for the first column, to which 198 µL of the bacterial suspension was added. In the first column, 2 µL of drug at its highest

concentration was added to the first well containing 198 µL of bacterial suspension. Then, 2-fold serial dilutions were carried out, and the plates were incubated for 5 days at 30°C. MIC were recorded by visual inspection. Assays were completed in triplicate in three independent experiments.

Zebrafish infection experiments

Zebrafish experiments were completed under European Union guidelines for the handling of laboratory animals and was approved by the Direction Sanitaire et Vétérinaire de l'Hérault and Comité d'Ethique pour l'Expérimentation Animale under reference 2020022815234677. Zebrafish infections were completed as previously described (Bernut et al., 2014a, 2015). Briefly, embryos at 30 hours post-fertilization were dechorionated and anaesthetized in tricaine prior to caudal vein injection with 2-3 nL of *M. abscessus* (~100 CFU/nL) expressing TdTomato fluorescent protein under the control of the pTEC27 plasmid. Following infection, embryos were recovered in fresh embryo water and placed in 24-well plates (2 embryos/well). Embryos were checked daily, with dead embryos removed from the well each day until 12 days post-infection. At 2-, 4- and 6-days post-infection, embryos were anaesthetized in tricaine and placed on 3% methylcellulose (w/v) for fluorescent microscopy using Zeiss Axio Zoom.V16 coupled with an Axiocam 503 mono (Zeiss). Bacterial cords were quantified based on size and shape of bacterial cords within the live zebrafish embryo, which was substantially larger in surrounding size and shape as compared to neighboring cells. Abscesses were quantified based on the large uncontrolled extracellular growth of fluorescent bacteria that exceeded 100 µm in diameter. All experiments displayed were performed in at least three independent experiments.

Macrophage infection assays

THP-1 macrophages were grown in RPMI medium supplemented with 10% fetal bovine serum (Sigma-Aldrich) and incubated at 37°C with 5% CO_2. THP-1 monocytes were differentiated in macrophages in the presence of 20 ng/mL phorbol myristate acetate in 24-well flat-bottom tissue culture microplates (10^5 cells/well) and incubated for 48 hrs at 37°C with 5% CO_2. Infection with *M. abscessus* strains (S, R, Δ1, Δ1C, Δ2, Δ2C, Δ3, Δ3C, Gtf3_ov, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3) carrying pTEC27 (MOI 2:1) was performed for 3 hrs at 37°C in the presence of 5% CO_2. Cells were carefully washed three times with 1× PBS and then incubated with $RPMI^{FBS}$ supplemented with 250 µg/mL amikacin for 2 hrs to kill extracellular bacteria. At various time points (3, 24, and 72 hpi), macrophages were washed three times with PBS and lysed with 100 µL of 1% Triton X100. Lysis was stopped by adding 900 µL PBS and serial dilutions were plated to monitor the intracellular bacterial counts. CFU were counted after 5 days of incubation at 37° C.

Immunofluorescence staining infected THP-1 macrophages

For microscopy-based infectivity assays, THP-1 cells were cultured on coverslips in 24-well plates at a density of 10^5 cells/well and incubated for 48 hrs at 37° C with 5% CO_2. Cells were infected with tdTomato-expressing *M. abscessus* (MOI 2:1) for 3 hrs, washed, treated with amikacin, fixed at 3, 24, and 72 hpi with 4% paraformaldehyde in PBS for 20 min permeabilized using 0.2% TritonX-100 for 20 min. After blocking with 2% BSA in PBS supplemented with 0.2% Triton X-100 for 20 min, cells were incubated with anti-CD43 antibody (Becton Dickinson; dilution 1:1000) for 1 hr and with a Alexa Fluor 488- or 594-conjugated anti-mouse secondary antibody (Molecular Probes, Invitrogen). Cells were then stained with 1 µg/mL 4',6-diamidino-2-phenylindole (DAPI; Becton Dickinson) for 5 min, washed with PBS, mounted onto microscope slides using Immumount (Calbiochem) and examined with either an epifluorescence microscope using a 63x lens objective or using a confocal microscope using a 40× objective (Zeiss LSM880). The average proportion of macrophages containing fewer than <5, 5−10, or >10 bacteria was quantified using Zeiss Axiovision software. Images were acquired by focusing on combined signals (CD43 in green and red fluorescent *M. abscessus*) and captured on a Zeiss Axioimager confocal microscope equipped with a 40x or 63× oil objective and processed using Zeiss Axiovision software. Scoring the number of bacilli present within macrophages were performed using ImageJ. Equal parameters for the capture and scoring of images were consistently applied to all samples.

Intra- and extracellular cords within or among the macrophages infected with the various R morphotype strains were imaged using confocal microscopy. High resolution confocal microscopy images were acquired with a LSM880 confocal microscope paired with an Airyscan module (ZEISS) with a 63x lens. The Zeiss LSM 880, Axio-Observer microscope is equipped with a Plan-Apochromat 63× Oil 1.4 NADIC M27 objective (Zeiss). The cords or macrophages containing green fluorescent beads were sectioned (mode Z-stacks) and slices were acquired every 0.18 µm, resulting in 38-42 serial slices per image. 3D imaging reconstruction of the 3D Airyscan confocal microscopy images was made with either Zen Black **(Figure S7A)** or IMARIS **(Figure S5)**. Each image channel was segmented individually by intensity thresholding and the final surface reconstruction was generated by an overlay of the three individual surfaces.

Infection of primary human monocyte-derived macrophages and intracellular CFU determination

Buffy coats from anonymous donors were obtained from the Etablissement Français du Sang (EFS), Montpellier, France. Peripheral blood mononuclear cells (PBMCs) were purified by Ficoll density gradient separation (Blanchet et al., 2010). Monocytes were isolated from PBMCs using a magnetic cell separation system with anti-CD14 mAb-coated microbeads (130-050-201; Miltenyi Biotec) and then cultured in complete IMDM supplemented with 50 ng/mL GM-CSF (every 2 days) at 37° C under a humidified 5% CO_2 atmosphere for 6 days. Cells were washed and detached with 20 mM EDTA in PBS and plated in a 24-well plate at a concentration of 10^5 cells/well with complete IMDM. Monocyte-derived macrophages were

infected with either S or Gtf3_ov strains for 3 hrs at a MOI of 2:1, washed, treated with 250 µg/mL amikacin for 2 hrs and replenished with fresh medium. To determine the number of internalized bacilli at 3 hpi, cells were lysed with 0.1% saponin and lysates plated on LB agar plates prior to CFU counting.

Fluorescent beads phagocytosis assay

The beads were added to the macrophages at a rate of 10 beads/cell. After 4 hrs of incubation at 37° C, macrophages were washed five times with PBS before labeling using anti-CD43 antibodies and Alexa Fluor 594-coupled anti-mouse secondary antibody. Quantification of the percentage of macrophages containing beads and the average number of beads per macrophage was done using an epifluorescence microscope.

Carbohydrate and antibody inhibition assays in macrophages

All carbohydrates (mannan, laminarin and methyl-α-D-glucopyranoside) were purchased from Sigma-Aldrich. For inhibition experiments, carbohydrates (100 µg/mL) were added to the cells for 1 hr in RPMI^{FCS} medium prior to the addition of either the GPL-coupled beads or infection with the following strains: S, Δ1, Δ2, Δ3, Gtf3_ov and Δ1,3. Quantification of intracellular CFU and determination of the percentage of infected macrophages and average number of beads per macrophage was performed as described above. The antibody inhibition experiments were performed in a similar manner, except that the cells were incubated with the specific neutralizing antibodies in RPMI^{FCS} medium for 3 hrs at a concentration of 50 µg/ml prior to determination of the percentage of macrophages that phagocytosed the beads and the average number of beads per macrophage.

Glycolipids extraction

GPL were extracted from bacteria grown on 7H10^{OADC} agar plates without detergent, first with chloroform/methanol/0.3%NaCl (9:10:3, v/v/v) and then by chloroform/methanol/0.3%NaCl (5:10:4, v/v/v). The combined solvent extracts were mixed for 5 min with chloroform and 0.3%NaCl (1:1, v/v) and centrifuged at 3,000 *g* for 5 min to separate the lower-organic phase from the aqueous phase. The upper aqueous layer was discarded and the lower-organic phase was evaporated under a stream of nitrogen and resuspended in chloroform/methanol (2:1, v/v). Apolar and polar lipids were subjected to TLC analysis using Silica gel 60 F_{254} plates (Merck). GPL were separated using chloroform/methanol/water (90:10:1, v/v/v) and PIMs using chloroform/acetic acid/methanol/water (40/25/6/3, v/v/v/v). Orcinol/sulfuric acid vapor was used to reveal glycolipids while phosphomolybdic acid/ethanol was preferred for peptidolipids.

GPL purification

2-Flash chromatography. The polar lipid fraction was impregnated with silica gel overnight. Impregnated silica was added on the top of a low-pressure silica gel column 15 µm, 4 g (INTERCHIM, France) and a linear gradient

was applied from 100 % chloroform to reach 85 % chloroform - 15 % methanol at 20 min using Puriflash (INTERCHIM, France). This ratio was maintained for 10 min. 6 mL fractions were collected and the presence of glycolipids was assayed by TLC. Fractions containing GPL were checked by MS then pooled, dried and dissolved in chloroform/methanol (2:1, v/v).

1-TLC. A second round of purification was performed: 150 µL of the GPL fraction was spotted on a 150 mm band on a 60 µm silica TLC plate with a glass back (20 cm x 20 cm) and migrated in a solution of chloroform/methanol/water (90:10:1, v/v/v). Plate sides were sprayed with orcinol/sulfuric acid and charred while the plate center was reversibly stained with iodide vapor to label lipids. GPL-containing silica was scraped and GPL were extracted in chloroform/methanol (2:1, v/v) under sonication for 1 hr. Silica was filtrated on glass wool and 0.2µm PTFE syringe filter. After TLC and MS analysis, purified GPL-1b was pooled from Δ1 and Δ1,3 while purified GPL-2a was prepared from Δ3 and WT *M. abscessus.* Purified GPL-2b was prepared from Δ1 and GPL-3 from WT *M. abscessus*.

Glycolipids saponification

10 µL of polar lipid fraction was dried under nitrogen, 200 µL sodium hydroxide 0.1 M in chloroform/methanol (2:1, v/v) was added and heated for 2 hrs at 37° C. After the reaction, 1 mL chloroform and 1 mL water were added. The mixture was agitated for 1 min then centrifuged for 30 sec. The lower chloroform phase was dried under nitrogen and dissolved in 150 µL chloroform/methanol (2:1, v/v).

Monosaccharides composition by GC/MS

For hydrolysis step, 1 µg mesoinositol, 20 µL GPL fraction and 1 mL trifluoroacetic acid 3 M were mixed then heated 4 hrs at 80° C and dried overnight. The reduction step was conducted for 4 hrs at room temperature in 500 µL $NaBH_4$ 10 mg/mL in 2 M NH_4OH. The reaction was stopped with dropwise addition of concentrated glacial acetic acid. Samples were dried at 55° C under a nitrogen stream by co-distillation with methanol/acetic acid three times, desiccated overnight and incubated in 500 µL anhydride acetic 4 hrs at 80° C. The reaction products were extracted three times with chloroform/water. The chloroform-rich phase was then filtered, dried and dissolved in 100 µL chloroform. 1 µL of monosaccharide derivatives was injected in splitless mode on a Solgel 1 MS 30 m x 0.25 mm x 0.25 µm capillary column with the following gradient temperature: 120° C to 230° C at 3° C/min, then to 270° C at 10° C/min. Compounds were detected after electronic impact at 70 eV on a HP-7820 gas chromatograph coupled to a 5976 single quad (Agilent Technologies, Santa Clara, US) in full scan mode from 50 Da to 500 Da. To quantify GPL prior to bead coating, similar experiments were conducted using 8 and 40 µg of Rha, 6-dTal and Man as external standards (Carbosynth, UK).

Bead coating

The protocol used was reported earlier (Schlesinger et al., 1994). Briefly, 15 µL containing $5.4.10^8$ of 1µm fluorescent (Em / Ex ; 505 nm / 515 nm) sulfate microspheres beads (THERMOFISHER SCIENTIFIC, San Jose, USA) were washed twice with 985 µL carbonate/bicarbonate 0.05M and resuspended in 1 mL PBS. Next to GC quantification, 40 µg of purified GPL were dissolved in 50 µL of water saturated butanol before adding 950 µL of beads in PBS. Coating was performed 2 hrs at 37° C under gentle agitation. Beads were washed twice with 1 mL PBS and blocked for 1 hr with 1 mL 5% BSA in PBS at 37 °C under gentle agitation. Beads were washed twice then transfer for storage at -20°C with 0.5% BSA in PBS. The detection of deoxyhexoses as trimethylsilyl derivatives in the preparation tube was negative. The bead shape was checked with confocal microscopy.

MALDI-TOF mass spectrometry

5 µL of 10 mg/mL dihydroxybenzoïc acid (DHB) in chloroform/methanol (1:2, v/v) were mixed with 5 µL of the sample extract in chloroform/methanol (2:1, v/v). 1-2 µL of sample was spotted on MALDI plate with a glass capillary tube. MS and MS^n spectra were acquired on an Axima Resonance (Shimadzu, Kyoto, Japan) in reflectron mode. For MS^2 experiments, ion selection was set from 250 to 500 and collision energy was tuned from 300 eV to 600 eV.

Nuclear Magnetic Resonance

Flash purified GPL were dried and dissolved in a mixture of $CDCl_3/CD_3OD$ (2:1, v/v) with 0.03 % trimethylsilane (EURISOTOP, France) three times then dissolved in a final volume of 270 µL. The sample was then introduced into a 3 mm glass tube (SHIGEMI, Allison Park, PA, US). A TBI probe was used to observe ^1H and ^{13}C nuclei at 293°K on an AVANCE II system (BRUKER BIOSPIN GmbH, Germany). Impulsion sequences used for homonuclear and heteronuclear experiments were from the manufacturer. After acquisition, phase correction and calibration on methanol signals were performed for $\delta\ ^1H$ and $\delta\ ^{13}C$.

Statistical analyses

Statistical analyses were performed on Prism 9.0 (Graphpad) and detailed for each figure legend: *$P \leq 0.05$, ** $P \leq 0.01$, and *** $P \leq 0.001$.

SUPPLEMENTAL INFORMATION

Figure S1. The unmarked deletion strategy was used to individually or simultaneously delete the three *M. abscessus* genes that encode GPL glycosyltransferases

(A) The *gtf1* gene is sandwiched between the *rmt4* and *atf1* genes; *gtf2* is wedged between *rmt3* and *fmt*; finally, *gtf3* is positioned to the right of *rmlA* and to the left of *rmlB*. Plasmids pUX1-katG-*gtf1*, pUX1-katG-*gtf2* and pUX1katG-*gtf3* were generated to remove the *Mab-gtf1*, *Mab-gtf2* and *Mab-gtf3* genes by double-homologous recombination. Genomic fragments of 1039, 1031 and 1014 bp corresponding of the left arms DNA sequences of the *Mab-gtf1*, *Mab-gtf2* and *Mab-gtf3*, respectively, were amplified by PCR and subcloned into the PacI and MfeI sites of plasmid pUX1-*katG* (Daher et al., 2020; Richard et al., 2019). The right arms DNA sequences of the *Mab-gtf1*, *Mab-gtf2* and *Mab-gtf3* were amplified from genomic DNA by PCR (1043, 1033 and 998 bp, respectively) and cloned into the MfeI and NheI sites of plasmid pUX1-*katG*. The resulting suicide plasmids (lacking motifs for episomal replication or mobile elements promoting chromosomal integration) were used to transform *M. abscessus*, in which they could only propagate *via* homologous recombination between the cloned sequences and their chromosomal homologous sequences. The first recombination events were selected in the presence of kanamycin. Single red fluorescent tdTomato-expressing clones were subjected to the second round of recombination selected on isoniazid and screened for double crossover phenotypes, i.e., loss of red fluorescence, sensitivity to kanamycin, and resistance to isoniazid. Dotted lines represent the size (indicated above each line) of the expected PCR products in *M. abscessus* WT, Δ1, Δ2, and Δ3. Black arrows represent the primers used for PCR analysis.

(B) PCR analysis demonstrating the deletion of *gtf1, gtf2, gtf3, gtf1/gtf2, gtf1/gtf3, gtf2/gtf3* and *gtf1/gtf2/gtf3* in the various mutant strains. Genomic DNA from WT bacteria was used to amplify the intact *gtf1* or *gtf2* or *gtf3* loci. Amplicons were subjected to sequencing to confirm the deletion of *gtf1* or *gtf2* or *gtf3* genes.

(C) The *gtf1, gtf2* and *gtf3* mutants were complemented using the integrative vector pMV306 containing the open reading frames of *gtf1* or *gtf2* or *gtf3* fused to a HA tag on the C-terminal side. In the complemented strains, *gtf1* expression was controlled by the *hsp60* promoter while *gtf2* and *gtf3* expression levels were dictated by their own promoters respectively. The molecular weight in kDa is indicated for each glycosyltransferase fused with an HA tag at the bottom of each schematic illustration.

(D) Western blot analysis of Δ1_C, Δ2_C, and Δ3_C strains expressing the different Gtf proteins in fusion with an HA tag (left panel). KasA was included as a loading control (right panel).

Figure S2. Mass spectrometry analysis of deacetylated polar glycolipids.

(A) MALDI-TOF-MS spectra acquired in positive mode of *Mab S* show signals patterns at *m/z* 1174/1202 for deacetylated GPL-2a and 1320/1348 for deacetylated GPL-3. Single deletion of gtf3 clearly showed a similar cluster at *m/z* 1174/1202 for deacylated GPL-2a.

(B) MALDI-TOF-MS^2 spectrum in positive mode of wild-type deacylated GPL-3 parent ion at *m/z* 1348 shows fragment Da (in bold) compared to MS^2 spectrum of native GPL-3 **(Figure 1E)**, which confirms the positions of acetyl groups on dTal.

(C) MALDI-TOF-MS^2 spectrum in positive mode of wild-type deacylated GPL-2a parent ion at *m/z* 1202 shows fragments ions with -84Da (in bold) compared to MS^2 spectrum of native GPL-2a (Figure 1D), which confirms the positions of acetyl groups on dTal.

Figure S3. Structural analysis of the GPL profile in various *gtfs* mutants.

(A) MALDI-TOF-MS spectrum in positive mode of low R_f GPL-2b isolated from Δ1 (upper panel) and MALDI-TOF-MS^2 spectrum of parent ion at *m/z* 1188 (lower panel) confirm the presence of a di-*O*-methylated di-rhamnosyl motif and the absence of dTal.

(B) MALDI-TOF-MS spectrum in positive mode of high R_f GPL-1b isolated from Δ1 (upper panel) and MALDI-TOF-MS^2 spectrum of parent ion at *m/z* 1042 (lower panel) confirm the presence of a single di-*O*-methyl Rha and the absence of dTal.

(C) MALDI-TOF-MS spectrum in positive mode of high R_f GPL-1b isolated from Δ1,3 (upper panel) and MALDI-TOF-MS^2 spectrum of parent ion at *m/z* 1042 (lower panel) confirm the presence of a single di-*O*-methyl Rha and the absence of dTal.

(D) ^{1}H-^{13}C HSQC spectrum of high R_f GPL-1b from Δ1,3 confirmed the presence of a single di-*O*-methyl Rha substituted in C3 and C4 positions. Another *O*-methyl group observed at ^{1}H-^{13}C 3.28/56.8 ppm was attributed to 3-methoxy lipid as previously described (Daher et al., 2020). Alpha-carbons from each amino-acid were also detected and correlated with respective beta carbons by ^{1}H-^{1}H COSY spectrum (data not shown). Two amino-alcohol alaninol and valinol are also observed, which confirmed the presence of the two variants GPL.

(E) TLC analysis of GPL fractions developed in $CHCl_3/CH_3OH/H_2O$ (90/10/1) and stained with phosphomolybdic acid revealed two intense bands corresponding to accumulated lipopeptide for mutant Δ1,2. Lower intensity was observed for mutant Δ*2*. The circled band was further analyzed by MALDI-MS and NMR.

(F) MALDI-TOF MS spectrum in positive mode of the lipopeptide isolated from Δ1,2 (**Figure S3E** circled lower Rf band) showed two ions at *m/z* 838 and 866. This mass shifted at *m/z* 866 and 894 for the higher R_f accumulating band.

(G) ^{1}H-^{1}H COSY spectrum allows the amino-acid sequencing of the lipopeptide isolated from Δ1,2 (**Figure S3E**, circled lower Rf band). Higher Rf lipopeptide did not show the presence of alaninol residue (data not shown).

Figure S4. Deletion of *gtf1* or *gtf2* promotes bacterial sedimentation in *in vitro* culture, abolishes the ability of bacilli to form thick and filamentous biofilms on a liquid surface and increases the susceptibility of *Mab* to rifabutin.

(A) The 3 glycosyltransferases are not involved in the growth of bacilli *in vitro*. Growth curve of parental S, wild-type R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Gtf3_ov, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3 strains in 7H9 medium at 37 °C. Data is representative of three independent experiments.

(B) Bacilli that do not express Gtf1 or Gtf2 sediment rapidly in *in vitro* culture. Like the parental S strain, but unlike the R variant, Δ3 does not sediment rapidly in liquid culture. However, the Δ1, Δ2, double and triple mutant strains do sediment rapidly in *in vitro* culture. Note that the Gtf3-ov strain partially sedimented for the same amount of time. Finally, the Δ2_C variant sedimented like an R strain but forming a hill at the bottom of the glass tube. Experiments were carried out three times.

(C) Strains belonging to the S morphotype form filamentous biofilms on the surface of the liquid M63 medium while R variants establish a thin layer dotted with spaced white spots. Pellicle-forming capabilities of *M. abscessus* S, R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Gtf3_ov, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3 strains. Experiments were carried out three times.

(D) Biofilm formation (liquid/air interface) and planktonic growth (solid/liquid interface) in glass tubes. The *M. abscessus* S, R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Gtf3_ov, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3 strains were grown, in static conditions at 37 °C, in M63 medium and in glass tubes. Pictures were taken after 6 days of incubation.

(E) Absence of gtf1 or gtf2 makes bacilli 4 times more susceptible to rifabutin. Drug susceptibility profile of *M. abscessus* S, R, Δ1, Δ1_C, Δ2, Δ2_C, Δ3, Δ3_C, Gtf3_ov, Δ1,2, Δ1,3, Δ2,3 and Δ1,2,3 strains to various antibiotics (cefoxitin, imipenem, bedaquiline, amikacin and rifabutin) in CaMBH. MIC (µg/mL) were determined following the CLSI guidelines.

Figure S5. Inactivation of the *gtf1* or *gtf2* genes in the smooth morphotype variant gives the bacilli the ability to establish cords *in cellulo.*

One immunofluorescent field per strain was taken at 3 days post-infection showing macrophages infected with *M. abscessus* expressing Tdtomato (red). The surface of the macrophages was detected using anti-CD43 antibodies (green). The nuclei were stained with DAPI (blue). The S, Δ1_C, Δ3 and Δ3_C strains belong to the smooth morphotype and therefore do not form cords *in cellulo*. In contrast to the smooth variants, the rough strains which are R, Δ1, Δ2, Δ2_C as well as the double and triple mutants form intracellular and extracellular cords *ex vivo*. C.: cord. The scale bars for all images taken in normal confocal mode represent 20 µm.

Two images for each of the R, Δ1 and Δ2 strains were taken in Z-stack Airyscan mode to illustrate in 3D the cords established inside the macrophages as well as outside the cells. A Z-stack series was collected with a Zeiss LSM880 Airyscan confocal microscope and images were assembled and reconstructed with Imaris 3D

software. I.C.: intracellular cord, E.C.: extracellular cord. The scale bars for the three 3D images showing either the intracellular cords or the other three extracellular cords represent 3 µm and 5 µm, respectively.

Figure S6. Uptake inhibition was not observed when pre-treated macrophages were infected with the strains: S or Δ1 or Δ2 or Δ3 or Δ1,3. Invasion of the strains (S or Δ1 or Δ2 or Δ3 or Δ1,3) was not reduced in macrophages pretreated with mannan, laminarin and methyl-α-D-glucopyranoside. CFU were determined at 3 hpi. Data are mean values ± SD for four independent experiments (each time in triplicate) (n=12).

Figure S7. Steric hindrance of MR and CR3 receptors reduces the average number of internalized GPL-3-coupled beads per macrophage

(A) Maximum intensity projection reconstruction of two three-dimensional macrophages (red) containing fluorescent beads coupled to GPL-3 (green). The two images were created by merging serial scans of extremely thin (0.18 µm) cell sections, observed under confocal microscopy (LSM880) in Airyscan Z-stack mode at 63x magnification. Showing in the y/z plane the depth of the two macrophages. The nuclei were stained with DAPI (blue).

(B) On average, macrophages pre-treated with mannan or laminarin or methyl-α-D-glucopyranoside ingested less GPL-3-coupled beads than those not pre-treated. Quantification of the average number of beads alone or coupled to GPL-1b, GPL-2a, GPL-2b or GPL-3 per macrophage was performed after 4 hours of phagocytosis. The red bar in each violin plot represents median. The colored area of a violin plot corresponds to the probability density of the data and contains first and third quartiles. Data are mean values ± SD for three independent experiments (each time in triplicate) (n=900 macrophages containing beads per condition). One-tailed Mann−Whitney test: ***$p < 0.001$.

(C) Representative fluorescent images showing the impact of neutralizing anti-CR3 and anti-Dectin1 antibodies on the internalization of GPL-2a- or GPL-3-coupled beads by macrophages. CD43 (red), DAPI (blue), fluorescent beads (green). White arrows indicate the presence of fluorescent beads inside macrophages. Scale bars represent 20 µm.

Movies 1 and 2. These two movies show two different macrophages that phagocytosed 3 or 8 green fluorescent beads coupled to GPL3. Image acquisition was performed using a confocal microscope in AiryScan Z-stack mode. Both videos were generated using the Zen program (blue edition).

Table S1. Primers used in this study. Restriction sites are in bold and underlined.

Number	5'→3' sequence	Restriction site	F (sense), R (antisense)
Cloning in pUX1-*katG*			
1	gaga**TTAATTAA**GAAATATCTGGCGACA TCGGTTCATAAGATTTTTTGAC	PacI	To clone the left arm of *gt* in pUX1-*katG* (F)
2	gaga**CAATTG**GTAACTCGCCAGCACAAATTTCATCGCTTC	MfeI	To clone the left arm of *gt* in pUX1-*katG* (R)
3	gaga**CAATTG**GAAAGTGCTGTCGCTGCCGCG	MfeI	To clone the right arm of *gtf1* in pUX1-*katG* (F)
4	gaga**GCTAGC**GCAATGACAAAGAACAGCAACAGATTCAGACTG	NheI	To clone the right arm of *gtf1* in pUX1-*katG* (R)
5	gaga**TTAATTAA**CTTGACAAAGCGGTACATCCGGTATTCG	PacI	To clone the left arm of *gt* in pUX1-*katG* (F)
6	gaga**CAATTG**CACGCTGACGGATGGTTCGACATC	MfeI	To clone the left arm of *gt* in pUX1-*katG* (R)
7	gaga**CAATTG**GCAGTCTCCGCTCAGATGACCTC	MfeI	To clone the right arm of *gtf2* in pUX1-*katG* (F)
8	gaga**GCTAGC**GATTCCTCTTGATCGCGATGCTGCG	NheI	To clone the right arm of *gtf2* in pUX1-*katG* (R)
9	gaga**TTAATTAA**TTGTCGTAGAAGTACAGCCCCGGTA	PacI	To clone the left arm of *gt* in pUX1-*katG* (F)
10	gaga**CAATTG**TTTCATGACGGTATAGACCTCCCCGAT	MfeI	To clone the left arm of *gt* in pUX1-*katG* (R)
11	gaga**CAATTG**TCTCTGGTCTCGGACCTACGCA	MfeI	To clone the right arm of *gtf3* in pUX1-*katG* (F)
12	gaga**GCTAGC**TTGTTTCTCACGTGCCAACCGCG	NheI	To clone the right arm of *gtf3* in pUX1-*katG* (R)
Primers to verify double homologous recombination			
13	GTGTTCGACTCGGTGGTCGTGATTCAC	-	To check double homologous recombinatio Δ1 (F)
14	GTTTTTCCACGACAAACCAACTGAAAGCC	-	To check double homologous recombinatio Δ1 (R)
15	TCGACGACGTTCATCACTCAACTATACCTATC	-	To check double homologous recombinatio Δ2 (F)
16	CTTTCAGTTGGTTTGTCGTGGAAAAACGCG	-	To check double homologous recombinatio Δ2 (R)
17	GAATCGAACGTTCCGGTGTCCAG	-	To check double homologous recombinatio Δ3 (F)
18	CATCGCGAAACTTCCCGACAACG	-	To check double homologous recombinatio Δ3 (R)
Cloning in pMV306-KanR			
19	gaga**CAATTG**CCATGAAATTTGTGCTGGCGAGTTACGGAACG	MfeI	To clone the *gtf1-HA* ORF pMV306 (F)
20	gaga**GTCGAC**CTAAGCGTAATCTGGAACATCGT ATGGGTATGCCGAACACCTCGCCTGCG	SalI	To clone the *gtf1-HA* ORF pMV306 (R)
21	gaga**TCTAGA**TGGACGACTCCGTACCATCGACGAA	XbaI	To clone the *gtf2-HA* ORF pMV306 (F)

22	gaga**GTCGAC**CTAAGCGTAATCTGGAACATCGTAT GGGTACACCGCTTTTCGGCCGGCGGT	SalI	To clone the *gtf2-HA* ORF pMV306 (R)
23	gagaga**GGTACC**ACCAGCGGTTTGTCGTAGACCG	KpnI	To clone the *gtf3-HA* ORF pMV306 (F)
24	gagaga**GAATTC**CTAAGCGTAATCTGGAACATCGTA TGGGTATGGGTGGGAAGCCTTTGCCTC	EcoRI	To clone the *gtf3-HA* ORF pMV306 (R)
Cloning in pMV261-KanR			
25	gagaga**TGGCCA**CCATGGCCGAGAATATTGACGTCCATTCTG	MscI	To clone the *gtf3-HA* ORF pMV261 (F)
26	gagaga**GAATTC**CTAAGCGTAATCTGGAACATCGT ATGGGTATGGGTGGGAAGCCTTTGCCTC	EcoRI	To clone the *gtf3-HA* ORF pMV261 (R)

Figure 1

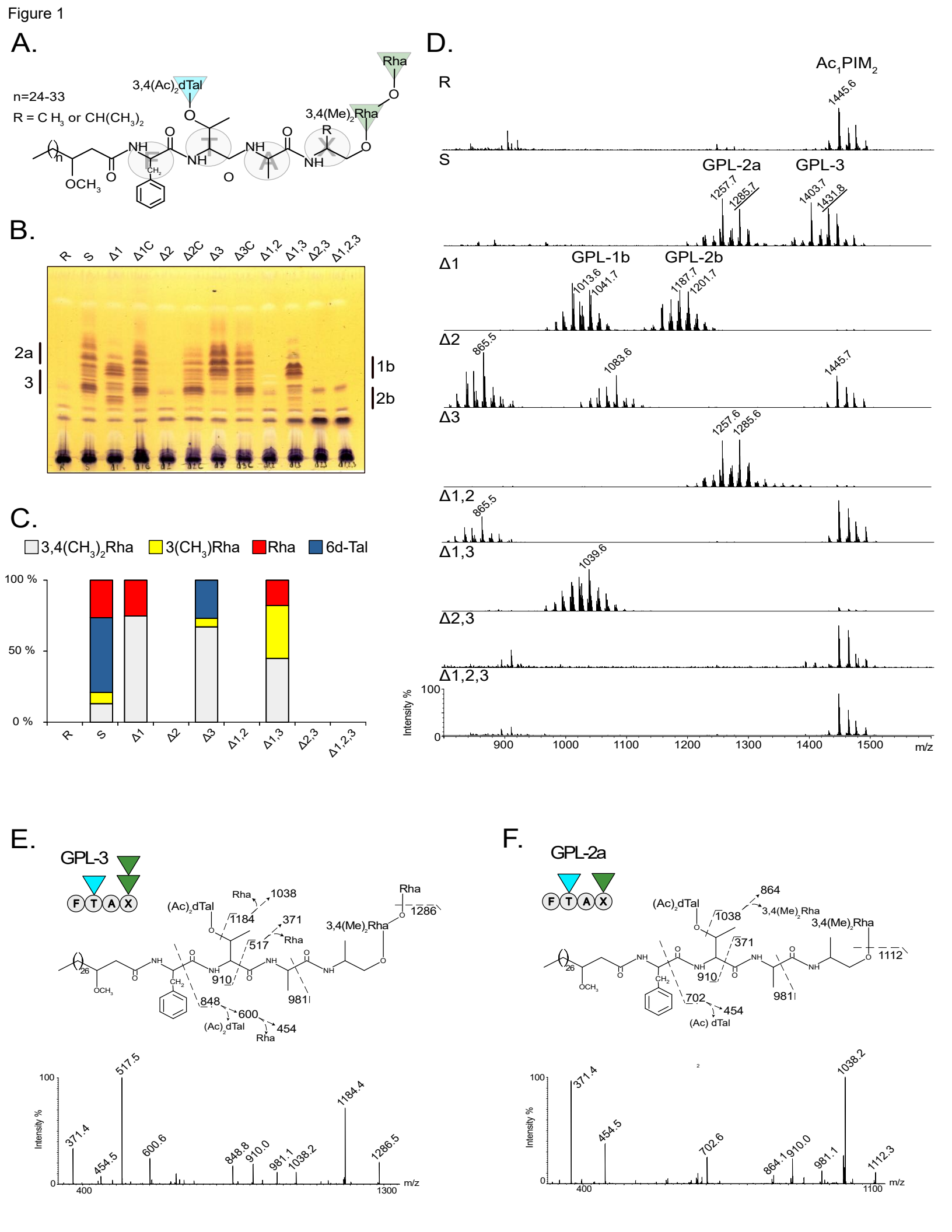

A.
n=24-33
R = C H3 or CH(CH3)2
3,4(Ac)2dTal
3,4(Me)2Rha
Rha
OCH3
B.
R S Δ1 Δ1C Δ2 Δ2C Δ3 Δ3C Δ1,2 Δ1,3 Δ2,3 Δ1,2,3
2a
3
1b
2b
C.
3,4(CH3)2Rha
3(CH3)Rha
Rha
6d-Tal
100 %
50 %
0 %
D.
Ac1PIM2
1445.6
GPL-2a
GPL-3
1257.7
1285.7
1403.7
1431.8
GPL-1b
GPL-2b
1013.6
1041.7
1187.7
1201.7
865.5
1083.6
1445.7
1257.6
1285.6
1039.6
Intensity %
m/z
E.
GPL-3
(Ac)2dTal
1038
1184
371
517
910
848
600
454
981
1286
3,4(Me)2Rha
F.
GPL-2a
864
1038
371
910
702
454
981
1112
(Ac) dTal

Figure 2

A.

GPLs glycoforms	Peptido Lipid	GPL-1a	GPL-1b	GPL-2a	GPL-2b	GPL3
	F T A X	F T A X	F T A X	F T A X	F T A X	F T A X
(detected m/z)	(866)		(1014/1042)	(1258/1286)	(1160/1188)	(1404/1432)
(saponified)	(-)		(-)	(+)	(-)	(+)
R						
S				X		X
Δ1			X		X	
Δ2	X					
Δ3				X		
Δ1,2	X					
Δ1,3			X			
Δ2,3						
Δ1,2,3						

B.

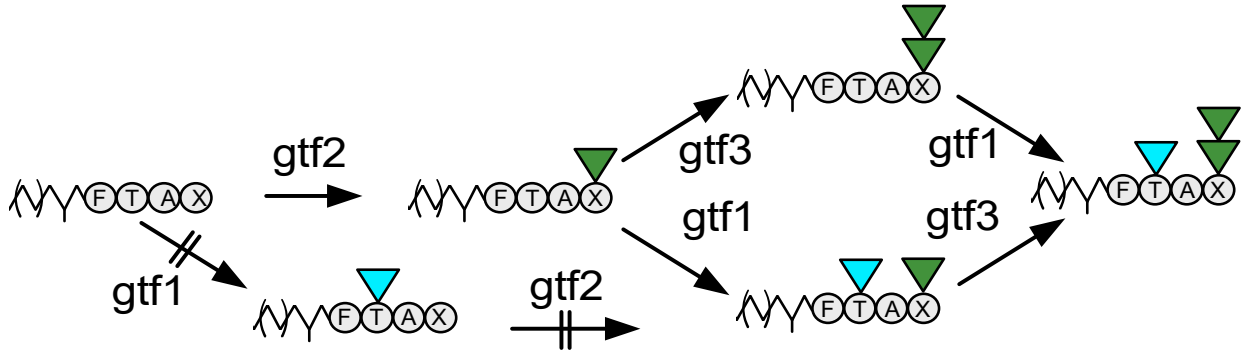


Figure 3

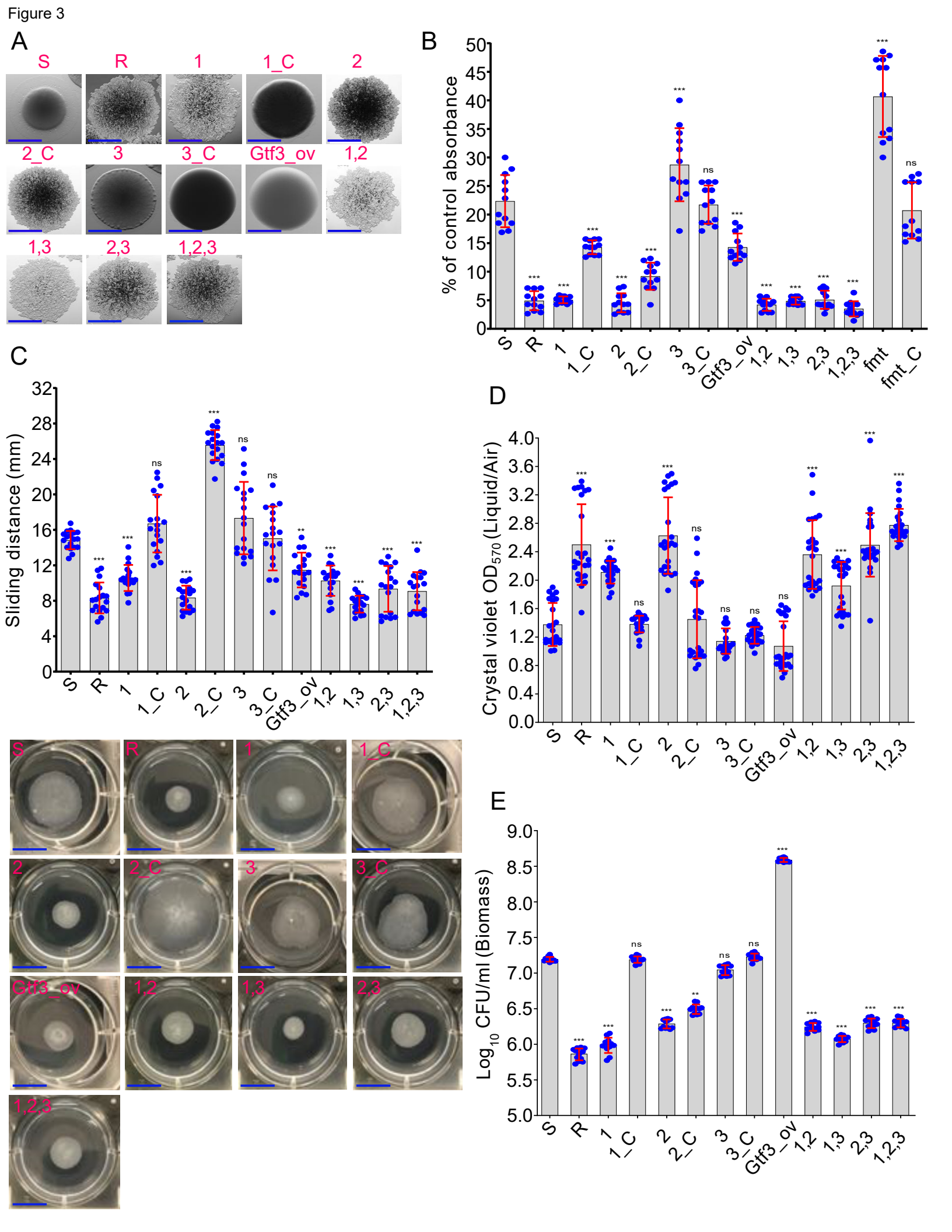

A
S
R
1
1_C
2
2_C
3
3_C
Gtf3_ov
1,2
1,3
2,3
1,2,3
B
% of control absorbance
fmt
fmt_C
C
Sliding distance (mm)
D
Crystal violet OD570 (Liquid/Air)
E
Log10 CFU/ml (Biomass)
ns

Figure 4

Click here to access/download;Figure;Figure 4.tif

gtf fish figure_230521.pdf

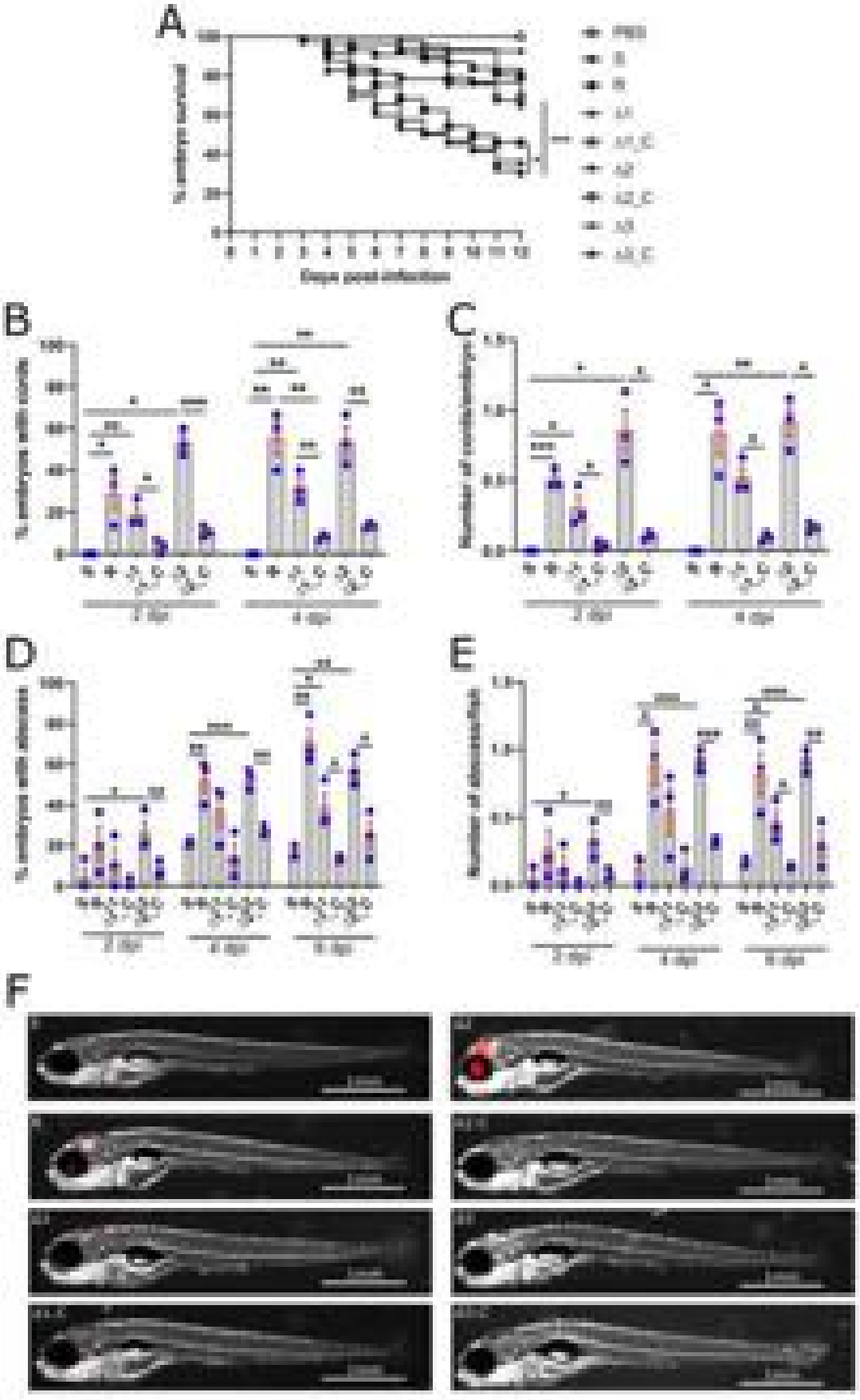


Figure 5

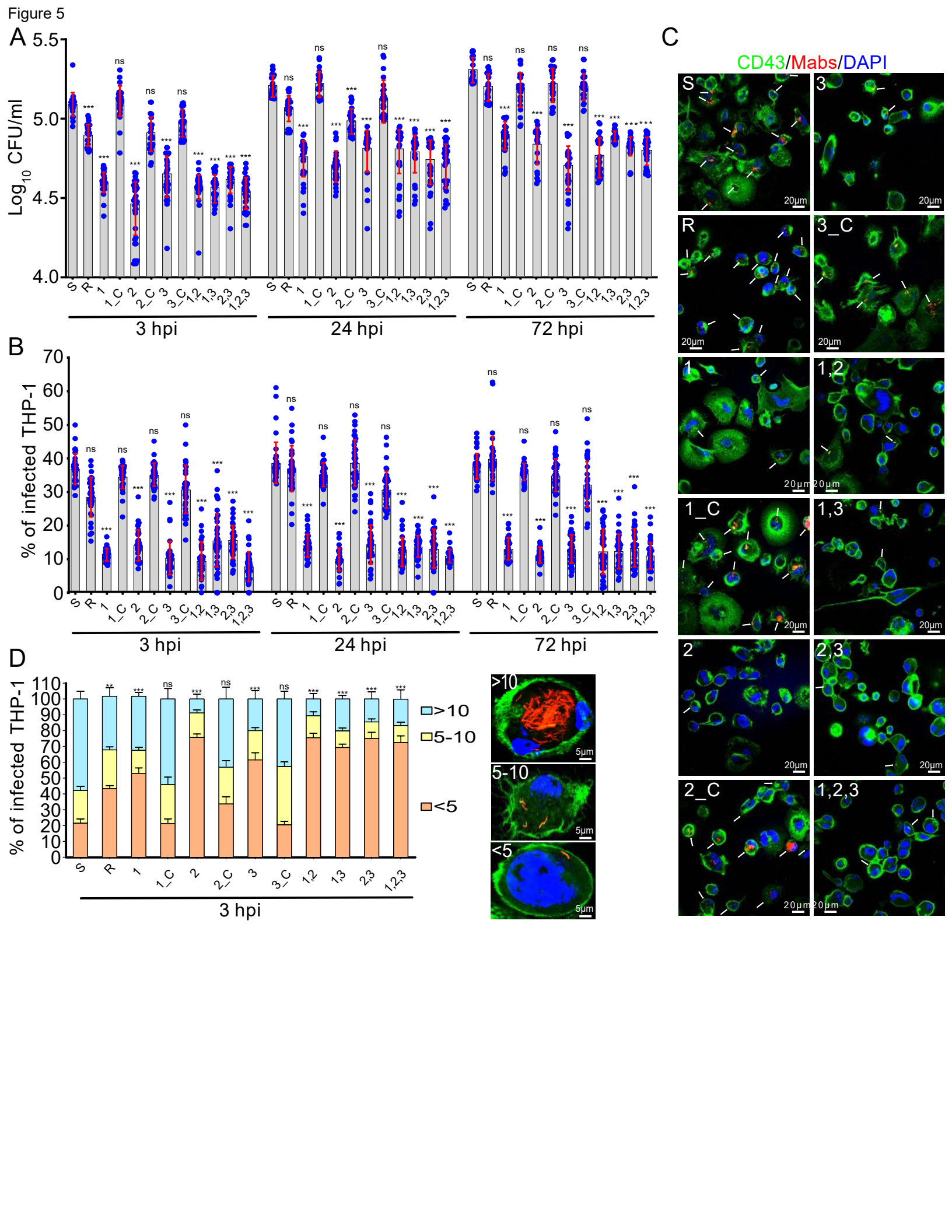

A
Log10 CFU/ml
3 hpi
24 hpi
72 hpi
B
% of infected THP-1
3 hpi
24 hpi
72 hpi
C
CD43/Mabs/DAPI
S
R
3
3_C
1
1,2
1_C
1,3
2
2,3
2_C
1,2,3
20µm
D
% of infected THP-1
3 hpi
>10
5-10
<5
5µm

Figure 6

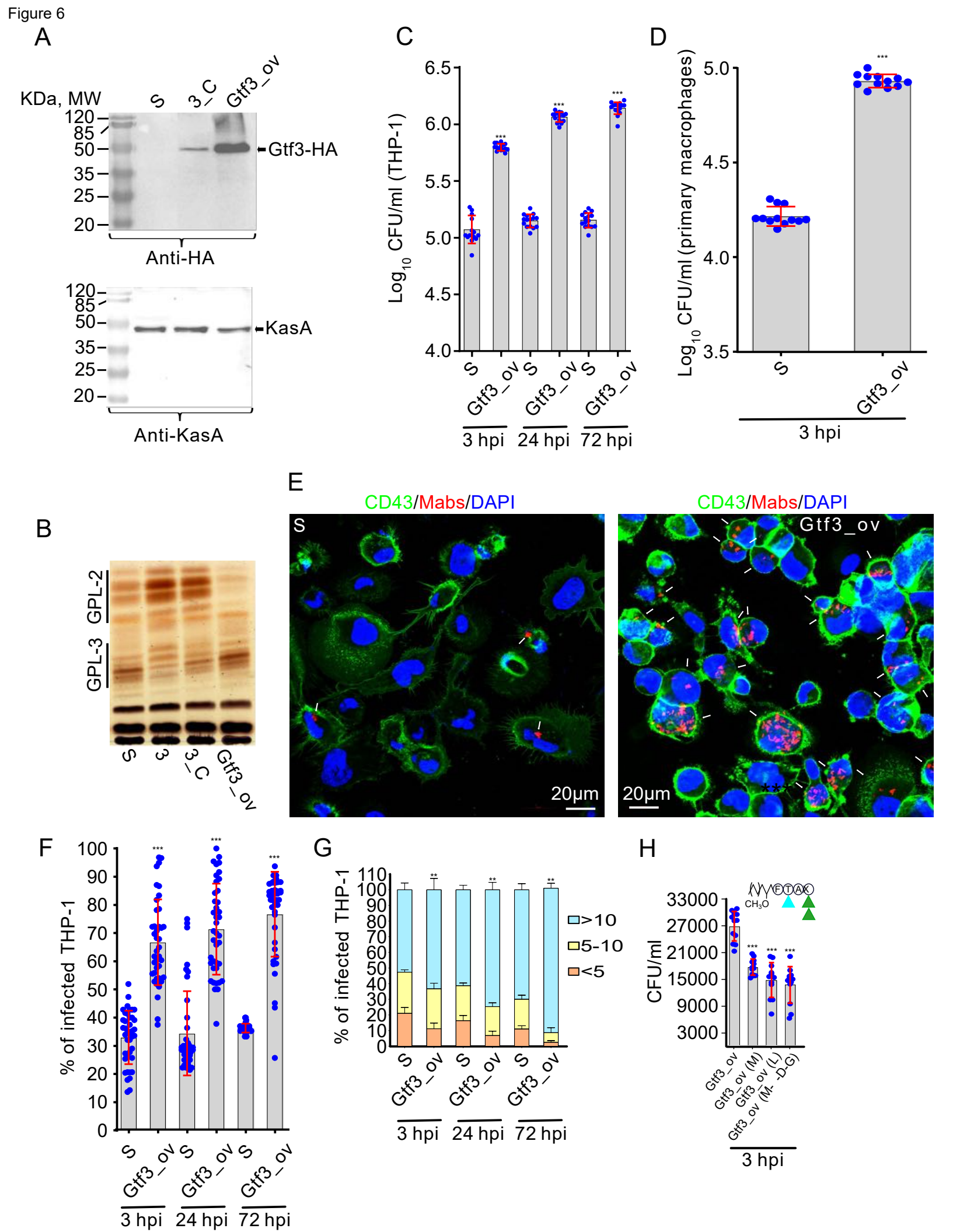
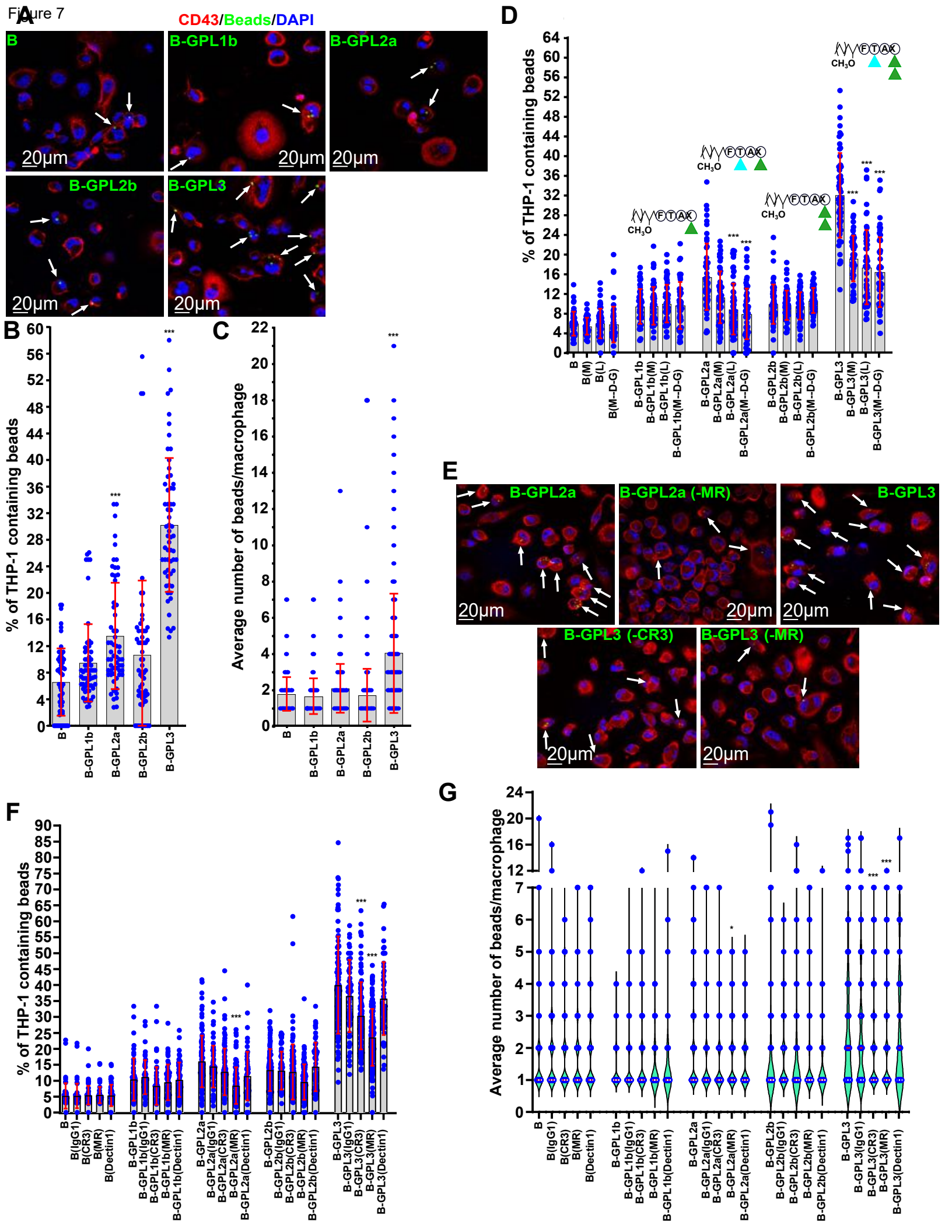
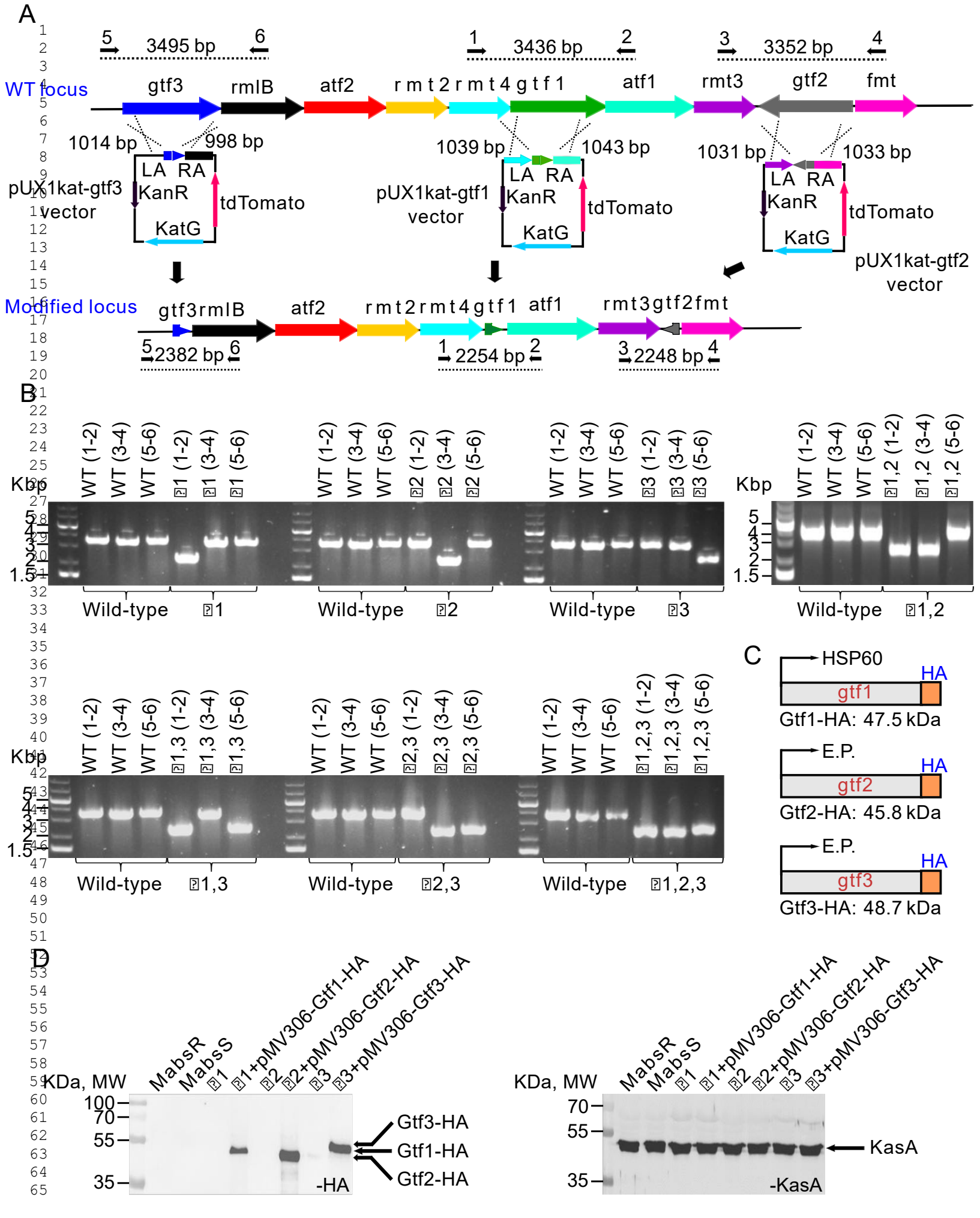

A
KDa, MW
S
3_C
Gtf3_ov
120
85
50
35
25
20
Gtf3-HA
Anti-HA
KasA
Anti-KasA
B
GPL-2
GPL-3
C
Log10 CFU/ml (THP-1)
3 hpi
24 hpi
72 hpi
D
Log10 CFU/ml (primary macrophages)
E
CD43/Mabs/DAPI
20μm
F
% of infected THP-1
G
>10
5-10
<5
H
CFU/ml
Gtf3_ov
Gtf3_ov (M)
Gtf3_ov (L)
Gtf3_ov (M- -D-G)

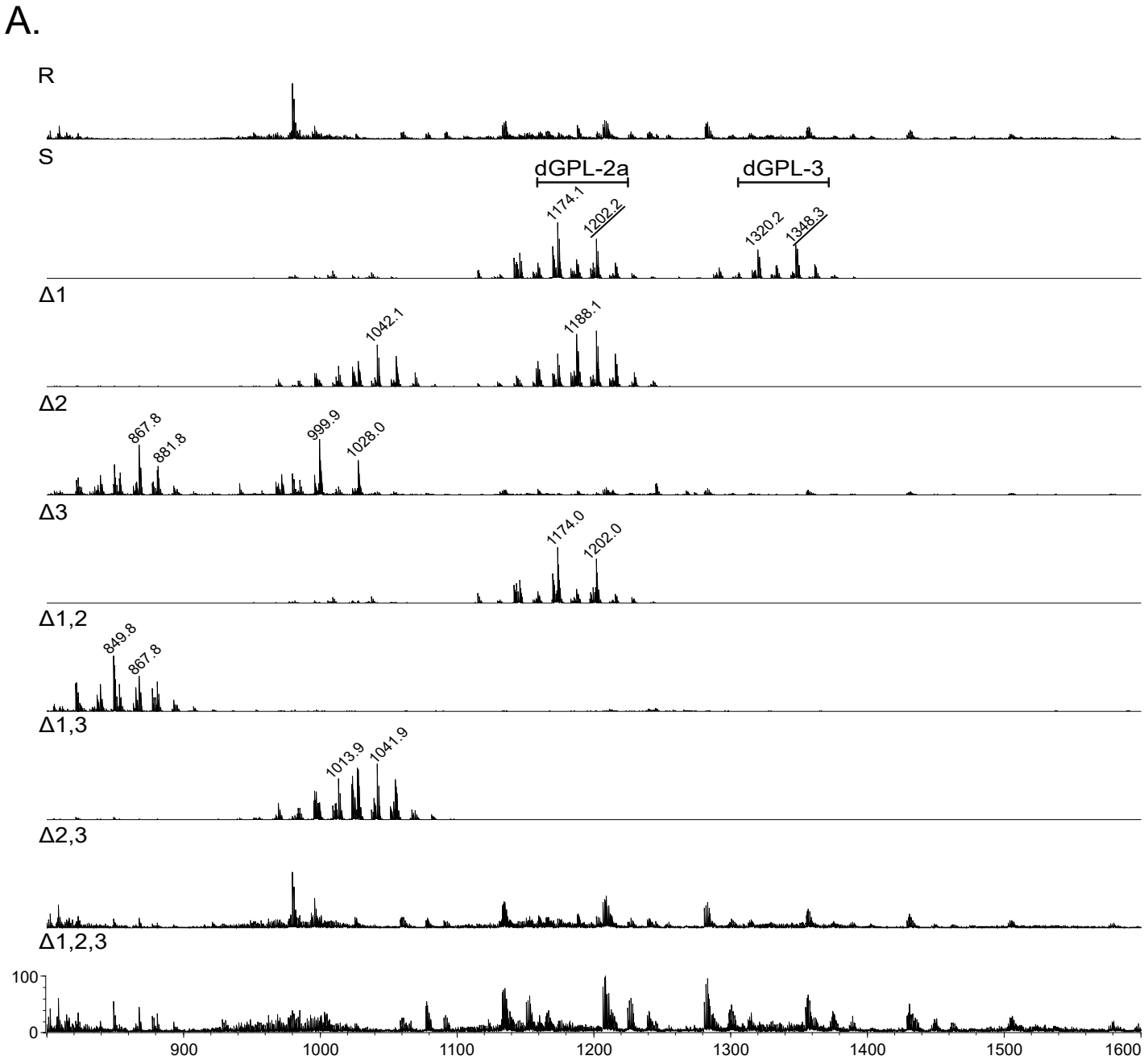
Figure 7

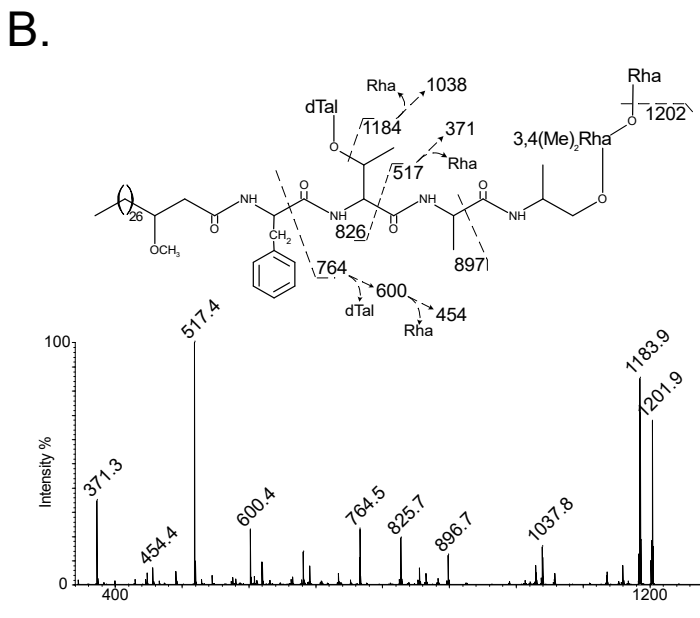

A
CD43/Beads/DAPI
B
B-GPL1b
B-GPL2a
B-GPL2b
B-GPL3
20μm
B
% of THP-1 containing beads
C
Average number of beads/macrophage
D
% of THP-1 containing beads
CH3O
FTAX
E
B-GPL2a
B-GPL2a (-MR)
B-GPL3
B-GPL3 (-CR3)
B-GPL3 (-MR)
F
% of THP-1 containing beads
G
Average number of beads/macrophage

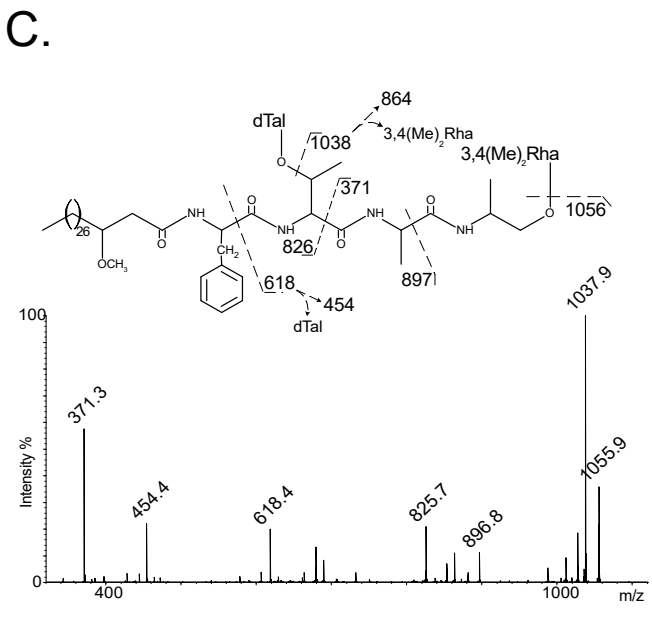
Supplemental Figure 1



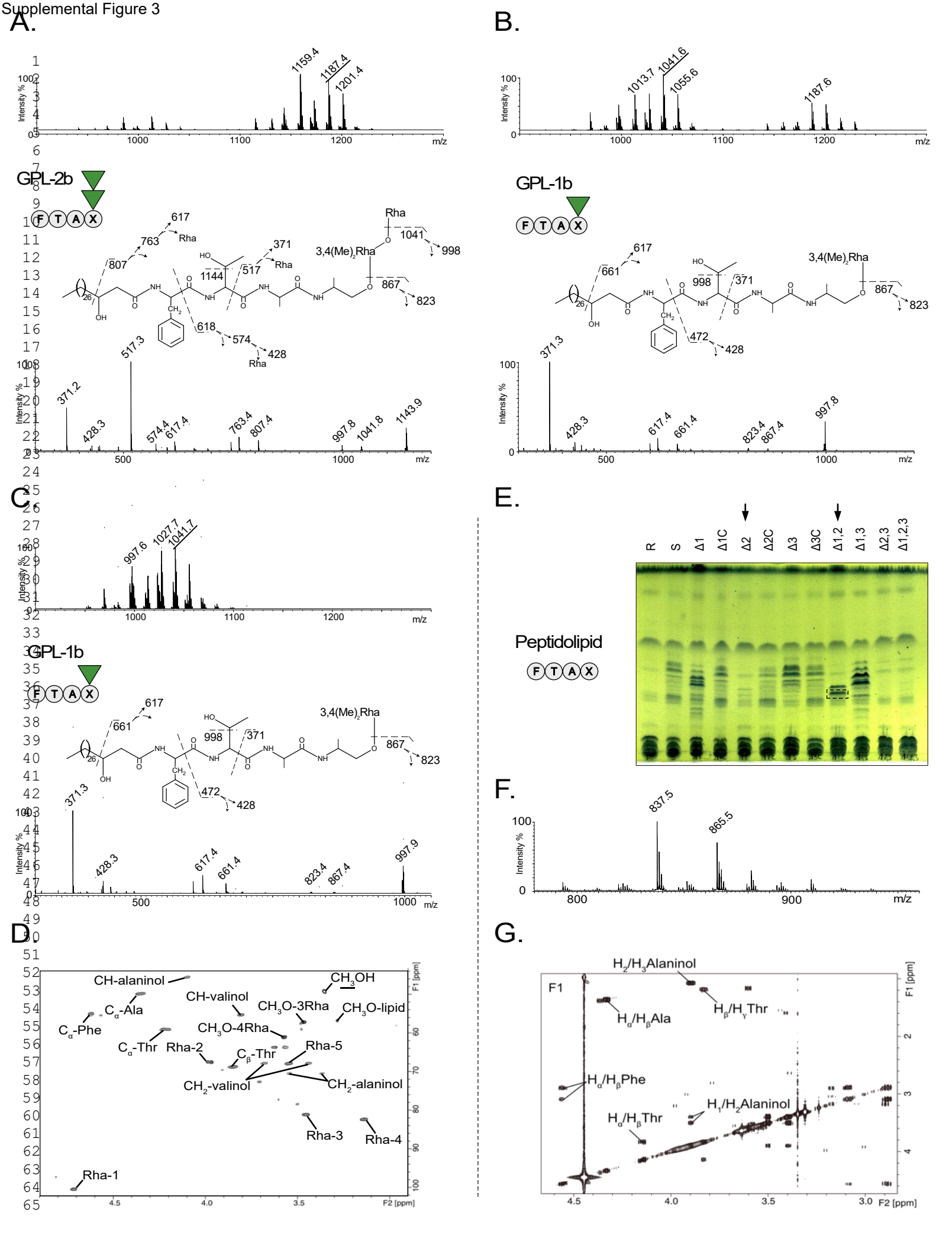
Supplemental Figure 2



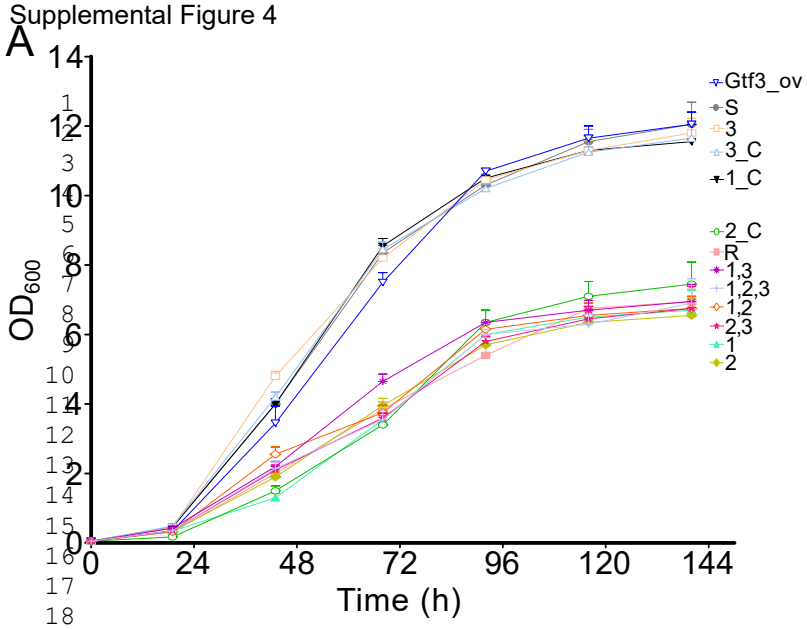


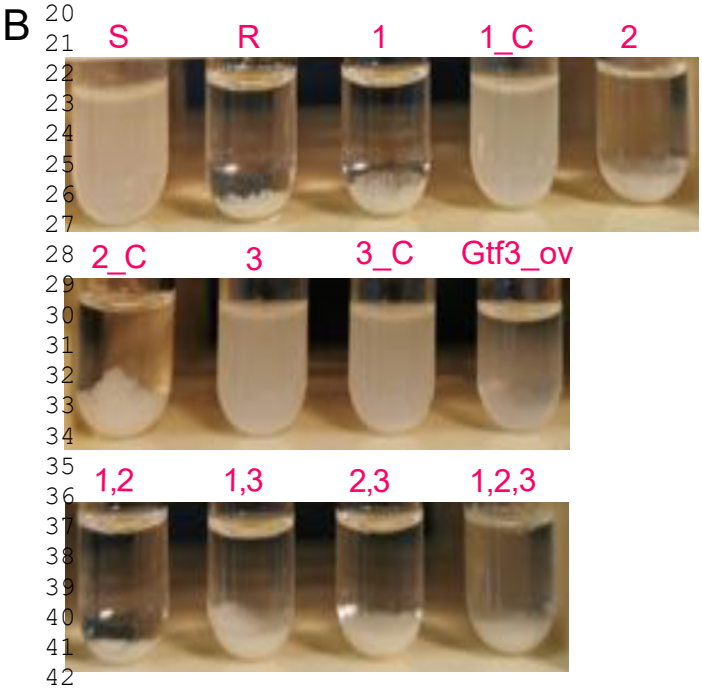


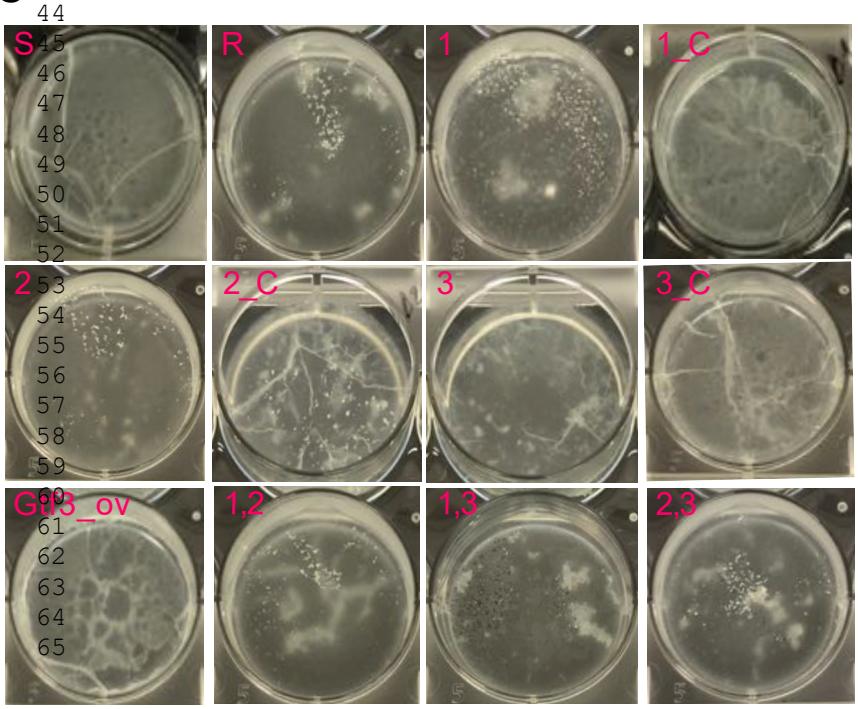
Supplemental Figure 3

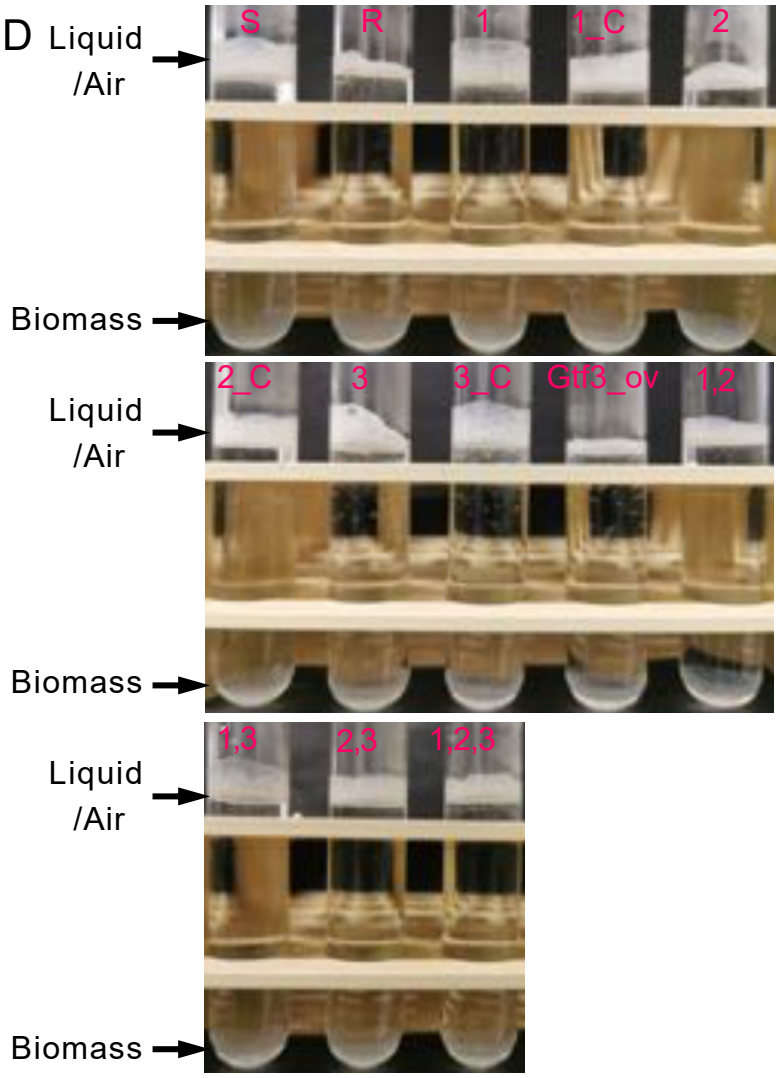


Supplemental Figure 4





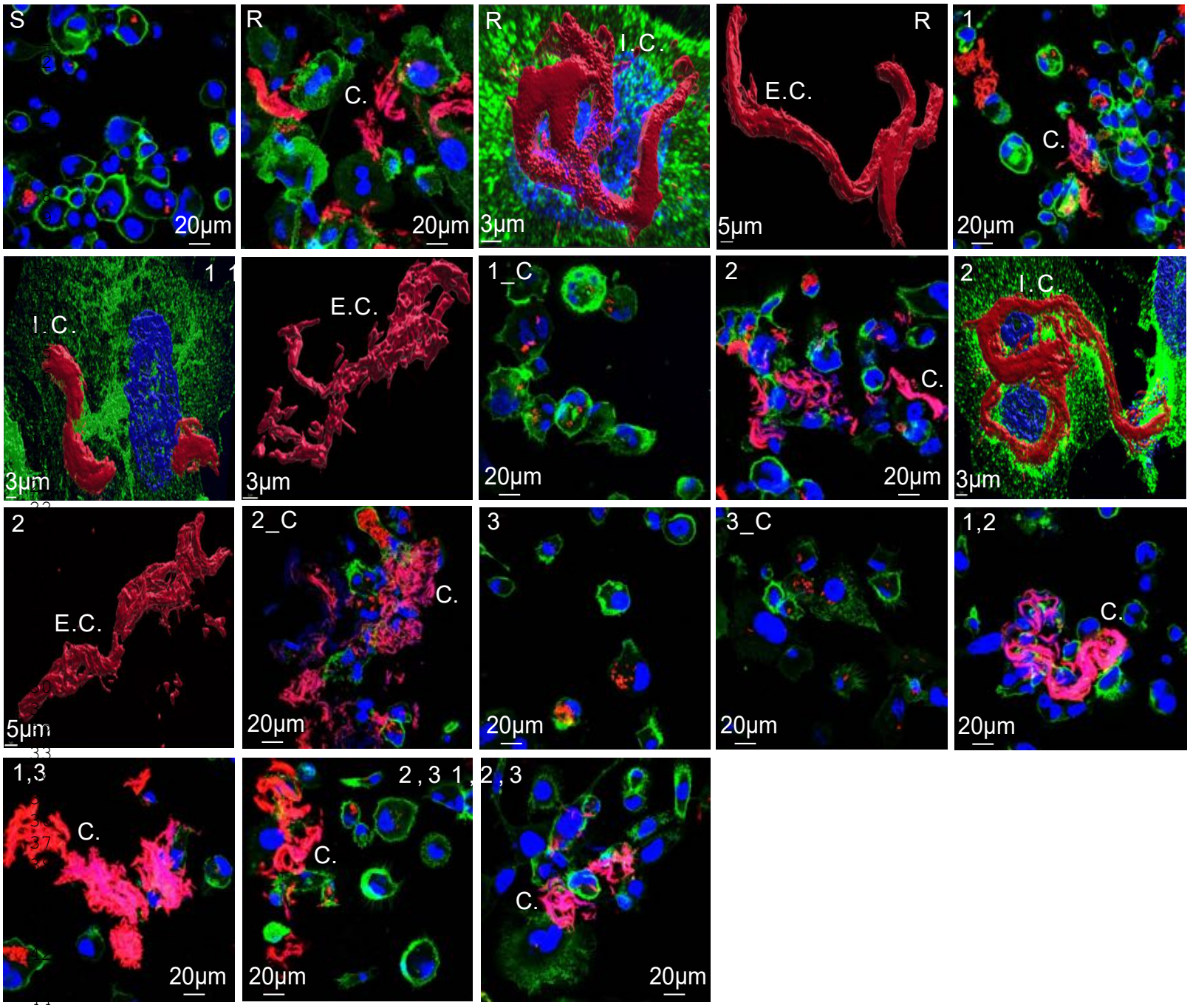




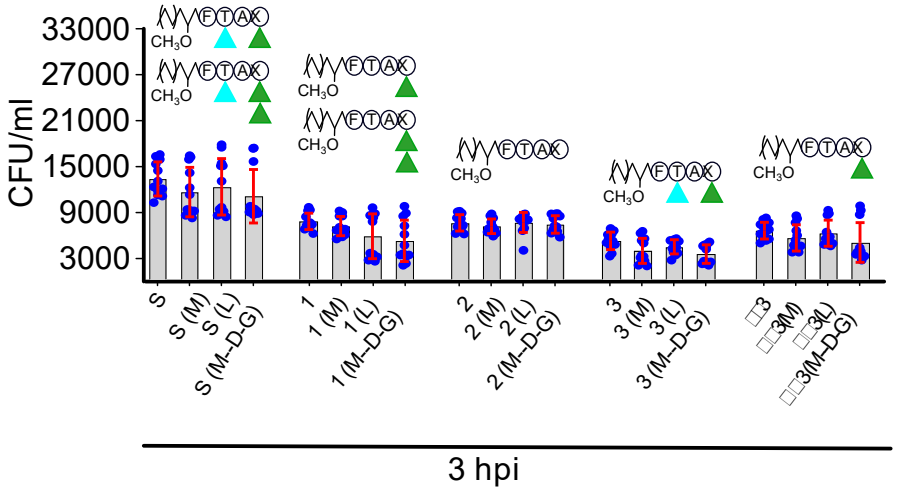
E

	CMI µg/ml				
	Cefoxitin	Imipenem	Bedaquiline	Amikacin	Rifabutin
S	64	16	0.06	32	50
R	64	32	0.06	16	12.5
1	64	32	0.06	32	12.5
1_C	64	16	0.06	32	50
2	64	32	0.06	32	12.5
2_C	64	32	0.03	32	50
3	64	16	0.06	32	50
3_C	64	16	0.06	32	50
Gtf3_ov	64	16	0.06	32	50
1,2	64	32	0.03	32	12.5
1,3	64	32	0.03	16	12.5
2,3	64	32	0.06	16	12.5
1,2,3	64	32	0.03	16	12.5

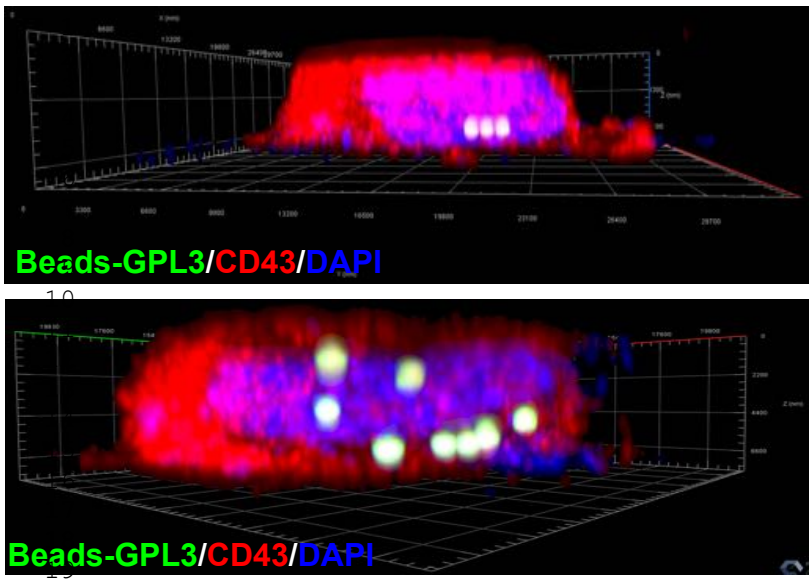
Supplemental Figure 5

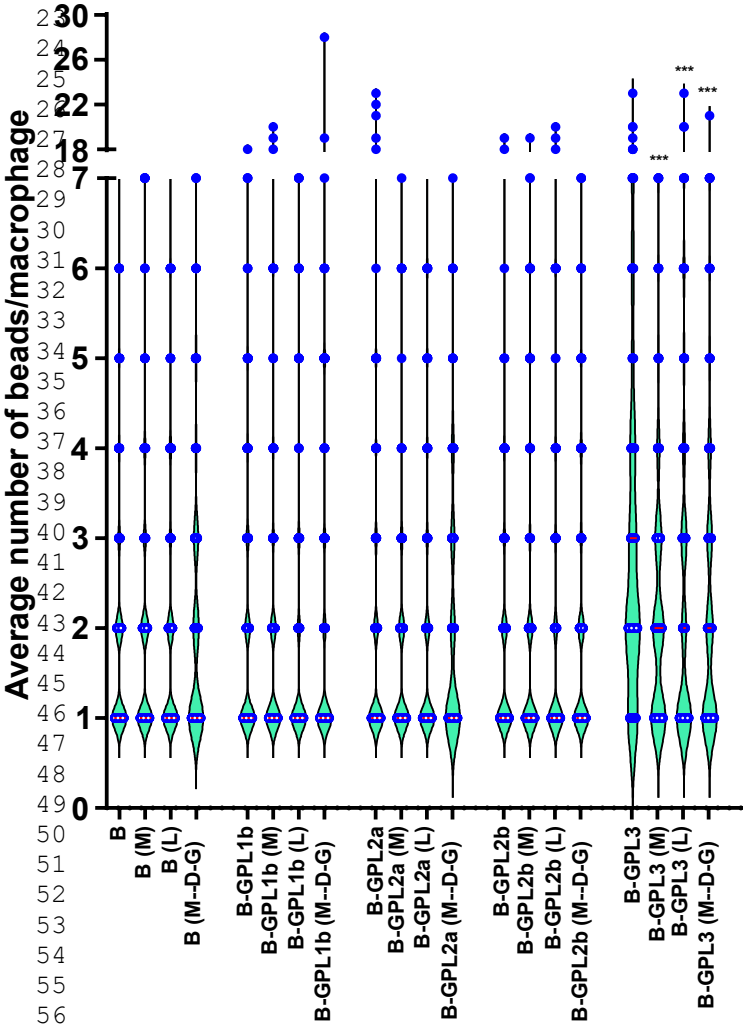


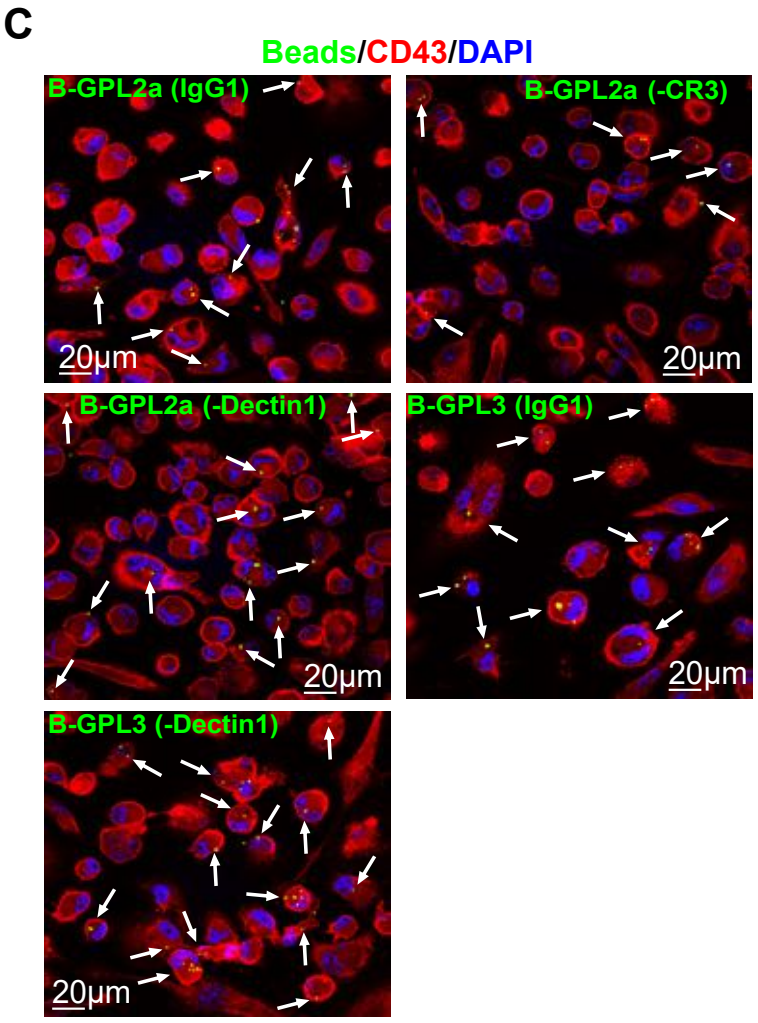
Supplemental Figure 6



Supplemental Figure 7







Video 1

Click here to access/download

Supplemental Videos and Spreadsheets

Movie1.avi

Video 2

Click here to access/download

Supplemental Videos and Spreadsheets

Movie2.avi

III.E.5.c Lien entre l'acétylation et la glycosylation des GPLs de Mabs S chez les mutants Δatfs (version soumise du 3ème article)

C'est travaux ont donné lieu à la rédaction d'un article dont je suis co-premier auteur et qui est actuellement en soumission. Les résultats seront présentés dans la version de l'article envoyée à l'éditeur et commentés en faisant référence aux figures originales de l'article.

L'étude structurale des GPLs s'est poursuivie avec la recherche de formes différemment acétylées parmi trois mutants d'acétyltranférases puisque le gène *atf* présent chez *Msmeg* est dupliqué chez *Mabs* en *atf1* et *atf2*. Notre objectif était de vérifier si la duplication du gène était associée à l'ajout successif par chacun des produits des deux gènes des groupements acétyles sur la position 3 puis 4 du 6 déoxy talose, introduisant ainsi une spécificité dans l'activité des enzymes.

Nous nous sommes d'abord rendu compte que la délétion des gènes *atf1* et *atf2* n'empêchait pas l'acétylation (Fig. 2), et qu'une compensation était assurée par 2 acétyltransférases codées par les gènes *MAB_1725c* et *MAB_3448*, hors du locus de biosynthèse des GPLs (Fig. 3). Alors que le mutant triple Δ1,2,1725c était capable de produire un GPL-3 mono-acétylé indifféremment sur les positions 3 ou 4, la délétion des quatre gènes était nécessaire pour n'observer que des GPLs complétement déacétylés (Fig. 4). De plus, deux effets métaboliques collatéraux et imprévisibles ont été constatés. Tout d'abord, la délétion d'*atf2* a entrainé une diminution de la proportion de GPL-2a comparativement au GPL-3 qui a été quantifiée dans toutes les souches mutantes. Ensuite, l'accumulation de GPL-3 dans ces dernières a provoqué l'apparition de GPLs *O*-méthylées sur le rhamnose terminal (Fig. 5D). Comparativement à *Msmeg*, il apparait donc que l'acétylation des GPLs est plus finement régulée chez *Mabs* aboutissant à une plus grande variété de formes nouvelles. Enfin, nous avons confirmé que la présence des groupements acétyles sur le 6dTal n'influençait pas la phagocytose par les macrophages des souches mutantes ainsi que des billes fluorescentes recouvertes de GPLs, mais modifiait progressivement la morphologie des colonies vers un aspect plus rugueux (Fig. 6).

III.E.5.d Conclusion

Les travaux menés sur le morphotype lisse sauvage de *Mabs* ont établi la fonction biochimique des gènes appartenant au locus de biosynthèse des GPLs. Dans le cas du gène *fmt*, ce dernier est impliqué dans l'*O*-méthylation de la fontion β-hydroxy du lipide. Les gènes *gtf1*, *gtf2* et *gtf3* ajoutent le 6 déoxy-talose, ainsi que les deux rhamnoses. Contrairement à *Msmeg*, les étapes de glycosylation sont plus finement régulées. L'initiation nécessite l'action de *gtf2* et l'ajout du premier rhamnose pour

que le gène *gtf1* transfère immédiatement le 6 déoxy-talose. Par contre, le produit du gène *gtf3* n'est pas complétement actif pour l'ajout du second rhamnose puisque le GPL-2a coexiste avec le GPL-3. En outre, la duplication des gènes *atf1* et *atf2* se répercute partiellement sur l'acétylation. En effet, les gènes MAB_1725c et MAB_3448 sont aussi impliqués dans l'acétylation. Ces quatre gènes présentent des fonctions redondantes pour l'acétylation des GPLs.

S'agissant du morphotype, la délétion seule de *fmt*, *gtf3* ou de l'un des 4 gènes codant les acétyltransférases n'a pas modifié le morphotype alors que la délétion de *gtf1*, *gtf2* ou des 4 gènes codant les acétyltransférases entraine une transition vers le morphotype rugueux. Nos travaux ont également mis en évidence le rôle biologique du gène *fmt* dans l'hydrophobicité de la paroi et les forces d'adhésion du bacille, la contribution prépondérante du GPL-3 dans la phagocytose in-vitro par les macrophages THP-1 des bacilles ainsi que le rôle du gène *atf2* dans l'homéostasie GPL-2a/GPL-3.

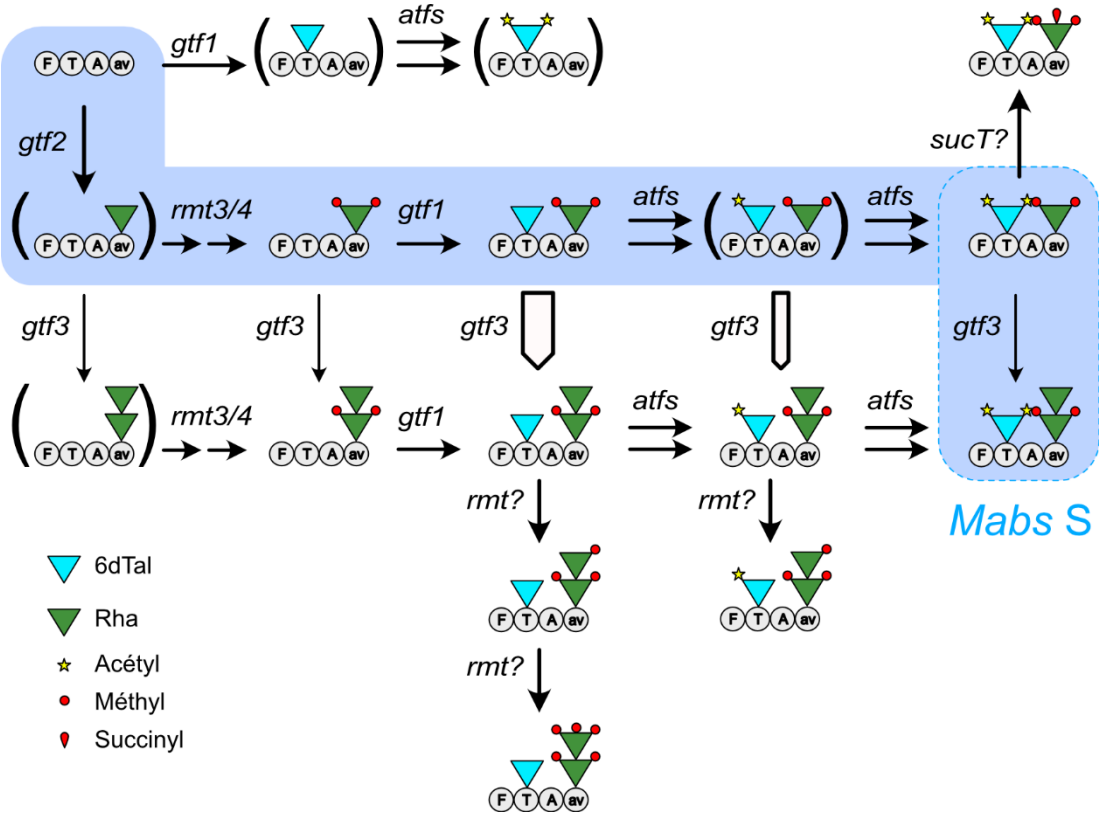


Figure 55 : Schéma de glycosylation des GPLs de *Mabs* S. D'après le contenu en GPLs des mutants étudiés de glycosyltransférases et d'acétyltransférases, la voie de biosynthèse indiquée en bleu aboutit à l'accumulation de GPL-2a et GPL-3 chez la forme sauvage. Les GPLs entre parenhèses n'ont pas été identifiés, supposant des impasses métaboliques (action de Gtf1 sur le lipopeptide) ou des cinétiques enzymatiques plus ou moins rapides (acétylation par les 4 Atfs).

Manuscript Click here to view linked References

Multiple *Mycobacterium abscessus O*-acetyltransferases influence glycopeptidolipid structure

Running title: Acetylation of GPL in *Mycobacterium abscessus*

Morgane Illouz[1,*], Louis-David Leclercq[2,*], Clara Dessenne[2], Graham Hatfull[3], Wassim Daher[1,4], Laurent Kremer[1,4,#], and Yann Guérardel[2,5#]

[1]Centre National de la Recherche Scientifique UMR 9004, Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France.
[2]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.
[3]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, US
[4]INSERM, IRIM, 34293 Montpellier, France.
[5]Institute for Glyco-core Research (iGCORE), Gifu University, Gifu, Japan

*These authors contributed equally to this work.

[#]Corresponding authors: yann.guerardel@univ-lille.fr, Tel: (+33) 3 20 43 69 41; laurent.kremer@irim.cnrs.fr, Tel: (+33) 4 34 35 94 47.

Keywords: *Mycobacterium abscessus*, glycopeptidolipid, cell wall, acetyltransferase, macrophage, mass spectrometry.

ABSTRACT

Mycobacterium abscessus is causing severe lung infections and clinical isolates can have either smooth (S) or rough (R) colony morphotypes. S, but not R variants have abundant cell wall glycopeptidolipids (GPL), consisting of a peptidolipid core substituted by a 6-deoxy-talose (6-dTal) and one or two rhamnose residues. Deletion of *gtf1*, encoding the 6-dTal transferase results in the S-to-R transition, mycobacterial cord formation, and increased virulence, underscoring the importance of 6-dTal in infection outcomes. However, since 6-dTal is di-*O*-acetylated, it is unclear whether the *gtf1* mutant phenotypes are related to the loss of the 6-dTal or by the absence of acetylation. Here, we addressed whether *M. abscessus atf1* and *atf2*, encoding two putative *O*-acetyltransferases located within the *gpl* biosynthetic locus, transfer acetyl groups to 6-dTal. Deletion of *atf1* and/or *atf2* did not drastically alter the GPL acetylation profile, suggesting the occurrence of additional enzymes with redundant functions. Two paralogs of *atf1* and *atf2, MAB_1725c* and *MAB_3448,* were identified. While deletion of these two genes had no effect on GPL acetylation, the triple *atf1-atf2-MAB_1725c* mutant did not synthetize fully-acetylated GPL. However, the quadruple mutant *atf1-atf2-MAB_1725c-MAB_3448* was totally devoid of acetylated GPL. Moreover, both triple and quadruple mutants accumulated hyper-methylated GPL. Deletion of *atf* genes resulted in subtle changes in colony morphology but had no effect on *M. abscessus* internalization by macrophages. Overall, these findings reveal the existence of functionally redundant *O*-acetyltransferases and suggest that *O*-acetylation influences the glycan moiety of GPL by deflecting biosynthetic flux in *M. abscessus*.

INTRODUCTION

Infections caused by non-tuberculous mycobacteria (NTM), and particularly *Mycobacterium abscessus*, are globally on the rise and remain difficult to treat due to their intrinsic resistance levels to most antibiotic classes (1). *M. abscessus* is a fast-growing bacteria, increasingly acknowledged as an emerging human pathogen, responsible for skin, soft tissue (2) and lung infections, mostly in patients with underlying lung disorders, such as cystic fibrosis (CF) patients (3, 4). *M. abscessus* displays either smooth (S) or rough (R) colonial morphotypes, to which distinct *in vitro* and *in vivo* phenotypes can be assigned. S variants are typified by the production of surface-associated glycopeptidolipids (GPL), which are lacking or produced at low levels in the R variants (5–7).

GPL are complex lipids **(Figure 1A)** comprising a D-Phe-D-*allo*Thr-D-Ala-L-alaninol peptide core assembled by the action of two non-ribosomal peptide synthetases, Mps1 and Mps2, and acylated with a 3-hydroxy/methoxy C_{24}-C_{33} fatty acid (8–10). It was recently shown that alaninol can be replaced by the branched amino-alcohol valinol or leucinol (11). The resulting lipopeptide is glycosylated with 6-deoxy-α-L-talose (6-dTal) on the *allo*-Thr residue by the action of the glycosyltransferase Gtf1, while the alaninol is substituted by a α-L-rhamnose (Rha) by Gtf2, resulting in the production of the less-polar diglycosylated GPL molecules, designated GPL-2a. In addition to GPL-2 that contain a 3,4-di-*O*-acetylated 6-dTal and a 3,4-di-*O*-methylated or 2,3,4-tri-*O*-methylated Rha (8, 12–14), *M. abscessus* also produces a more polar GPL, designated GPL-3, through the addition of a 2,3,4-tri-hydroxylated Rha to the alaninol-linked 3,4-di-*O*-methyl Rha, a reaction catalyzed by glycosyltransferase Gtf3 (15). GPL-3 is structurally identical in *M. abscessus* and *M. smegmatis*, but is more abundant in *M. abscessus* (8). Once synthesized, GPL-2a and GPL-3 are translocated across the inner membrane by the MmpL4a/MmpL4b transporters (16–19) and inserted into the outer leaflet of the mycomembrane where they are exposed to the surface of the bacilli (20).

Numerous studies showed that the presence or loss of GPL influences susceptibility to antibiotics, sliding motility, biofilm formation (5, 7, 16, 21–23), bacterial surface hydrophobicity (20, 24), cord

formation (5, 16, 25, 26), interaction with host macrophages (13, 27) and induction of a pro-inflammatory responses (28). The role of S/R morphotypes to influence disease outcomes is supported by cellular and animal models, indicating increased pathogenesis of R forms relative to S forms (16, 26, 29, 30). In particular, in the zebrafish model of infection the S-to-R transition is associated with increased bacterial loads, production of mycobacterial cords, abscess formation and enhanced larval mortality (29). Understanding the molecular mechanisms supporting the S-to-R morphotypic switch is of foremost importance and clinically relevant as evidenced by epidemiological surveys highlighting the predominance of R strains in patients with severe lung diseases and chronic colonization of the airways in CF patients (31, 32). Genomic and transcriptomic analyses originally revealed the presence of various insertions or deletions in the R strains, mostly found in *mps1*, *mps2* and *mmpL4b* genes participating either in the biosynthesis or transport of GPL and responsible for the S-to-R transition in *M. abscessus* (33–35). A recent in-depth dissection study of the GPL glycome provided insights into the biological function of the different monosaccharides and demonstrated that deletion of either *gtf1* or *gtf2* (but not *gtf3*) resulted also in the S-to-R transition with enhanced cording and increased virulence in zebrafish embryos (15). These findings emphasize the importance of 6-dTal in infection outcomes and in the interaction with host macrophages. However, because 6-dTal is di-*O*-acetylated, it is not known whether the phenotypes associated with the *gtf1* mutant are a direct consequence of the absence of the 6-dTal or whether this may be due to loss of the 6-dTal-substituted acetyl groups. In addition, whether the presence/absence of 6-dTal acetylation influences the GPL structure is not well understood. Thus, identification of the *O*-acetyl transferases (Atf) and the contribution of *O*-acetylation to the overall GPL structure and biological function in *M. abscessus* remain to be established.

The aim of this study was to further advance our understanding of the genetic requirements for the biosynthesis and biological functions of the high-GPL producing variants by focusing on the Atf-related enzymes involved in GPL *O*-acetylation. Bioinformatic analyses together with genetic studies were combined to identify and generate single and multiple deletion mutants lacking

between one and four *atf* genes and further characterized by thorough biochemical and structural GPL analyses. These mutants were also exploited to address the contribution of *O*-acetylation during the synthesis of GPL, their role in bacterial morphology and internalization of *M. abscessus* in human macrophages.

RESULTS

Generation of *atf1* and *atf2* acetyltransferases deleted strains

While previous studies aimed at deciphering the contribution of the different moieties composing GPL (*O*-methylation of the lipid core (24) or glycosylation of the peptide backbone (15)) to the biological functions of this glycolipid, the importance of GPL acetylation on C3 and C4 positions of the 6-dTal of GPL-2a and GPL-3 **(Figure 1A)** remains unsolved. The *M. abscessus gpl* biosynthetic locus encompasses two genes coding for putative acetyltransferases, referred to as *atf1 (MAB_4106c)* and *atf2 (MAB_4110c)* (8), contrasting with *M. smegmatis* which possesses a single *atf* gene *(MSMEG_0390)* (36) **(Figure 1B)**; all of these are predicted to be membrane localized and contain ten putative transmembrane domains. To investigate the contribution and biological functions of Atf1 and Atf2 in GPL synthesis, *atf1* and *atf2* were deleted either individually or simultaneously in a high-GPL-producing *M. abscessus* S variant (CIP104536^T; identical to ATCC19977), using an unmarked deletion system that allows multiple gene deletions in *M. abscessus* (37). In brief, the strategy involves double homologous recombination leading to the deletion of the *atf1* and/or *atf2* open reading frames from the *gpl* locus, based on the pUX1-*katG* suicide vector **(Table S1)** carrying a kanamycin resistance (*kan*^{R}) cassette, a tdTomato red fluorescence marker, and a *katG* cassette that confers susceptibility to isoniazid (INH) in *M. abscessus*. The DNA segments flanking *atf1* or *atf2* were cloned into pUX1-*katG,* to give pUX1-*katG-atf1* and pUX1-*katG-atf2* **(Figure S1A)**. Transformants that integrated the plasmids into their specific loci *via* homologous recombination were selected for resistance to kanamycin and had visible red fluorescence. After a second homologous recombination event, loss of the plasmid resulted in recovery of *Δatf1* (designated Δ1) and *Δatf2* (designated Δ2) non-fluorescent colonies

that were INH-resistant and Kan-susceptible. PCR/sequencing of the progenitor and the deletion strains using primers listed in **Table S2**, confirmed the genotype of these *atf* mutants **(Figure S1B)**. Functional complementation of Δ1 and Δ2 was done through specific integration at the *attB* chromosomal site (38) of an intact copy of *atf1* and *atf2* with an HA-tag placed at the 3'-end and under the control of the *hsp60* promoter **(Figure S2A)**. Western blot analyses of the crude lysates using anti-HA antibodies was carried out by loading equal amounts of proteins (confirmed by the KasA protein internal control) **(Figure S2B)**. This revealed the presence of single bands, presumably corresponding to Atf1-HA and Atf2-HA, although these migrate somewhat faster their predicted sizes. These observations are consistent with complementation of Δ1 and Δ2.

Deletion of *atf1* and/or *atf2* does not suppress GPL acetylation

To qualitatively analyze the impact of *atf1* and/or *atf2* deletion, we analyzed the GPL from WT, Δ1, Δ2, and Δ1,2 strains. GPL were prepared by sequential extraction of apolar and polar lipid fractions and analyzed by thin layer chromatography (TLC). Apolar fractions all exhibited similar profiles **(Figure S3A)** and were not further studied. In contrast, the GPL profiles in polar fractions of Δ1, Δ2, and Δ1,2 differed from the parental S strain **(Figure 2A)**, whereas other glycolipids are expressed at similar levels **(Figure S3B).** In particular, the band intensity of GPL-2a significantly decreased in Δ1, Δ2 and Δ1,2 compared to WT. TLC analysis also showed an increased intensity for a set of bands with slower mobility, potentially more polar than GPL-3 in all three mutants; this was more pronounced in Δ1,2 **(Figure 2A, grey arrow)**.

The MALDI-MS spectrum of polar fraction from the WT strain reveals the presence of a complex set of GPL dominated by multiple isomers of GPL-3 (major signals at *m/z* 1404 and 1432) and GPL-2a (major signals at *m/z* 1258 and 1286) **(Figure 2B),** as reported previously (13, 15, 24, 39). The presence of di-*O*-acetylated GPL-3 was confirmed in all four samples by MS/MS analysis of the parent *m/z* 1404 ion that produced the diagnostic fragment ion M-248 at *m/z* 1156, typifying the loss of 3,4 di-*O*-acetylated 6-dTal **(Figure 2C)**. Detailed analysis of the fragmentation spectra showed several fragment ions with 28 a.m.u. differences that originate from either the variation of

the length of the lipid chain (C_{26} or C_{28}) or the variation of the peptide sequence (alaninol or valinol), generating intrinsic heterogeneity of GPL-3, designated Ala-C_{28} or Val-C_{26} **(Figure 2C)**. Noticeably, the internal Rha residue linked to the amino alcohol is 3,4 di-*O*-methylated. MS/MS analysis of the signal at *m/z* 1404 in all four strains resulted in the same fragmentation spectrum, demonstrating that Δ1 and Δ2 synthesize similar di-*O*-acetylated GPL-3 **(Figure S4A)**. Comparable observations were made with the isomeric structure of GPL-3 at *m/z* 1432 **(Figure S4B)**. As for GPL-3, GPL-2a was observed in all four strains with similar structures. However, in accordance with the TLC analysis, the intensity of GPL-2a MS signals at *m/z* 1258 and 1286 were strongly decreased in Δ2 and Δ1,2 compared to GPL-3 signals at *m/z* 1404 and 1432 **(Figure 2B)** but not in Δ1 **(Figure S5)**. Importantly, complementation of Δ1 and Δ2 restored the parental GPL profile **(Figure S3C)**. At this stage of the study, we could not identify the accumulating product in Δ1, Δ2 and Δ1,2 **(Figure 2A, grey arrow)** because of spectral overlap, but it was later identified as monoacetylated GPL-3 (see below).

Together, these observations show that the single and double disruptions of *atf1* and *atf2* induced moderate under-acetylation of GPL-3 associated with an unexpected reduction of GPL-2a amounts compared to GPL-3, which was more pronounced in the absence of *atf2*.

Bioinformatic analyses identify two *atf* paralogues in *M. abscessus*.

The fact that the acetylation profile was incompletely modified in Δ1,2 raised the possibility of additional enzymes with redundant functions, encoded by genes outside of the *gpl* locus. BLAST analyses using either Atf1 or Atf2 as queries identified two putative acetyltransferase candidates, encoded by *MAB_1725c* and *MAB_3448*, respectively. Interestingly, *MAB_1725c* is part of a previously identified prophage (designated propphiATCC19977-1) in *M. abscessus* ATCC19977 **(Figure 3A),** and can spontaneously induce to give viable phage particles (40, 41). The prophage is integrated into *attB-5* overlapping a host $tRNA^{Met}$ gene and has 113 predicted coding sequences (35, 40, 41). In contrast, *MAB_3448* belongs to a cluster of genes coding for proteins with unknown function or expressing lipolytic activity, such as MAB_3447c or MAB_3452c **(Figure**

3A). Multiple amino acid alignments show the high sequence conservation between all four proteins **(Figure 3B)**. MAB_1725c shares 65% and 53% identity with Atf1 and Atf2 while MAB_3448 shares 66% and 57% identity with Atf1 and Atf2, respectively. MAB_1725c and MAB_3448 share 79% identity **(Figure 3C)**. This high sequence conservation suggests these enzymes may be functionally redundant. Like Atf1 and Atf2, MAB_1725c and MAB_3448 are predicted to be membrane localized.

Loss of *atf1*, *atf2*, *MAB_1725c* and *MAB_3448* severely impact the GPL composition.

To investigate the potential contribution of *MAB_1725c* and *MAB_3448* to GPL acetylation, these genes were individually or simultaneously deleted in *M. abscessus* using the genetic approach employed to generate Δ1 and Δ2, leading to Δ1725c, Δ3448 and Δ1725c,3448 **(Figure S1C)**. The deletions were confirmed by PCR/sequencing **(Figure S1D)** and the mutants subjected to lipid analyses. TLC **(Figure 2A)** as well as MALDI-MS profiling **(Figure S5)** of Δ1725c, Δ3448 and Δ1725c,3448 closely mirror the WT patterns, demonstrating that *MAB_1725c* and *MAB_3448* are not essential for acetylation of GPL in a WT strain. Next, we tested the possibility that these enzymes express redundant activities with Atf1 and/or Atf2 by comparing the GPL profiles of Δ1, Δ2, Δ1,2 and Δ1,2,1725c and Δ1,2,3448 as well as in the quadruple mutant (Δ1,2,1725c,3448) lacking all four genes. Δ1,2,3448 showed by TLC and MALDI-MS analyses a GPL profile comparable to Δ1,2 **(Figure 2A and S5)**. Conversely, Δ1,2,1725c exhibited a different profile characterized by a strong reduction of GPL-2a and GPL-3 but with the concomitant increase in intensity of the lower Rf bands **(Figure 2A)**. Furthermore, two intense sets of bands appeared just below the GPL-3 **(Figure 2A, white and grey arrows)** and one between GPL-3 and GPL-2a **(Figure 2A, black arrow)**, suggesting that biosynthetic intermediates accumulate concomitantly to the disappearance of the final products. A similar trend was observed in Δ1,2,1725c,3448, culminating with the total disappearance of GPL-2a and GPL-3 bands **(Figure 2A)**. The polar glycolipid content was also assessed by TLC using other developing solvents, but no significant changes outside GPL were observed **(Figure S3D and S3E)**. Consistently, MALDI-MS spectra

of Δ1,2,1725c and Δ1,2,1725c,3448 displayed very different profiles to the WT strain, characterized by the strong reduction of GPL-3 signals at *m/z* 1404 and 1432 in the former and their complete disappearance in the latter **(Figure 2B)**. The MALDI-MS spectrum of Δ1,2,1725c showed a new cluster of signals dominated by a pair of ions at *m/z* 1362 and 1390, tentatively identified as mono-*O*-acetylated GPL-3 based on a 42 a.m.u reduction compared to GPL-3 signals at *m/z* 1404 and 1432. A cluster of ions from *m/z* 1320 to 1362 was identified in Δ1,2,1725c,3448 **(Figure 2B)**, requiring further investigation.

M. abscessus lacking *atf1*, *atf2*, *MAB_1725c* and *MAB_3448* produce non-acetylated and hyper-methylated GPL.

To explore the complex nature of the structural modifications observed in the different mutant strains, we first focused on the quadruple mutant. Indeed, deletion of all four *atf* genes induced substantial biochemical alterations, resulting in the simplest GPL profile of all studied mutants, characterized by the accumulation of three lipids, as observed on the TLC **(Figure 2A, black, grey and white arrows)**. The structures of these three components were individually determined in order to provide the basis for analyzing the more complex triple mutant. First, the polar lipid fraction isolated from the quadruple mutant was separated by preparative TLC into three fractions, referred to as Δ1,2,1725c,3448-F1, Δ1,2,1725c,3448-F2 and Δ1,2,1725c,3448-F3 **(Figure 4A)**. Δ1,2,1725c,3448-F1, which co-elutes with di-acyl trehalose (DAT), showed two major ions at *m/z* 1320 and 1348 on the MALDI-MS spectrum **(Figure 4B)**, tentatively attributed to non-acetylated GPL-3 based on the 84 a.m.u decrease from ions *m/z* 1404 and 1432. MS^2 fragmentation confirmed that the GPL signals at *m/z* 1320 was made of two GPL-3 isomers Ala-C_{28} and Val-C_{26}, substituted by non-acetylated 6-dTal and Rha residues, owing to the observation of signals M-6dTal at *m/z* 1156 and M-Rha at *m/z* 1174, respectively **(Figure S6A)**. The GPL signal at *m/z* 1348 was identified as a GPL-3 isomer Val-C_{28} substituted by identical monosaccharides **(Figure S6B)**. Δ1,2,1725c,3448-F2 comprises two bands with intermediate Rf between Δ1,2,1725c,3448-F1 and GPL-3 **(Figure 4A)**. MALDI-MS analysis showed a group of three intense ions dominated

by the ion at *m/z* 1334 **(Figure 4B)**. MS^2 fragmentation generated a signal at *m/z* 1170 indicative of the loss of non-acetylated 6-dTal as well as a M-160 ion at *m/z* 1174, indicative of the loss of terminal mono-*O*-methyl Rha from non-acetylated GPL-3 isomers Ala-C_{28} and Val-C_{26} **(Figure S7A).** Parent ion at *m/z* 1362 displayed a similar MS^2 pattern, highlighting the loss of non-acetylated 6-dTal and mono-*O*-methyl Rha, which typifies a GPL-3 Val-C_{28} form **(Figure S7B)**. It should be noted that tri-*O*-methylated GPL-3 with identical signal at *m/z* 1334 was previously reported in *M. smegmatis* (42). Finally, Δ1,2,1725c,3448-F3 was characterized as a single high mobility band on TLC, associated to a set of major ions at *m/z* 1334, 1362 and 1390 on MALDI-MS **(Figure 4A and 4B)**. The 28 a.m.u increase in the parent ion over Δ1,2,1725c,3448-F2, coupled with the Y2-fragment ion generated by MS/MS at *m/z* 531, suggested the presence of two additional *O*-methyl groups. These were located on the terminal Rha owing to the MS^3 fragmentation of the Y2 ion at *m/z* 531 that generated a secondary fragment M-188 ion at m/z 343, confirming the presence of a terminal tri-*O*-methyl Rha on GPL-3 Ala-C_{28} **(Figure S8A)**. Similar MS^2 fragmentation patterns were observed for parent ion at *m/z* 1390, which typifies GPL-3 Ala-C_{30} and Val-C_{28} isobars with terminal tri-*O*-methyl Rha **(Figure S8B)**. The presence of tri-*O*-methyl Rha was then confirmed by GC/MS analysis as demonstrated below.

Together, these data demonstrate that quadruple mutant Δ1,2,1725c,3448 completely lost the ability to synthesize acetylated GPL but accumulated a set of non-acetylated GPL-3 isomers substituted by two, three or five methoxy groups **(Figure 4B)**.

***M. abscessus* lacking *atf1*, *atf2* and *MAB_1725c* synthesizes mono-*O*-acetylated GPL-3.**

Next, we examined in detail the GPL pattern of the triple Δ1,2,1725c mutant that showed a more complex pattern than the quadruple mutant Δ1,2,1725c,3448, presumably because of the presence of partially acetylated products. The polar lipids from Δ1,2,1725c were separated by preparative TLC into four fractions, designated Δ1,2,1725c-F1 to Δ1,2,1725c-F4 **(Figure 4C)**, and subjected to MALDI-MS and MS/MS analysis **(Figure 4D)**. Fraction Δ1,2,1725c-F1 was identified as de-*O*-acetylated GPL-3 with a terminal Rha residue based on identical Rf, MALDI-MS and

MALDI-MS^2 fragmentation patterns than Δ1,2,1725c,3448-F1 **(Figures 4C, 4D, S9A and S9B)**. Similarly, based on TLC mobility, MS, MS^2 and MS^3 analyses **(Figure 4C, 4D and S8C)**, Δ1,2,1725c-F4 was identified as non-acetylated GPL-3 Ala-C_{28} with a terminal tri-*O*-methyl Rha, as observed in Δ1,2,1725c,3448-F3. In contrast, Δ1,2,1725c-F2 and Δ1,2,1725c-F3 were shown to contain differently methylated mono-*O*-acetylated GPL-3 species. In particular, Δ1,2,1725c-F2 was characterized by a set of three ions dominated by a signal at *m/z* 1362 **(Figure 4D)**. MS^2 analysis of parent ion at *m/z* 1362 showed the presence of a complex mixture consisting of two major GPL-3 isomers with valinol- or alaninol-containing peptides identified as mono-*O*-acetylated GPL-3 Val-C_{26} and GPL-3 Ala-C_{28} substituted by terminal Rha residues **(Figure S10A)**. Alaninol- and valinol-containing peptides were differentiated by simultaneous detection of fragment ions at *m/z* 489 and 517 and those resulting from the cleavage of terminal glycosidic linkages at *m/z* 1156 and 1216. The second major ion in Δ1,2,1725c-F2 observed at *m/z* 1390 was identified as mono-*O*-acetylated GPL-3 Val-C_{28} according to primary fragment M-206 at *m/z* 1184 while the fragment M-146 at *m/z* 1244 confirmed the presence of a terminal Rha residue **(Figure S10B)**. Fraction Δ1,2,1725c-F3, that is exclusively observed in Δ1,2,1725c showed three major ions at *m/z* 1348, 1376 and 1404 by MALDI-MS **(Figure 4D)** that were further subjected to MS^2 and MS^3 experiments. Of interest, the MS^2 spectrum of parent ion at *m/z* 1376 yielded signals at *m/z* 1170 and 1216 generated by the cleavage of terminal mono-*O*-acetylated 6-dTal and mono-*O*-methyl Rha residues **(Figure S11A)**, as confirmed by MS^3 fragmentation of ion at *m/z* 1170. MS^3 Y2-fragment ions at *m/z* 503 or 531 and secondary fragments produced by Y3-fragments at *m/z* 792 or 820 typified two GPL-3 isomers Val-C_{26} and Ala-C_{28}, both characterized by the presence of mono-*O*-acetylated 6-dTal and terminal mono-*O*-methyl Rha. The 28 a.m.u higher mass ion at *m/z* 1404 was similarly identified as GPL-3 Val-C_{28} based on MS^2 analysis **(Figure S11B)**. The exclusive presence of *O*-acetyl groups in fractions Δ1,2,1725c-F2 and Δ1,2,1725c-F3 was further confirmed by ^{1}H NMR analyses that showed a clear signal attributable to C$\underline{H_3}$-CO- group at δ2.14 ppm **(Figure S12)**.

Overall, detailed GPL mapping in the multiple *atf*-deleted strains established that the loss of at least three acetyl transferases present inside and outside the *gpl* locus is required to significantly reduce GPL acetylation while disruption of all four *atf*-related genes is needed to abrogate GPL acetylation in *M. abscessus*.

Deletion of all four *atf* genes disturbs the synthesis of GPL glycan moiety.

To address whether acetylation regulates the synthesis of GPL glycan moiety, the relative quantity of each deoxyhexose was determined in the various *atf* mutants. To do so, GPL were first chemically de-acetylated to afford a complete purification from acyl ester glycolipids such as trehalose-based compounds, in particular DAT. Then, the monosaccharide composition of the resulting purified deacetylated GPL (dGPL) was established by GC/MS **(Figure 5A).** This revealed that the overall proportion of Rha residues slightly increased in all mutants lacking *atf2* (below 60% in WT and up to 65% in Δ1,2,1725c,3448) but remained unchanged in all other mutants as well as in the Δ2-complemented strain. In agreement with above-mentioned data, GC/MS analysis revealed the presence of about 3% 2,3,4 tri-*O*-methyl Rha in Δ1,2,1725c and Δ1,2,1725c,3448, but not in the other strains.

As saponification released the acetyl groups, the migration of the resulting dGPL on TLC exclusively relies on the rhamnan moiety **(Figure 5B)** which considerably simplifies the migration pattern as compared to their corresponding native GPL **(Figure 2A)**. Standard dGPL-3 and dGPL-2a were obtained by saponification of GPL-3 and GPL-2a **(Figure S13A)** while standard deacetylated tri- and penta-*O*-methylated-GPL-3 were obtained from Δ1,2,1725c,3448-F2 and F3 **(Figure S13B)**. Based on the mobility of those standard molecules, the TLC analysis of deacetylated GPL fractions confirmed that the relative quantity of dGPL-2a was strongly decreased in all *atf2*-deficient mutants, with a complete loss of dGPL-2a in the quadruple mutant **(Figure 5B)**. Simultaneously, di-, tri- and penta-*O*-methylated dGPL-3 strongly accumulated. The impact of *atf* deletion on reduced dGPL-2a production was quantified by computing the integration values of MALDI-MS diagnosis ion signals at *m/z* 1320 and 1348 for dGPL-3 and 1202 and 1174

for dGPL-2a **(Figure S13C)**. Determination of the dGPL-2a/dGPL-3 ratio clearly indicated that the deletion of *atf2* was sufficient to reduce, by at least 10-fold, the relative quantity of GPL-2a with respect to GPL-3 **(Figure 5C)**. In comparison, the deletion of *atf1*, *MAB_1725c* or *MAB_3448* failed to or only marginally reduced the GPL-2a content.

GPL acetylation influences colonial morphology but not uptake of *M. abscessus* by macrophages.

It is well established that the absence or the presence of GPL determines the morphotype of *M. abscessus* colonies on agar medium (5, 16, 33). In addition, that a *gtf1* mutant lacking 6-dTal acquires a rough morphotype (15), prompted us to compare the colonial morphology of the single, double, triple and quadruple *atf* mutants. Observation of individual colonies on Tryptic Soy agar failed to reveal morphological differences in Δ1, Δ1725c, Δ3448 and Δ1725c,3448 and all appeared similar to the parental S strain **(Figure 6A)**; none of these have altered acetylation patterns. Unexpectedly, Δ2 and Δ1,2 showed a different morphology than the parental S strain, but distinct from the typical rough and dry features of the R strain. Strains Δ1,2,1725c, Δ1,2,3448 and Δ1,2,1725c,3448 all exhibited morphological differences, possibly as a consequence of the increased GPL-2a/GPL-3 ratio **(Figure 6A)**. In contrast to the Δ*gtf1* mutant lacking 6-dTal with a pronounced corded phenotype (15), none of the acetylation mutants displayed the typical serpentine cords of the R variant (data not shown), suggesting that the corded structure of Δ*gtf1* is primarily caused by the loss of the 6-dTal rather than by the absence of the acetyl substituents. Similar to the *M. abscessus* R strain, only mutants lacking *atf2* rapidly sedimented in liquid culture, again underscoring a possible role of the reduced GPL-2a/GPL-3 ratio in bacterial aggregation **(Figure 6B)**. The *in vitro* growth curves of the *atf* mutants in 7H9 broth at 37 °C were comparable to the one of the parental S strain, reaching a plateau after 2 days **(Figure 6C)**. This suggests that the single or multiple deletions of the *atf* genes do not influence the replication rate of *M. abscessus* in planktonic cultures.

We next infected human THP-1 macrophages for 3 hrs with *M. abscessus* S, Δ1,2, Δ1,2,1725c, Δ1,2,3448 and Δ1,2,1725c,3448 prior assessing phagocytosis by CFU counting. As shown in **Figure 6D**, all strains were equally taken up by macrophages, suggesting that GPL acetylation is not required for internalization of *M. abscessus* by macrophages. To confirm these observations, we determined the contribution of GPL acetylation in phagocytosis using acetylated *versus* de-acetylated GPL-2a and acetylated *versus* de-acetylated GPL-3-coated fluorescent beads. Quantification of the proportion of bead-containing cells indicated that macrophages internalized significantly more GPL-2a-coated beads than uncoated beads **(Figure 6E)**. As reported previously, THP-1 cells phagocytosed more GPL-3-coated beads than GPL-2a-coated beads, thus highlighting the contribution of the terminal Rha in this process (15). However, no significant differences between dGPL-2a *versus* GPL-2a or between dGPL-3 v*ersus* GPL-3 were noticed **(Figure 6E)**. High-resolution confocal imaging confirmed the intracellular localization of the different GPL-counted beads **(Figure 6F)**.

Collectively, the structure-function comparison using either whole bacilli or purified GPL-beads indicates that acetylation of 6-dTal does not participate in the uptake of *M. abscessus* by macrophages.

DISCUSSION

Acetylation of GPL has been proposed to regulate various biological processes such as sliding motility and biofilm formation in *M. smegmatis* (36) or macrophage apoptosis in *M. abscessus* through the interaction of di-*O*-acetylated GPL with mitochondrial cyclophilin D (13). While *M. smegmatis* produces essentially di-glycosylated GPL-2, *M. abscessus* synthesizes equivalent amounts of GPL-2a and GPL-3, which can be easily distinguished by TLC and mass spectrometry. Previous work demonstrated that disruption of the unique *atf* gene in *M. smegmatis* abolished di-*O*-acetylated GPL-2 synthesis (8, 36). However, in the present work, our attempts to quantitatively inhibit GPL acetylation in *M. abscessus* by simultaneously deleting *atf1* and *atf2*, located within the *gpl* biosynthetic locus, were unsuccessful despite the fact that Atf1 and Atf2 sequentially transfer acetyl groups to 6-dTal when overexpressed in the *M. smegmatis atf* mutant (8). Indeed, single and simultaneous disruptions of *atf1* and *atf2* genes only resulted in a slight decrease of GPL acetylation. These observations challenge the initial hypothesis that *atf1* and *atf2* duplication results in the sequential addition of acetyl groups on positions 3 and 4 on the 6-dTal in *M. abscessus*. Unexpectedly, disruption of *atf2* induced a notable decrease of di-*O*-acetylated GPL-2a correlating with an accumulation of mono- and di-*O*-acetylated GPL-3. Since all mutants lacking *atf2* sedimented efficiently in liquid culture, similarly to *M. abscessus* R, it may be proposed that this aggregation property is linked to the increased proportion of GPL-3 over GPL-2a. This hypothesis is strengthened by the fact that *M. abscessus* S overproducing the rhamnosyl-transferase Gtf3, also associated with an increased GPL-3/GPL-2a ratio, aggregated rapidly in liquid culture, like the original R strain (15).

We posit here the existence of additional factors, encoded outside the *gpl* cluster, capable of functionally rescuing the lack of acetylation in the Δ1,2 mutant. To advance our understanding of the GPL acetylation genetic requirements, two additional Atf-related enzymes potentially involved in GPL *O*-acetylation were identified, *MAB_1725c* and *MAB_3448*. Whereas individual or simultaneous deletion of these genes did not significantly modify the GPL profile, multiple deletions of *atf1*, *atf2* and *MAB_1725c* resulted in a high accumulation of mono-*O*-acetylated

GPL-3. Further deletion of *MAB_3448* in Δ1,2,1725c,3448 induced complete loss of GPL acetylation, as assessed by the accumulation of de-*O*-acetylated GPL-3. In addition to the loss of 6-dTal acetylation in Δ1,2,1725c and Δ1,2,1725c,3448 strains, additional methylation of the terminal Rha residue occurred in GPL-3, generating several highly methylated by-products that are not observed in WT *M. abscessus*. Overall, observation of the individual GPL profiles of single and multiple mutants of *M. abscessus* reflects a complex compensatory mechanism between the different *atf* genes that does not exist in *M. smegmatis*.

Three hypotheses for the biosynthesis of GPL in *M. abscessus* can be drawn from these observations, as summarized in **Figure 5D**. **(1)** The four identified putative acetyltransferases show partially redundant activities toward GPL-3 and probably GPL-2a. Specifically, deletions of *atf1* and *atf2* generate a modest accumulation of mono-acetylated GPL-3, which is exacerbated when both genes are deleted, thus suggesting that Atf1 and Atf2 can individually acetylate mono-acetylated GPL-3 to synthesize di-acetylated GPL-3. Conversely, deletion of *MAB_1725c* had no effect on GPL acetylation but completely inhibited the synthesis of di-acetylated GPL-3 and dramatically increased the accumulation of mono- and non-acetylated GPL-3 when deleted in Δ1,2. This strongly suggests that MAB_1725c can acetylate both mono- and non-acetylated GPL-3. Finally, one can observe that the individual deletion of *MAB_3448* induced no phenotype and that the Δ1,2,3448 showed the same phenotype than Δ1,2. This, combined with the fact that the *atf* quadruply deleted strain does not produce any acetylated GPL, suggests that MAB_3448 cannot use mono-acetylated GPL as a substrate to synthesize di-acetylated GPL. **(2)** Gtf3 has a higher activity on non-acetylated GPL-2a than acetylated GPL. We have previously demonstrated that Gtf3 uses di-acetylated GPL-2a to synthesize GPL-3 by transferring an additional Rha onto the internal Rha residue (15). The observation that GPL-2a content was strongly reduced when the *atf* genes were deleted put into light a totally unexpected regulatory process occurring during GPL synthesis that cannot be directly explained by the intrinsic activities of the enzymes. However, it may be explained by a modification in the metabolic processing of GPL-2a due to an increased activity of Gtf3 on non-acetylated GPL that would deplete the pool of GPL-2a toward

GPL-3. If so, acetylation of GPL would appear as a regulatory signal that maintains the GPL-2a/GPL-3 balance **(3)**. In a similar manner, the built-up of under-acetylated GPL-3 in mutant strains is coupled to the unexpected accumulation of hyper-methylated species that are not found in the WT-strain. It may be inferred that methylation acts as a compensatory mechanism to keep the overall hydrophobicity of GPL under the control of methyltransferases showing residual activity on under-acetylated GPL-3. Indeed, three genes *rmt2*, *rmt3* and *rmt4* located within the *gpl* locus **(Figure 1)** have orthologs in *M. smegmatis* and *M. avium* that are involved in the methylation of the first rhamnose (43, 44). Although attractive, these hypotheses should ideally be confirmed by studying the enzymatic activities of recombinantly produced acetyltransferases using individual acceptor substrates and identifying methyltransferases involved in the modification of GPL in *M. abscessus*.

Previous findings indicated that 6-dTal is a critical determinant promoting bacterial cell entry into macrophages (15), prompting us to investigate whether this was due to the loss of the monosaccharide and/or caused by the absence of the two acetyl groups. Quantitative analysis of the proportion of macrophages containing the various *atf* mutants at 3 hours post-infection revealed that all strains were equally taken up by THP-1 cells, thus ruling out a role of GPL acetylation in the early interaction between the bacilli and the macrophage surface. This was also supported when determining the proportion of macrophages containing beads coated with chemically de-acetylated GPL-2a *versus* di-acetylated GPL-2a or de-acetylated GPL-3 *versus* di-acetylated GPL-3. Thus, results from whole bacteria or purified GPL-coated beads suggest that GPL acetylation is unlikely to participate in the recognition between GPL and the macrophage receptor(s). From these results, it can be inferred that the reduced uptake of the *gtf1* mutant is likely a consequence of the absence of the 6-dTal rather than of the acetyl groups. Given the importance of GPL-3 in the internalization of *M. abscessus* S by macrophages (15), one would expect mutant strains with increased GPL-3/GPL-2a ratios to be more efficiently internalized by THP-1 cells than the WT strain, which is not the case for the *atf2*-deletion mutants. This could be due to hyper-*O*-methylation of GPL compensating for the loss of acetylation during interaction

between the bacilli and the host receptor(s). This hypothesis will be explored further in future research on the roles of the rhamnosyl-*O*-methyl transferases Rmt2, Rmt3 and Rmt4 in GPL structure and function.

An intriguing observation is that MAB_1725c is encoded by prophage prophiATCC19977-1, grouped within Subcluster MabA1. Prophages are prevalent and highly diverse among *M. abscessus* strains (41). However, *atf*-like genes are generally only present within Cluster MabA prophages, and only about half of all MabA prophages have an *atf* gene (41). These include prophage prophiGD17-2 and RNAseq analysis shows that its *atf* gene is lysogenically expressed – albeit at low levels – and can at least partially *O*-acetylate GPL in Δ1,2 **(Figure S14)**. Thus, at least in Δ1,2, the Atf encoded by the prophage participates in GPL biosynthesis and, therefore, in host cell wall assembly. These observations also suggest that GPL composition could vary among *M. abscessus* clinical isolates depending on the prophage status. Finally, we note that no *M. smegmatis* phages code for *atf* genes, but *Gordonia* phage Doggs and *Arthrobacter* phage Faja both code for Atf-like enzymes, and although GPL are not well characterized in these strains, it suggests that prophages may play broad roles in cell wall chemistry.

MATERIAL AND METHODS

Mycobacterial strains, growth conditions and reagents. All bacterial strains used for this study are listed in **Table S3**. Bacteria were grown in Middlebrook 7H9 broth (BD Difco) supplemented with 0.025% Tyloxapol and 10% oleic acid, albumin, dextrose, catalase (OADC enrichment) or on Middlebrook 7H10 agar (Difco) containing 10% OADC enrichment ($7H10^{OADC}$) at 37 °C, with antibiotics, if required. A Bio-Rad Gene pulser (25 µF, 2500 V, 800 Ohms) was used to transform electrocompetent mycobacteria. After transformation, strains carrying pTEC27 (Addgene, plasmid 30182) were selected in the presence of 1 mg/mL hygromycin for pTEC27 plasmid and strains carrying the pMV306 derivatives were selected on 250 µg/mL kanamycin.

Deletion of *atf* genes and complementation in *M. abscessus*. All deletion mutants were generated in the smooth (S) variant of the reference strain $CIP104536^{T}$ using the suicide vector pUX1-*katG* by double homologous recombination (37). Briefly, the upstream and downstream gene regions were PCR-amplified using the primers listed in **Table S2** and ligated into the PacI/NheI-linearized pUX1-*katG*. After transformation, bacteria were selected on $7H10^{OADC}$ supplemented with 250 µg/mL kanamycin with a visual screening of red fluorescent colonies, which have undergone the first homologous recombination. The second homologous recombination event was induced by INH counter-selection and selected on $7H10^{OADC}$ with 50 µg/mL INH and screening for non-fluorescent colonies (15, 37). This unmarked deletion system allowed to successively delete all four *atf* genes in the *M. abscessus*. Complementation of the single *atf* mutants was done by introducing the complementation plasmids generated using the integrative pMV306 **(Table S1)**. Genes in fusion with an HA-tagging sequence were amplified by PCR under the control of the *hsp60* promoter (15). Proper gene deletion and all plasmids were verified by DNA sequencing.

Western blotting. Bacteria were harvested, resuspended in PBS, and disrupted by bead beating using 1 mm diameter glass beads and a Mixer Mill MM 301 (Retsch, Germany) for 2 pulses of 3

min at 30Hz. Protein concentration was assessed using the BCA Protein Assay Reagent kit (Pierce), according to the manufacturer's instructions. Equal amounts of proteins (10 µg) were separated by 12% SDS-PAGE, transferred onto a nitrocellulose membrane, probed for 1 hr with either rat anti-HA (Sigma) or rat anti-KasA (loading control) antibodies. Membranes were washed and incubated for 45 min with goat anti-rat antibody conjugated to HRP (dilution 1:5000; Abcam). Bands were revealed using a SuperSignal West Femto (ThermoFisher Scientific) and a ChemiDoc MP system (Bio-Rad laboratories).

Colonial morphology, sedimentation and growth. Colony morphology was assessed from log phase cultures (OD_{600}=1) and streaked on Tryptic Soy agar. Plates were incubated 4 days at 37 °C and then imaged using a Zeiss microscope equipped with a Zeiss Plan Neo Fluor Z13/0.25 FWD objective. Images were taken with an Axiocam503 monochrome (Zeiss) camera and processed using ZEN 2 (blue edition). Sedimentation of bacterial aggregates was followed as reported earlier (24). Growth was initiated by inoculating mid-log phase cultures into fresh 7H9 at an OD_{600} of 0.05. Cultures were incubated at 37 °C with shaking and a sample of 1 mL of serially diluted culture was harvested every day and plated onto LB agar. After 4 days at 37 °C, colonies were counted to determine the colony forming unit (CFU).

Glycolipids extraction. Bacteria grown on $7H10^{OADC}$ agar plates without detergent were collected and lyophilized. 50 mg of bacterial pellets were weighted and apolar lipid fractions were first extracted in order to avoid TDM interference. GPL were extracted from the polar lipid fraction, first with chloroform/methanol/0.3%NaCl (9/10/3, v/v/v) and then with chloroform/methanol/0.3%NaCl (5:10:4, v/v/v). The combined solvent extracts were mixed for 5 min with chloroform and 0.3%NaCl (1:1, v/v) and centrifuged at 3,000 *g* for 5 min to separate the lower-organic phase from the aqueous phase. The upper aqueous layer was discarded and the lower-organic phase was evaporated under a stream of nitrogen and resuspended in chloroform/methanol (2:1, v/v).

TLC analysis. Apolar and polar lipids were subjected to TLC analysis. 10 µL of extract were spotted along 0.5 mm lane with glass capillary on Silica gel 60 F_{254} plates (Merck). GPL were separated in one or two migrations at 4 °C using chloroform/methanol/water (90:10:1, v/v/v) and more polar lipids were visualized with chloroform/methanol/water (65:25:4 or 30:8:1, v/v/v). Glycolipids were revealed by spraying the plates with orcinol in 20% sulfuric acid and charring.

GPL purification. Preparative TLC was performed on polar extracts: 300 µL were spotted on a 160 mm lane on a 60 µm silica gel plate with a glass back (20 cm x 20 cm) and migrated in a solution of chloroform/methanol/water (90:10:1, v/v/v). Plates were reversibly colored with iodide vapor to label bands, which were scrapped and GPL were further extracted in chloroform/methanol (2:1, v/v) under sonication for 1 hr. Silica was respectively filtrated on glass wool and through a 0.2 µm PTFE syringe filter.

Glycolipids saponification. 300 µL of the polar lipid fraction was dried under nitrogen, 1 mL sodium hydroxide 0.1 M in chloroform/methanol (1:1, v/v) was added and heated overnight at 37 °C. One mL butanol and 1 mL water were added and the mixture vortexed for 1 min and centrifuged for 30 sec. The upper butanolic phase was dried under nitrogen and dissolved in 300 µL chloroform/methanol (2:1, v/v).

Itol-acetates derivation. For hydrolysis step, 1 µg mesoinositol was added to the deacetylated GPL fraction then 1 mL 3M TFA was mixed, heated 4 hrs at 80 °C, dried and desiccated overnight. The reduction step was conducted for 4 hrs at room temperature in 500 µL $NaBH_4$ 10 mg/mL in 2M NH_4. The reaction was stopped with concentrated glacial acetic acid. Samples were dried at 55 °C under a nitrogen stream by co-distillation with methanol/acetic acid five times, dessicated overnight. Peracetylation was done by incubation in 500 µL anhydride acetic 4 hrs at 80 °C. The reaction products were extracted several times with chloroform/water. The chloroform-rich phase was then filtered, dried and dissolved in 500 µL chloroform. For GC-FID, 1 µL of itol-acetate

derivatives was injected in splitless mode with automatic sampler on a Solgel 1 MS 30 m x 0.25 mm x 0.25 µm capillary column with the following gradient temperature: 120 °C to 230 °C at 3 °C/min, then to 270 °C at 10 °C/min. Compounds were detected with flame ionization detector on a HP-7820 gas chromatograph (Agilent Technologies, Santa Clara, US). Previously determined retention times were used to identify each deoxyhexoses (15).

MALDI-TOF mass spectrometry. Before spotting 1 µL on the MALDI plate with a glass capillary tube, 10 µL of 20 mg/mL dihydroxybenzoïc acid (DHB) in chloroform/methanol (1:2, v/v) were mixed with 10 µL of the sample extract in chloroform/methanol (2:1, v/v). MS and MS^n spectra were acquired on an Axima Resonance (Shimadzu, Kyoto, Japan) in reflectron mode. For MS^2 experiments, ion selection was set from 250 to 500 Δm and collision energy was tuned from 300 eV to 600 eV. The GPL-2a/GPL-3 ratio was determined by dividing relative intensities of detected ion's pair at *m/z* 1174 and 1320 or 1202 and 1348 following 1/2, 1/3, or 1/5 sample dilution with the matrix.

Nuclear Magnetic Resonance. TLC-purified GPL were dried and dissolved in a mixture of $CDCl_3/CD_3OD$ (2:1, v/v) with 0.03% trimethylsilane (Eurisotop, France) three times then dissolved in a final volume of 270 µL. Samples were introduced into a 3 mm glass tube (Shigemi, Allison Park, PA, US). A TBI probe was used to observe ^1H and ^{13}C nuclei at 293K on an AVANCE II system (Bruker Biospin GmbH, Germany). Impulsion sequences used for homonuclear and heteronuclear experiments were from the manufacturer. After acquisition, phase correction and calibration on methanol signals were performed for δ ^1H and δ ^{13}C.

Internalization of *M. abscessus* by macrophages. THP-1 macrophages were grown in RPMI medium supplemented with 10% fetal bovine serum (Sigma-Aldrich) ($RPMI^{FBS}$), differentiated in the presence of 20 ng/mL phorbol myristate acetate in 24-well flat-bottom tissue culture microplates (10^5 cells/well) and incubated for 48 hrs at 37°C with 5% CO_2. Infection with *M.*

abscessus strains was performed for 3 hrs at 37 °C in the presence of 5% CO_2 with 2 bacteria per cell (MOI 2:1). Cells were carefully washed three times with PBS and then incubated with $RPMI^{FBS}$ supplemented with 250 µg/mL amikacin for 2 hrs to kill extracellular bacteria. The medium containing amikacin was discarded and cells were washed three times with PBS prior to assess CFU by lysing cells with 100 µL of 1% Triton X100 and plating serial dilutions of the homogenates. CFU were counted after 4 days of incubation at 37° C, as described earlier (15).

Fluorescent beads phagocytosis assay. Fluorescent beads were coated with purified acylated and de-acylated GPL-2a and GPL-3, as reported previously (15). THP-1 cells were incubated for 4 hrs at 37 °C with fluorescent beads (10 beads/macrophage), washed and stained using anti-CD43 antibodies and Alexa Fluor 594-coupled anti-mouse secondary antibody and DAPI. Percentage of macrophages containing beads were quantified using an epifluorescence microscope (15).

Statistical analyses.

Statistical analyses were carried out with Prism 9.0 (Graphpad). Details are given in the legend of each figure. ns, $P \geq 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$.

AUTHOR CONTRIBUTIONS

MI, LDL, and WD conducted experiments and analyzed the data. CD conducted experiments. YG and LK analyzed the data and wrote the manuscript with the help of MI, LDL, WD and GH. LK conceived the idea of the project.

ACKNOWLEDGMENTS

This study was supported by the French National Research Agency ANR-19-CE15-0012-01 (SUNLIVE) to LK and YG, and National Institutes of Health grant GM131729 to GFH. We acknowledge Vaincre la Mucoviscidose (RF20200502678) and the Association Grégory Lemarchal for funding the PhD fellowship of MI. We are grateful to the Montpellier RIO Imaging facility and to the PAGes core facility (http://plateforme- pages.univ-lille1.fr) and UAR 2014 - US 41 - Plateformes Lilloises en Biologie et Santé for providing the scientific and technical environment conducive to achieving this work. We thank SEA-PHAGES students Megan Ulbrich, Erin Bogler, and Dylan Leonard for their work on phages Doggs and Faja.

CONFLICT OF INTEREST

The funders had no role in study design, data collection, interpretation, or the decision to submit the work for publication. GFH receives support through a collaborative research agreement with Janssen Inc.

REFERENCES

1. Johansen, M. D., Herrmann, J.-L., and Kremer, L. (2020) Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. *Nat. Rev. Microbiol.* **18**, 392–407

2. Brown-Elliott, B. A., Nash, K. A., and Wallace, R. J. (2012) Antimicrobial susceptibility testing, drug resistance mechanisms, and therapy of infections with nontuberculous mycobacteria. *Clin. Microbiol. Rev.* **25**, 545–582

3. Sermet-Gaudelus, I., Le Bourgeois, M., Pierre-Audigier, C., Offredo, C., Guillemot, D., Halley, S., Akoua-Koffi, C., Vincent, V., Sivadon-Tardy, V., Ferroni, A., Berche, P., Scheinmann, P., Lenoir, G., and Gaillard, J.-L. (2003) *Mycobacterium abscessus* and children with cystic fibrosis. *Emerg. Infect. Dis.* **9**, 1587–1591

4. Esther, C. R., Esserman, D. A., Gilligan, P., Kerr, A., and Noone, P. G. (2010) Chronic *Mycobacterium abscessus* infection and lung function decline in cystic fibrosis. *J. Cyst. Fibros. Off. J. Eur. Cyst. Fibros. Soc.* **9**, 117–123

5. Howard, S. T., Rhoades, E., Recht, J., Pang, X., Alsup, A., Kolter, R., Lyons, C. R., and Byrd, T. F. (2006) Spontaneous reversion of *Mycobacterium abscessus* from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. *Microbiol. Read. Engl.* **152**, 1581–1590

6. Medjahed, H., Gaillard, J.-L., and Reyrat, J.-M. (2010) *Mycobacterium abscessus*: a new player in the mycobacterial field. *Trends Microbiol.* **18**, 117–123

7. Gutiérrez, A. V., Viljoen, A., Ghigo, E., Herrmann, J.-L., and Kremer, L. (2018) Glycopeptidolipids, a Double-Edged Sword of the *Mycobacterium abscessus* Complex. *Front. Microbiol.* **9**, 1145

8. Ripoll, F., Deshayes, C., Pasek, S., Laval, F., Beretti, J.-L., Biet, F., Risler, J.-L., Daffé, M., Etienne, G., Gaillard, J.-L., and Reyrat, J.-M. (2007) Genomics of glycopeptidolipid biosynthesis in *Mycobacterium abscessus* and *M. chelonae*. *BMC Genomics*. **8**, 114

9. Burbaud, S., Laval, F., Lemassu, A., Daffé, M., Guilhot, C., and Chalut, C. (2016) Trehalose Polyphleates Are Produced by a Glycolipid Biosynthetic Pathway Conserved across Phylogenetically Distant Mycobacteria. *Cell Chem. Biol.* **23**, 278–289

10. Vats, A., Singh, A. K., Mukherjee, R., Chopra, T., Ravindran, M. S., Mohanty, D., Chatterji, D., Reyrat, J.-M., and Gokhale, R. S. (2012) Retrobiosynthetic approach delineates the biosynthetic pathway and the structure of the acyl chain of mycobacterial glycopeptidolipids. *J. Biol. Chem.* **287**, 30677–30687

11. Wiersma, C. J., Belardinelli, J. M., Avanzi, C., Angala, S. K., Everall, I., Angala, B., Kendall, E., de Moura, V. C. N., Verma, D., Benoit, J., Brown, K. P., Jones, V., Malcolm, K. C., Strong, M., Nick, J. A., Floto, R. A., Parkhill, J., Ordway, D. J., Davidson, R. M., McNeil, M. R., and Jackson, M. (2020) Cell Surface Remodeling of *Mycobacterium abscessus* under Cystic Fibrosis Airway Growth Conditions. *ACS Infect. Dis.* **6**, 2143–2154

12. Villeneuve, C., Etienne, G., Abadie, V., Montrozier, H., Bordier, C., Laval, F., Daffe, M., Maridonneau-Parini, I., and Astarie-Dequeker, C. (2003) Surface-exposed glycopeptidolipids of *Mycobacterium smegmatis* specifically inhibit the phagocytosis of mycobacteria by human macrophages. Identification of a novel family of glycopeptidolipids. *J. Biol. Chem.* **278**, 51291–51300

13. Whang, J., Back, Y. W., Lee, K.-I., Fujiwara, N., Paik, S., Choi, C. H., Park, J.-K., and Kim, H.-J. (2017) *Mycobacterium abscessus* glycopeptidolipids inhibit macrophage apoptosis and bacterial spreading by targeting mitochondrial cyclophilin D. *Cell Death Dis.* **8**, e3012

14. Miyamoto, Y., Mukai, T., Nakata, N., Maeda, Y., Kai, M., Naka, T., Yano, I., and Makino, M. (2006) Identification and Characterization of the Genes Involved in Glycosylation Pathways of Mycobacterial Glycopeptidolipid Biosynthesis. *J. Bacteriol.* **188**, 86–95

15. Daher, W., Leclercq, L.-D., Johansen, M. D., Hamela, C., Karam, J., Trivelli, X., Nigou, J., Guérardel, Y., and Kremer, L. (2022) Glycopeptidolipid glycosylation controls surface

properties and pathogenicity in *Mycobacterium abscessus*. *Cell Chem. Biol.* **29**, 910-924.e7

16. Bernut, A., Viljoen, A., Dupont, C., Sapriel, G., Blaise, M., Bouchier, C., Brosch, R., de Chastellier, C., Herrmann, J.-L., and Kremer, L. (2016) Insights into the smooth-to-rough transitioning in *Mycobacterium bolletii* unravels a functional Tyr residue conserved in all mycobacterial MmpL family members. *Mol. Microbiol.* **99**, 866–883

17. Medjahed, H., and Reyrat, J.-M. (2009) Construction of *Mycobacterium abscessus* defined glycopeptidolipid mutants: comparison of genetic tools. *Appl. Environ. Microbiol.* **75**, 1331–1338

18. Nessar, R., Reyrat, J.-M., Davidson, L. B., and Byrd, T. F. (2011) Deletion of the mmpL4b gene in the *Mycobacterium abscessus* glycopeptidolipid biosynthetic pathway results in loss of surface colonization capability, but enhanced ability to replicate in human macrophages and stimulate their innate immune response. *Microbiol. Read. Engl.* **157**, 1187–1195

19. Viljoen, A., Dubois, V., Girard-Misguich, F., Blaise, M., Herrmann, J.-L., and Kremer, L. (2017) The diverse family of MmpL transporters in mycobacteria: from regulation to antimicrobial developments. *Mol. Microbiol.* **104**, 889–904

20. Viljoen, A., Viela, F., Kremer, L., and Dufrêne, Y. F. (2020) Fast chemical force microscopy demonstrates that glycopeptidolipids define nanodomains of varying hydrophobicity on mycobacteria. *Nanoscale Horiz.* **5**, 944–953

21. Ryan, K., and Byrd, T. F. (2018) *Mycobacterium abscessus*: Shapeshifter of the Mycobacterial World. *Front. Microbiol.* **9**, 2642

22. Madani, A., Ridenour, J. N., Martin, B. P., Paudel, R. R., Abdul Basir, A., Le Moigne, V., Herrmann, J.-L., Audebert, S., Camoin, L., Kremer, L., Spilling, C. D., Canaan, S., and Cavalier, J.-F. (2019) Cyclipostins and Cyclophostin Analogues as Multitarget Inhibitors That Impair Growth of *Mycobacterium abscessus*. *ACS Infect. Dis.* **5**, 1597–1608

23. Johansen, M. D., Daher, W., Roquet-Banères, F., Raynaud, C., Alcaraz, M., Maurer, F. P., and Kremer, L. (2020) Rifabutin Is Bactericidal against Intracellular and Extracellular Forms of *Mycobacterium abscessus*. *Antimicrob. Agents Chemother.* **64**, e00363-20

24. Daher, W., Leclercq, L.-D., Viljoen, A., Karam, J., Dufrêne, Y. F., Guérardel, Y., and Kremer, L. (2020) O-Methylation of the Glycopeptidolipid Acyl Chain Defines Surface Hydrophobicity of *Mycobacterium abscessus* and Macrophage Invasion. *ACS Infect. Dis.* **6**, 2756–2770

25. Viljoen, A., Gutiérrez, A. V., Dupont, C., Ghigo, E., and Kremer, L. (2018) A Simple and Rapid Gene Disruption Strategy in *Mycobacterium abscessus*: On the Design and Application of Glycopeptidolipid Mutants. *Front. Cell. Infect. Microbiol.* **8**, 69

26. Bernut, A., Herrmann, J.-L., Kissa, K., Dubremetz, J.-F., Gaillard, J.-L., Lutfalla, G., and Kremer, L. (2014) *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess formation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E943-952

27. Roux, A.-L., Viljoen, A., Bah, A., Simeone, R., Bernut, A., Laencina, L., Deramaudt, T., Rottman, M., Gaillard, J.-L., Majlessi, L., Brosch, R., Girard-Misguich, F., Vergne, I., de Chastellier, C., Kremer, L., and Herrmann, J.-L. (2016) The distinct fate of smooth and rough *Mycobacterium abscessus* variants inside macrophages. *Open Biol.* **6**, 160185

28. Roux, A.-L., Ray, A., Pawlik, A., Medjahed, H., Etienne, G., Rottman, M., Catherinot, E., Coppée, J.-Y., Chaoui, K., Monsarrat, B., Toubert, A., Daffé, M., Puzo, G., Gaillard, J.-L., Brosch, R., Dulphy, N., Nigou, J., and Herrmann, J.-L. (2011) Overexpression of proinflammatory TLR-2-signalling lipoproteins in hypervirulent mycobacterial variants. *Cell. Microbiol.* **13**, 692–704

29. Bernut, A., Herrmann, J.-L., Ordway, D., and Kremer, L. (2017) The Diverse Cellular and Animal Models to Decipher the Physiopathological Traits of *Mycobacterium abscessus* Infection. *Front. Cell. Infect. Microbiol.* **7**, 100

30. Catherinot, E., Clarissou, J., Etienne, G., Ripoll, F., Emile, J.-F., Daffé, M., Perronne, C., Soudais, C., Gaillard, J.-L., and Rottman, M. (2007) Hypervirulence of a rough variant of the *Mycobacterium abscessus* type strain. *Infect. Immun.* **75**, 1055–1058

31. Catherinot, E., Roux, A.-L., Macheras, E., Hubert, D., Matmar, M., Dannhoffer, L., Chinet, T., Morand, P., Poyart, C., Heym, B., Rottman, M., Gaillard, J.-L., and Herrmann, J.-L. (2009) Acute respiratory failure involving an R variant of *Mycobacterium abscessus*. *J. Clin. Microbiol.* **47**, 271–274

32. Jönsson, B. E., Gilljam, M., Lindblad, A., Ridell, M., Wold, A. E., and Welinder-Olsson, C. (2007) Molecular epidemiology of *Mycobacterium abscessus*, with focus on cystic fibrosis. *J. Clin. Microbiol.* **45**, 1497–1504

33. Pawlik, A., Garnier, G., Orgeur, M., Tong, P., Lohan, A., Le Chevalier, F., Sapriel, G., Roux, A.-L., Conlon, K., Honoré, N., Dillies, M.-A., Ma, L., Bouchier, C., Coppée, J.-Y., Gaillard, J.-L., Gordon, S. V., Loftus, B., Brosch, R., and Herrmann, J. L. (2013) Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent *Mycobacterium abscessus*. *Mol. Microbiol.* **90**, 612–629

34. Park, I. K., Hsu, A. P., Tettelin, H., Shallom, S. J., Drake, S. K., Ding, L., Wu, U.-I., Adamo, N., Prevots, D. R., Olivier, K. N., Holland, S. M., Sampaio, E. P., and Zelazny, A. M. (2015) Clonal Diversification and Changes in Lipid Traits and Colony Morphology in *Mycobacterium abscessus* Clinical Isolates. *J. Clin. Microbiol.* **53**, 3438–3447

35. Dedrick, R. M., Smith, B. E., Garlena, R. A., Russell, D. A., Aull, H. G., Mahalingam, V., Divens, A. M., Guerrero-Bustamante, C. A., Zack, K. M., Abad, L., Gauthier, C. H., Jacobs-Sera, D., and Hatfull, G. F. (2021) *Mycobacterium abscessus* Strain Morphotype Determines Phage Susceptibility, the Repertoire of Therapeutically Useful Phages, and Phage Resistance. *mBio*. **12**, e03431-20

36. Recht, J., and Kolter, R. (2001) Glycopeptidolipid acetylation affects sliding motility and biofilm formation in *Mycobacterium smegmatis*. *J. Bacteriol.* **183**, 5718–5724

37. Richard, M., Gutiérrez, A. V., Viljoen, A., Rodriguez-Rincon, D., Roquet-Baneres, F., Blaise, M., Everall, I., Parkhill, J., Floto, R. A., and Kremer, L. (2019) Mutations in the MAB_2299c TetR Regulator Confer Cross-Resistance to Clofazimine and Bedaquiline in *Mycobacterium abscessus*. *Antimicrob. Agents Chemother.* **63**, e01316-18

38. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., and Hatfull, G. F. (1991) New use of BCG for recombinant vaccines. *Nature*. **351**, 456–460

39. Jia Khor, M., Broda, A., Kostrzewa, M., Drobniewski, F., and Larrouy-Maumus, G. (2021) An Improved Method for Rapid Detection of *Mycobacterium abscessus* Complex Based on Species-Specific Lipid Fingerprint by Routine MALDI-TOF. *Front. Chem.* **9**, 715890

40. Ripoll, F., Pasek, S., Schenowitz, C., Dossat, C., Barbe, V., Rottman, M., Macheras, E., Heym, B., Herrmann, J.-L., Daffé, M., Brosch, R., Risler, J.-L., and Gaillard, J.-L. (2009) Non mycobacterial virulence genes in the genome of the emerging pathogen *Mycobacterium abscessus*. *PloS One*. **4**, e5660

41. Dedrick, R. M., Aull, H. G., Jacobs-Sera, D., Garlena, R. A., Russell, D. A., Smith, B. E., Mahalingam, V., Abad, L., Gauthier, C. H., and Hatfull, G. F. (2021) The Prophage and Plasmid Mobilome as a Likely Driver of *Mycobacterium abscessus* Diversity. *mBio*. **12**, e03441-20

42. Gao, J., and Sampson, N. S. (2014) A GMC oxidoreductase homologue is required for acetylation of glycopeptidolipid in *Mycobacterium smegmatis*. *Biochemistry*. **53**, 611–613

43. Jeevarajah, D., Patterson, J. H., Taig, E., Sargeant, T., McConville, M. J., and Billman-Jacobe, H. (2004) Methylation of GPLs in *Mycobacterium smegmatis* and *Mycobacterium avium*. *J. Bacteriol.* **186**, 6792–6799

44. Patterson, J. H., McConville, M. J., Haites, R. E., Coppel, R. L., and Billman-Jacobe, H. (2000) Identification of a Methyltransferase from *Mycobacterium smegmatis* Involved in Glycopeptidolipid Synthesis. *J. Biol. Chem.* **275**, 24900–24906

Figure 1: Prediction of GPL acetyltransferases in *M. abscessus.* (A) The two major structures of GPL are schematized with a lipid chain (green), a peptide core (blue) and a glycan moiety (pink). Tri-glycosylated GPL, designated GPL-3, is composed of a 6-dTal and two Rha residues. Di-glycosylated GPL, designated GPL-2a, possesses only one Rha. The acetyl groups on C3 and C4 positions of 6-dTal are highlighted by a red circle. Internal Rha residues are methylated in C3 and C4 positions. **(B)** *M. smegmatis* and *M. abscessus gpl* loci encode enzymes for the synthesis, modification and transport of GPL.

Figure 2: The GPL profile varies according to deleted acetyltransferases. (A) TLC analysis of the polar lipid fraction of wild type (WT) and acetyltransferases mutant strains. Previously characterized GPL-2a and GPL-3 mobility (15, 24) are shown on the left side (2a and 3); newly observed accumulating GPL are marked on the right side as white, grey and black arrows; diacyl-trehalose isomers (DAT) identified by NMR are highlighted with an asterisk. **(B)** MALDI-MS spectra in positive mode of the polar lipid fraction of wild type (WT) and selected acetyltransferases mutant strains. Previously characterized GPL-2a and GPL-3 (15, 24) MS signals are shown on top of the spectra. **(C)** MS^2 fragmentation spectrum of the selected parent ion at *m/z* 1404. Fragmentation patterns are illustrated for two GPL-3 isomers with alaninol and C_{28} lipid (Ala-C_{28}) or valinol and C_{26} lipid (Val-C_{26}). Each one produces the same fragment ions when di-*O*-acetylated 6-dTal and/or Rha are lost, but Y fragments ions are shifted by 28 m.u. along the peptidic backbone.

Figure 3: Paralogues of *atf1* and *atf2* are present in the *M. abscessus* genome. (A) Genomic organization of two putative acetyltransferases encoding genes, *MAB_1725c* and *MAB_3448*. **(B)** Alignment of the protein sequence of the *M. smegmatis* GPL acetyltransferase (Atf) with the putative *M. abscessus* GPL acetyltransferases. **(C)** Cross-tabulation of the percentage of identity between the five acetyltransferases.

Figure 4: GPL from *M. abscessus* Δ1,2,1725c and Δ1,2,1725c,3448 are under-acetylated and hyper-methylated. (A, B) GPL profile of Δ1,2,1725c,3448; **(C, D)** GPL profile of Δ1,2,1725c. **(A, C)** TLC analyses of Δ1,2,1725c and Δ1,2,1725c,3448 before and after separation in three (F1 to F3) or four (F1 to F4) fractions, respectively, as compared to wild type (WT). White, grey and black arrow correspond to the accumulating products observed on TLC plates of total polar fractions. **(B and D)** MALDI-MS spectra of Δ1,2,1725c,3448-F1 to F3 and Δ1,2,1725c-F1 to F4 following separation of total polar fractions. Increasing *m/z* values are associated with additional acetyl group (yellow star) on the 6-dTal and *O*-methyl groups (red circles) on the terminal Rha. GPL schemes are drawn according to single-letter amino-acids code and SNFG nomenclature.

Figure 5: Deletion of the *atf* genes alters the GPL pattern. (A) Monosaccharide composition of purified de-acetylated GPL. The relative proportion of Rha residues increased in all mutants lacking *atf2*, as well as *O*-methyl Rha residues in Δ1,2,1725c and Δ1,2,1725c,3448. **(B)** TLC analysis of the saponified polar lipid fraction highlights strong modifications of GPL profiles in all mutants deleted in *atf1* and *atf2* genes. Deacetylated dGPL-2a and dGPL-3 were attributed owing to saponification of the purified GPL-2a and GPL-3 as well as purified fractions from Δ1,2,1725c and Δ1,2,1725c,3448 **(Figure S13A and B)**. The composition of rhamnan moiety is shown to the right of the TLC. **(C)** dGPL-2a/dGPL-3 ratios in all mutant strains were established by integrating relative intensities of ions at *m/z* 1174 *versus* 1320 as well as 1202 *versus* 1348. In all mutants lacking *atf2* gene, the relative proportion of GPL-2a is significantly reduced. Median values are shown and results (n=8) were analyzed using the Friedman *t* test. *P ≤ 0.05, ** P ≤ 0.01**,** ***P ≤ 0.001. **(D)** Proposed biosynthetic pathway of GPL modification by glycosylation (*gtf*), acetylation (*atf*) and methylation (*rmt*). Acetyl groups are illustrated by yellow stars on the 6-dTal and *O*-methyl groups by red circles. All four identified acetyltransferases show partial overlapping activities. MAB_1725c is postulated to be active on both non-acetylated and mono-acetylated GPL, whereas MAB_3448 to be active only on non-acetylated GPL. The rhamnosyltransferase Gtf3 is thought to be more

active on non-acetylated substrates, as depicted by the width of the arrows, which would explain the accumulation of under-acetylated GPL-3 that, in turn, serves as substrates for the rhamnosyl-methyltansferases (Rmt) to synthetize the over-methylated GPL-3 identified in the mutants. Under-acetylated GPL-2a (in brackets) are not observed in any of the studied strains as they represent biosynthetic intermediates that do not accumulate.

Figure 6: Impact of GPL acetylation on bacterial growth, morphology and sedimentation. (A) Colony morphology of the wild-type *M. abscessus* smooth and rough strains, *atf* mutants and complemented strains. Scale bars represent 1 mm. **(B)** Sedimented bacterial aggregates are indicated by arrows. **(C)** Growth curve of reference strains (S and R), Δ1,2, Δ1,2,1725c, Δ1,2,3448 and Δ1,2,1725c,3448 at 37°C for 5 days with agitation. Experiment was performed 3 times with one measure per day per strain. **(D)** Internalization by macrophages of different *atf* mutants and wild-type strain (S) after 3 hrs of infection (MOI 2:1). Intracellular bacteria are expressed as CFU/mL and error bars represent standard deviations. Four independent experiments performed in triplicate are plotted on the graph and analyzed using the two-tailed Mann-Whitney *t* test. **(E)** Influence of GPL acetylation in macrophage infection was evaluated with fluorescent beads coated with di- (GPL-2a) or tri-glycosylated (GPL-3) GPL, either di- or de-acetylated (dGPL) using 10 beads/cell. Non-coated beads were included as a control group. Each symbol represents the percentage of bead-containing macrophages in one field. Experiments were conducted 3 times (n=60) and data are expressed as mean values ± SD. Results were analyzed using the Mann-Whitney *t* test. ns, non-significant; ****P ≤ 0.0001. **(F)** Representative fields of fluorescent beads (green) internalized by THP-1 cells. DAPI (blue) was used to stain the nuclei and CD43 antibodies (red) to detect the cell membrane. Scale bar, 50 µm.

Figure 1

Click here to access/download;Main Figure (High Resolution);Figure 1.pdf

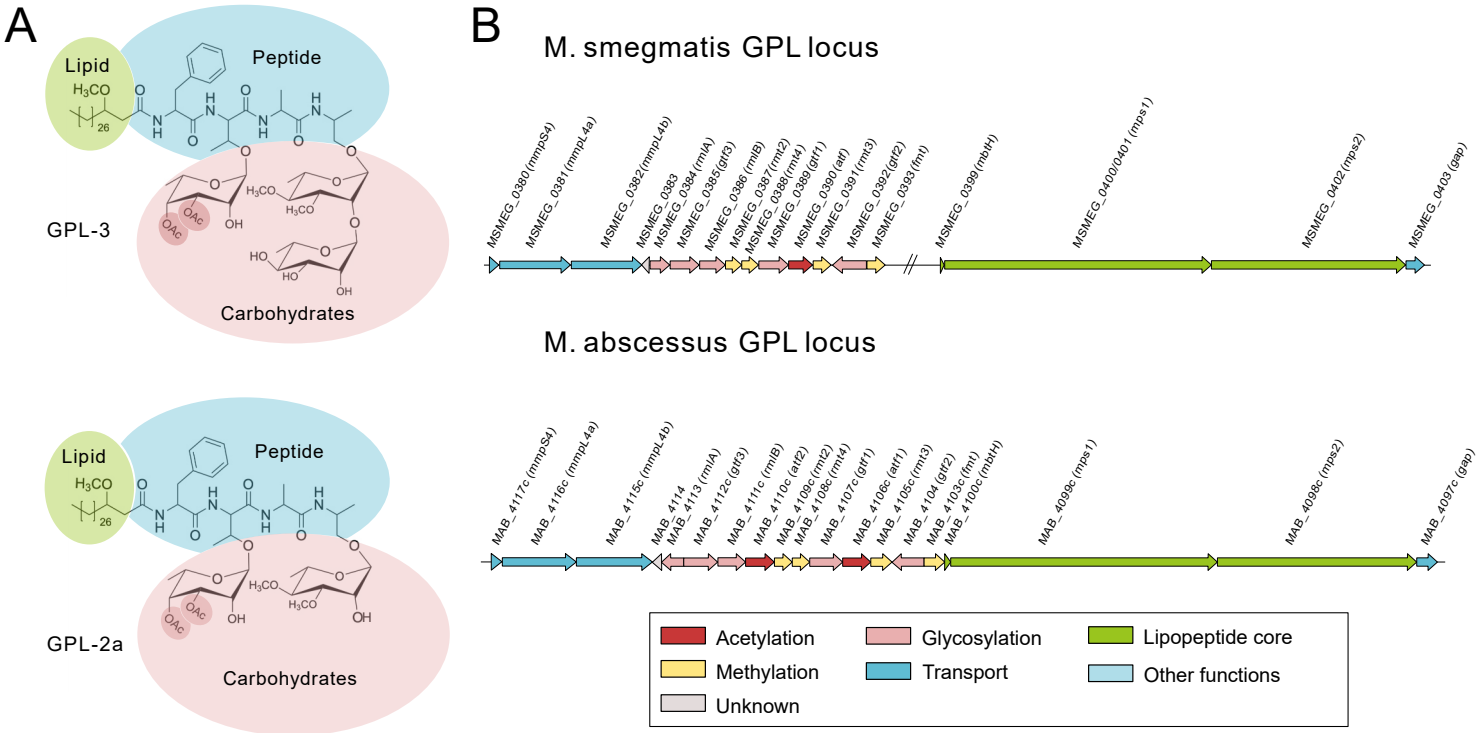


Figure 1

Figure 2

Click here to access/download;Main Figure (High Resolution);Figure 2.pdf

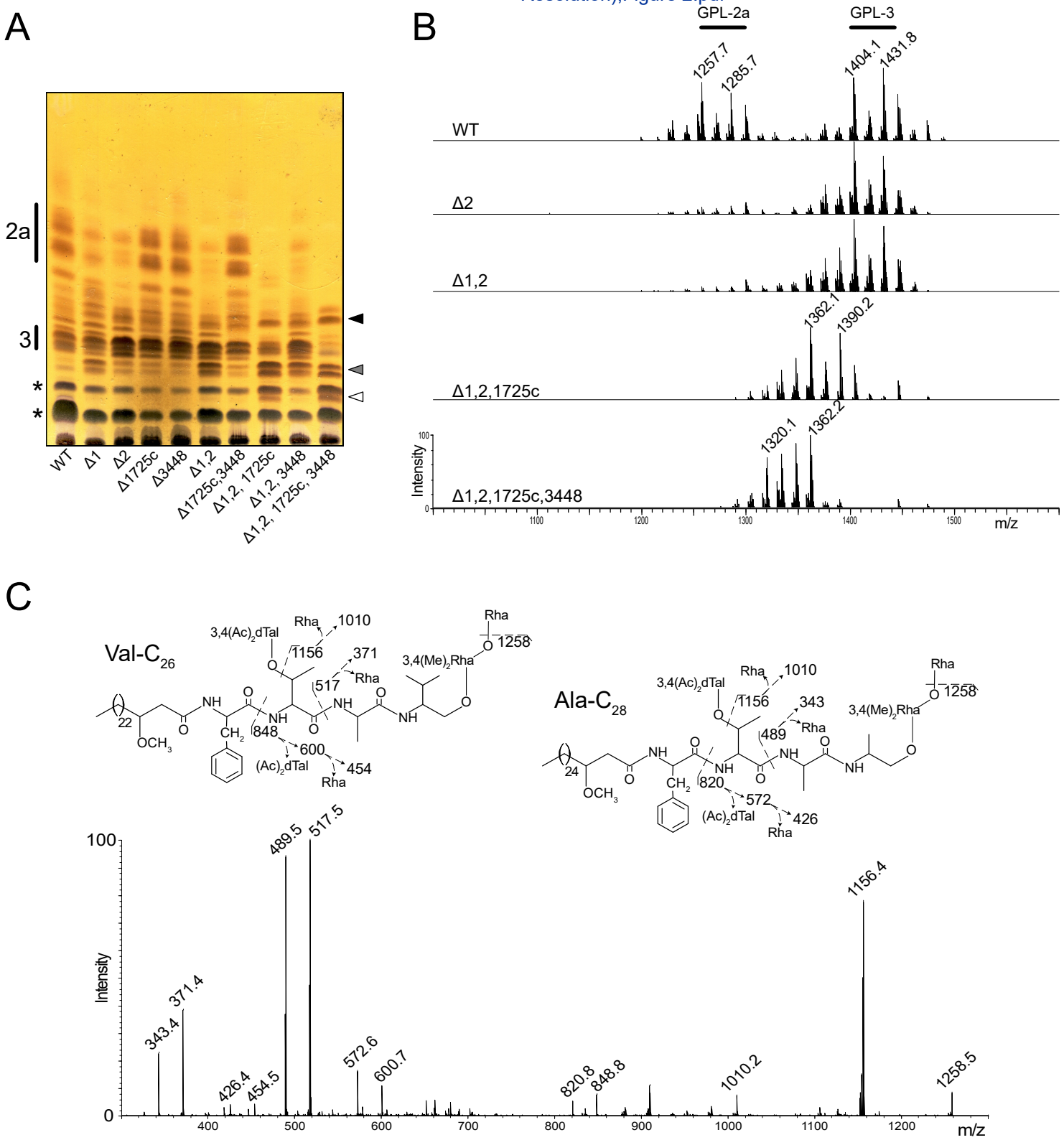
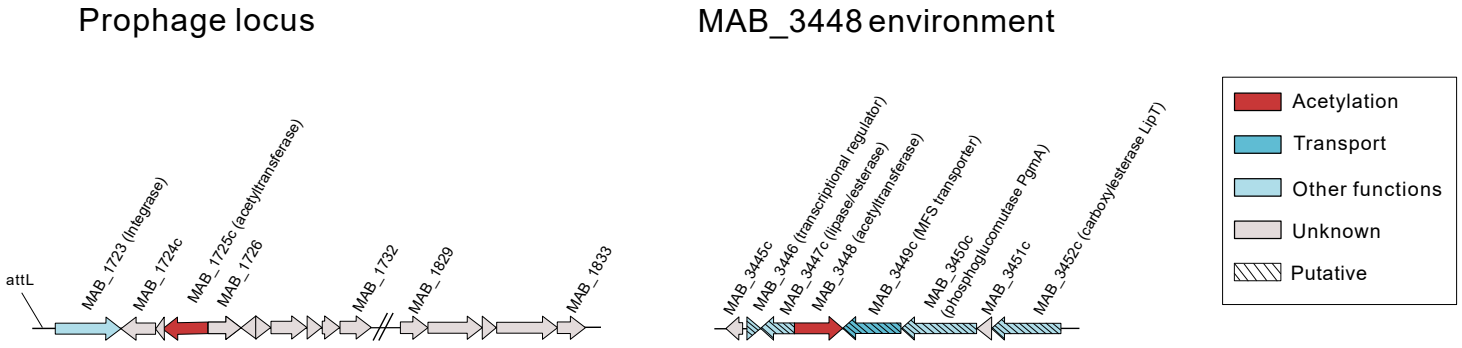


Figure 2

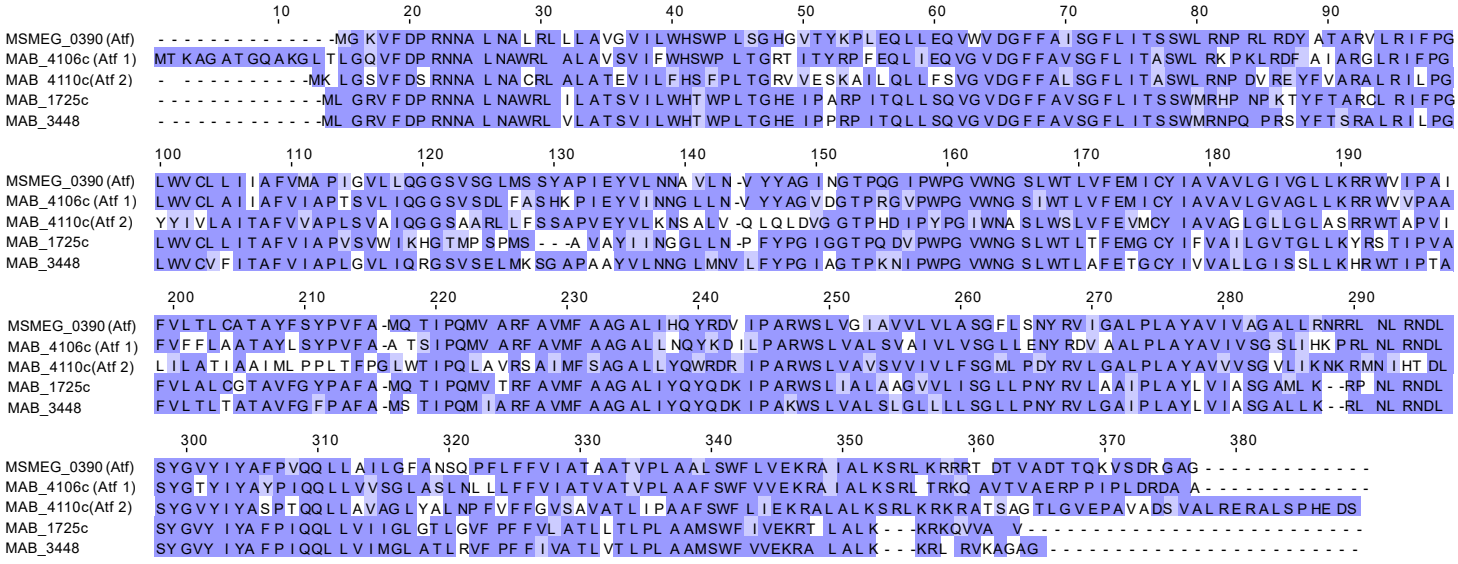
Figure 3

Click here to access/download;Main Figure (High Resolution);Figure 3.pdf

A



B



C

	Atf MSMEG_0390	Atf1 MAB_4106c	Atf2 MAB_4110	MAB_1725c	MAB_3448
Atf MSMEG_0390	100.00	71.94	55.83	65.78	68.31
Atf1 MAB_4106c	71.94	100.00	55.80	64.71	65.80
Atf2 MAB_4110c	55.83	55.80	100.00	53.53	56.52
MAB_1725c	65.78	64.71	53.53	100.00	78.82
MAB_3448	68.31	65.80	56.52	78.82	100.00

Figure 3

Click here to access/download;Main Figure (High Resolution);Figure 4.pdf

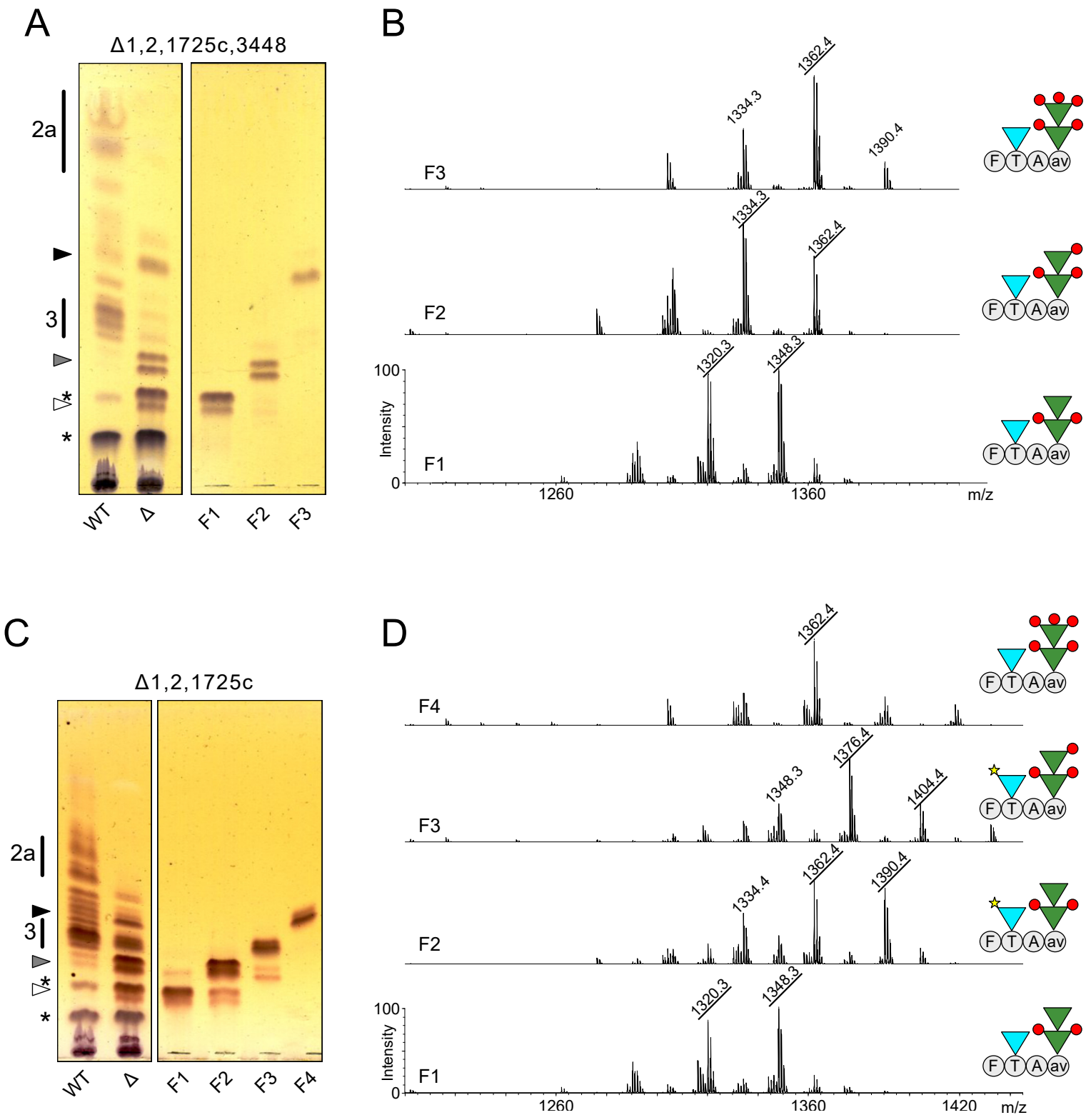


Figure 4

Click here to access/download;Main Figure (High Resolution);Figure 5.pdf

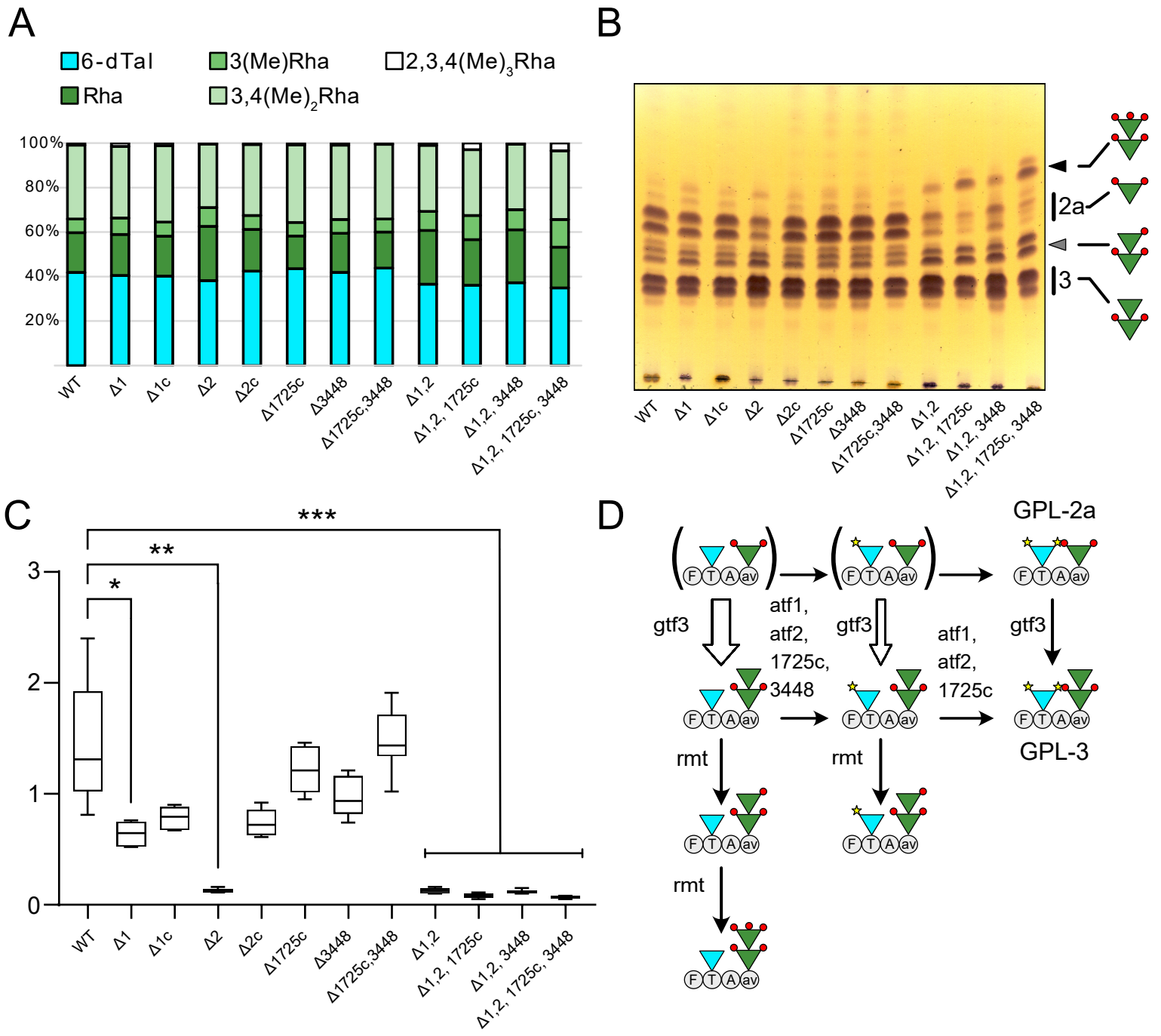


Figure 5

Figure 6

Click here to access/download;Main Figure (High Resolution);Figure 6.pdf

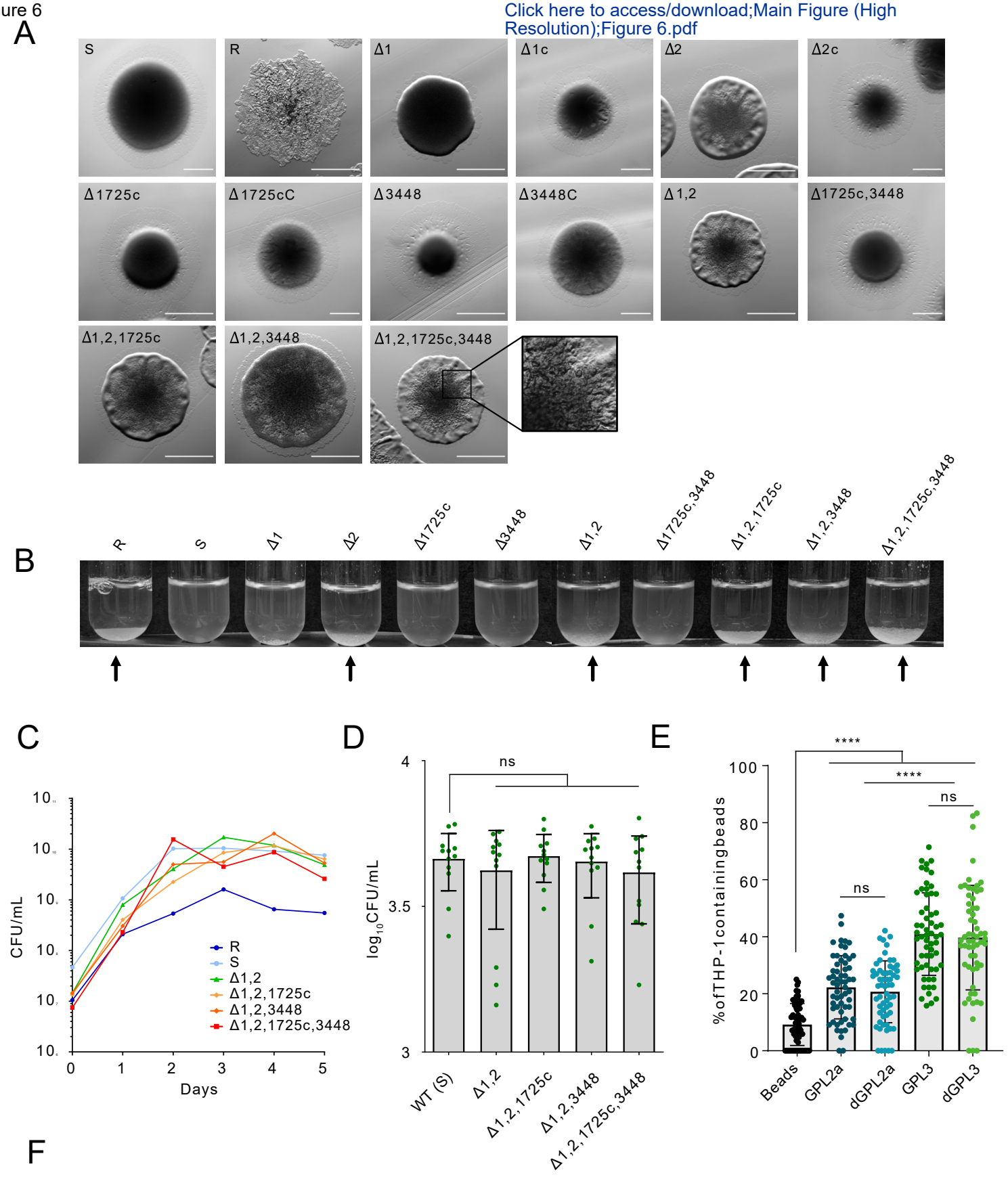


Figure 6

AUTHOR CREDIT STATEMENT

MI, LDL, and WD conducted experiments and analyzed the data. CD conducted experiments. YG and LK analyzed the data and wrote the manuscript with the help of MI, LDL, WD and GH. LK conceived the idea of the project.

Supporting information

Multiple *Mycobacterium abscessus O*-acetyltransferases influence glycopeptidolipid structure

Morgane Illouz[1,&], Louis-David Leclercq[2,&], Clara Dessenne[2], Graham Hatfull[3], Wassim Daher[1,4], Laurent Kremer[1,4,#], and Yann Guérardel[2,5#]

[1]Centre National de la Recherche Scientifique UMR 9004, Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France.
[2]Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France.
[3]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, US
[4]INSERM, IRIM, 34293 Montpellier, France.
[5]Institute for Glyco-core Research (iGCORE), Gifu University, Gifu, Japan

[&]These authors contributed equally to this work.

[#]For correspondence: E-mails yann.guerardel@univ-lille.fr, Tel: (+33) 3 20 43 69 41; laurent.kremer@irim.cnrs.fr, Tel: (+33) 4 34 35 94 47.

Running title: Acetylation of GPL in *Mycobacterium abscessus*

Keywords: *Mycobacterium abscessus*, glycopeptidolipid, cell wall, acetyltransferase, macrophage, mass spectrometry.

SUPPLEMENTARY FIGURES

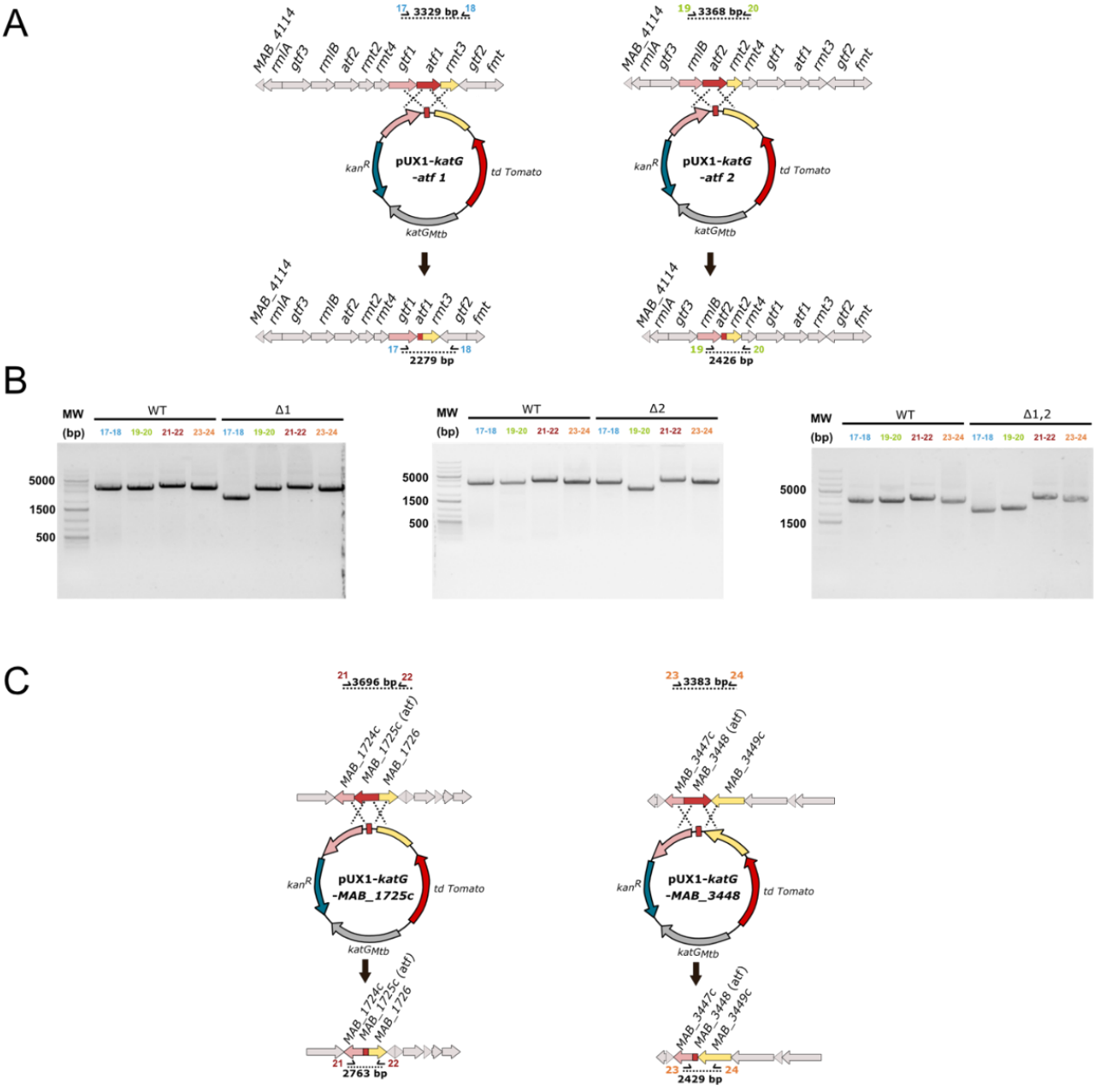


Figure S1: Generation of the *atf* mutants in *M. abscessus*. (A, C) The gene of interest to be inactivated in the bacterial chromosome is colored in red. The upstream and downstream regions of the gene of interest are in pink and yellow, respectively. Crosses represent homologous recombination events. Primers used are listed in **Table S2**. The intact locus is drawn above the plasmid and the modified locus is shown below the plasmid. **(B, D)** PCR profile of the different mutants compared to WT strain. To amplify the different loci, genomic DNA was used as template. Sequencing of the amplicons was performed to confirm the proper deletions.

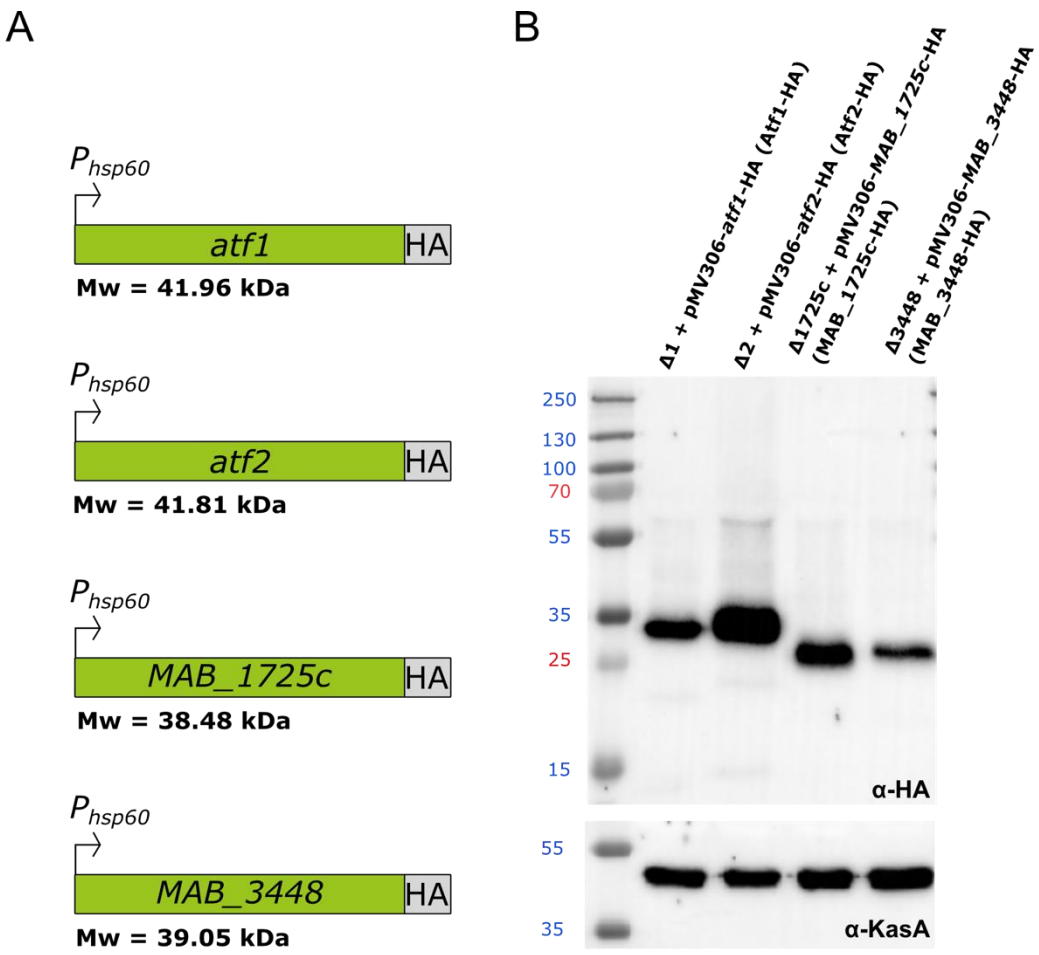


Figure S2: Acetyltransferase complementation in *M. abscessus* single mutants. (A) Illustrative diagrams of the gene of interest fused with an HA-tag on the C terminal side produced under the control of the *hsp60* promoter. The molecular weight of the protein with the HA-tag is indicated below each diagram. **(B)** Western blot analysis showing the production of the four Atf proteins in fusion with a HA-tag (upper panel). Kas A is shown as an internal loading control (lower panel).

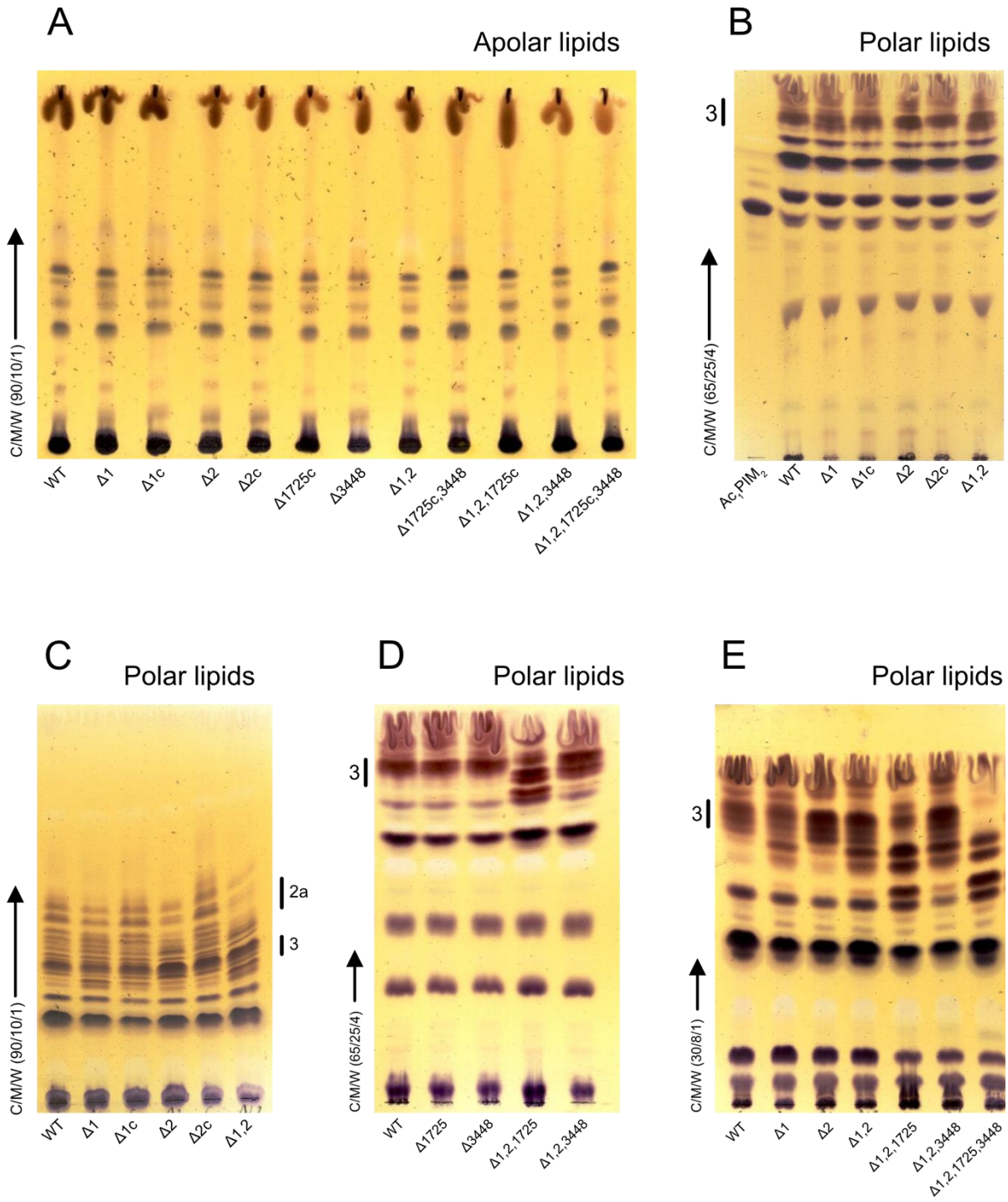


Figure S3: TLC analysis of polar and apolar fractions from mutant strains developed in different solvents. Only GPL showed significant modifications in mutant strains. **(A)** Petroleum ether extracted apolar lipid fraction of wild type (WT) and all acetyltransferases mutant strains. **(B, C)** Polar glycolipids from wild type (WT), Δ1 and complemented strain, Δ2 and complemented strain, Δ1,2 showed phosphatidyl *myo*-inositol mannosides and GPL profile. **(D, E)** Polar glycolipids from WT, Δ1, Δ2, Δ1725c, Δ3448 as well as Δ1,2 and Δ1,2,1725c and Δ1,2,3448 mutants.

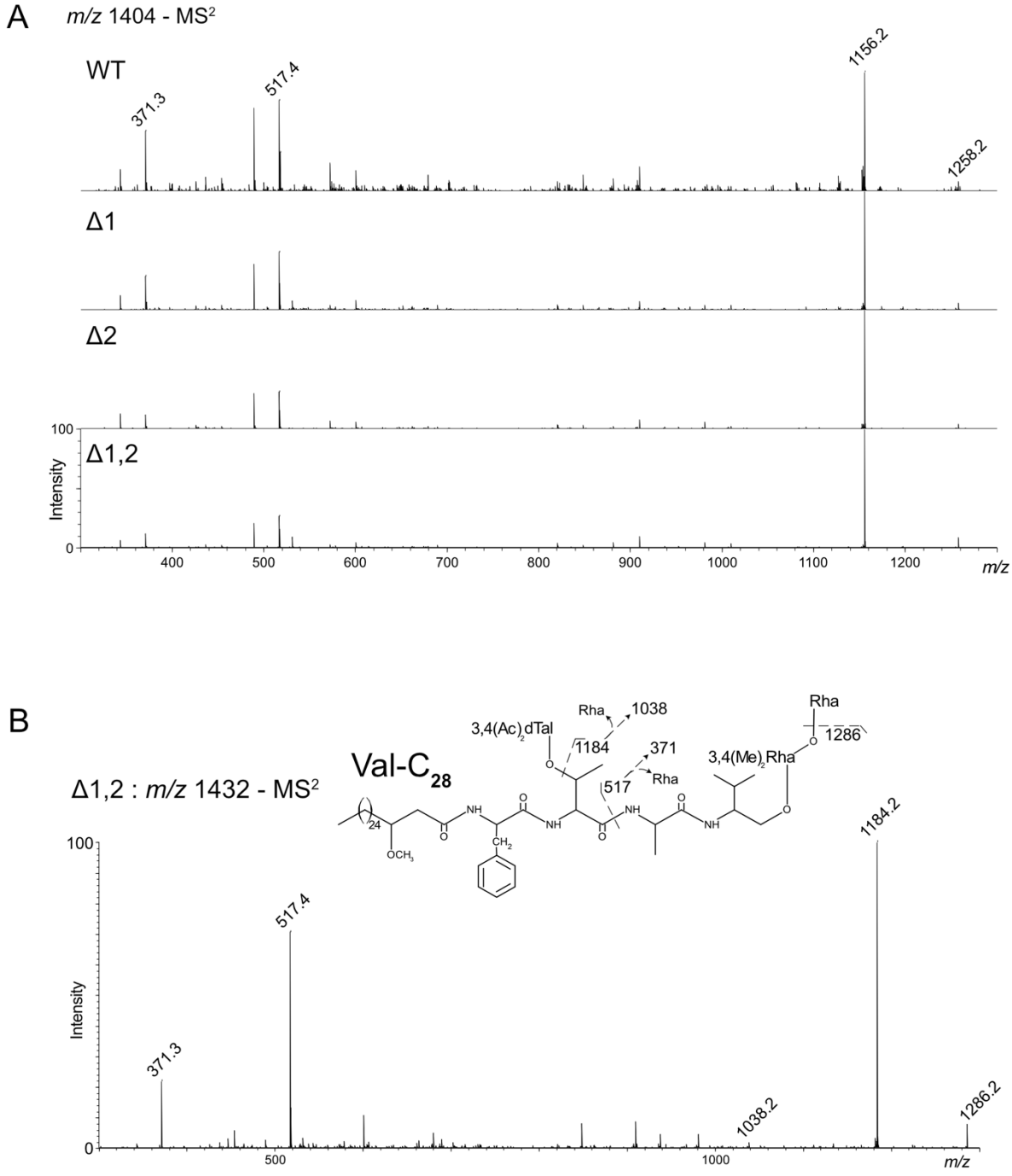


Figure S4: (A) MALDI-MS^2 spectra in positive mode of the selected parent ion at *m/z* 1404 from wild type (WT), Δ1, Δ2 and Δ1,2 mutant strains. **(B)** MALDI-MS2 spectra in positive mode of the selected parent ion at m/z 1432 of Δ1,2. Fragmentation scheme shows the presence of 3,4 di-*O*-acetylated 6d-Tal in valinol form with C28 lipid (Val-C28).

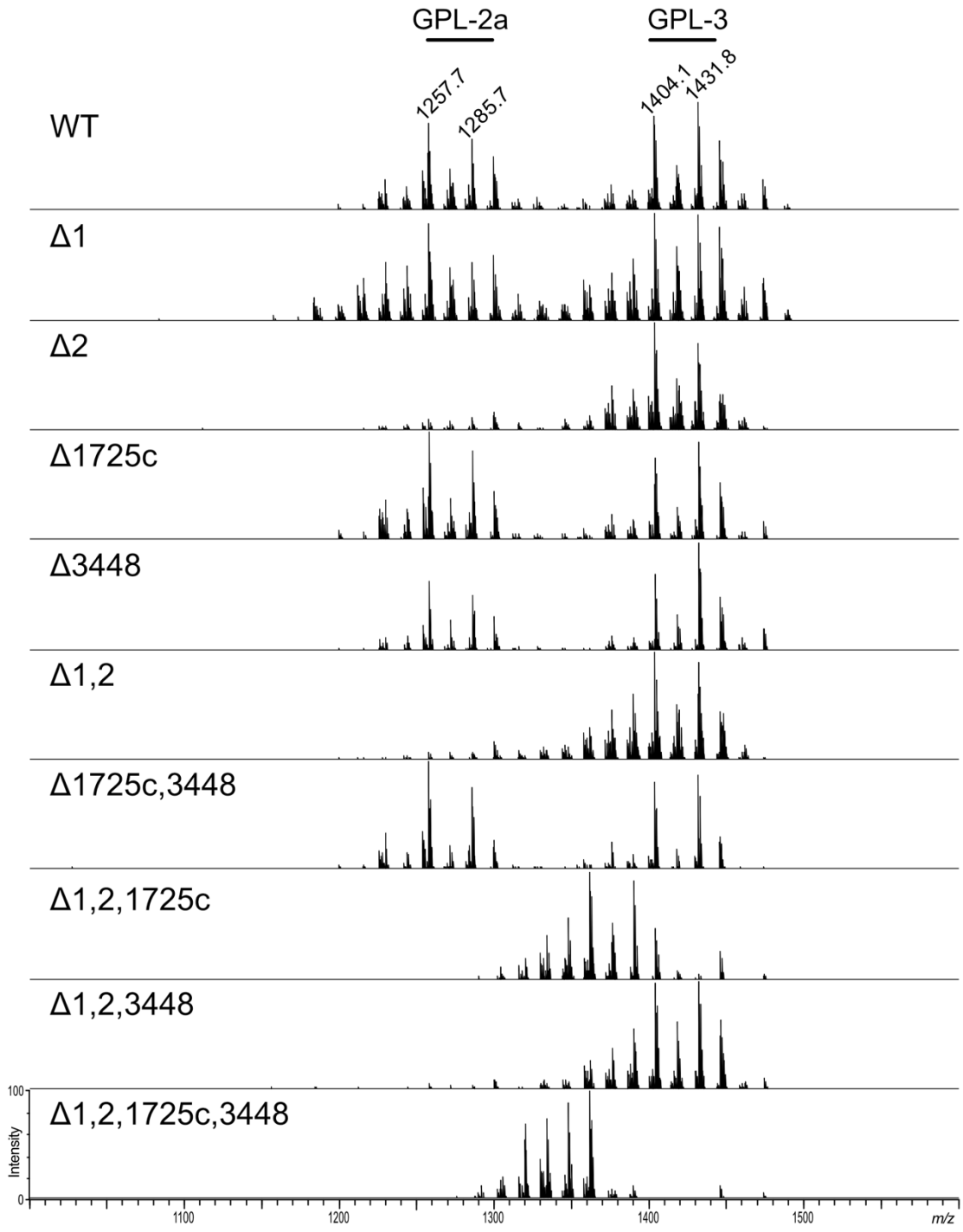


Figure S5: MALDI-MS spectra in positive mode of the polar lipid fraction from wild type and acetyltransferase mutants. Previously characterized GPL-2a and GPL-3 (1, 2) are shown on top of the spectra.

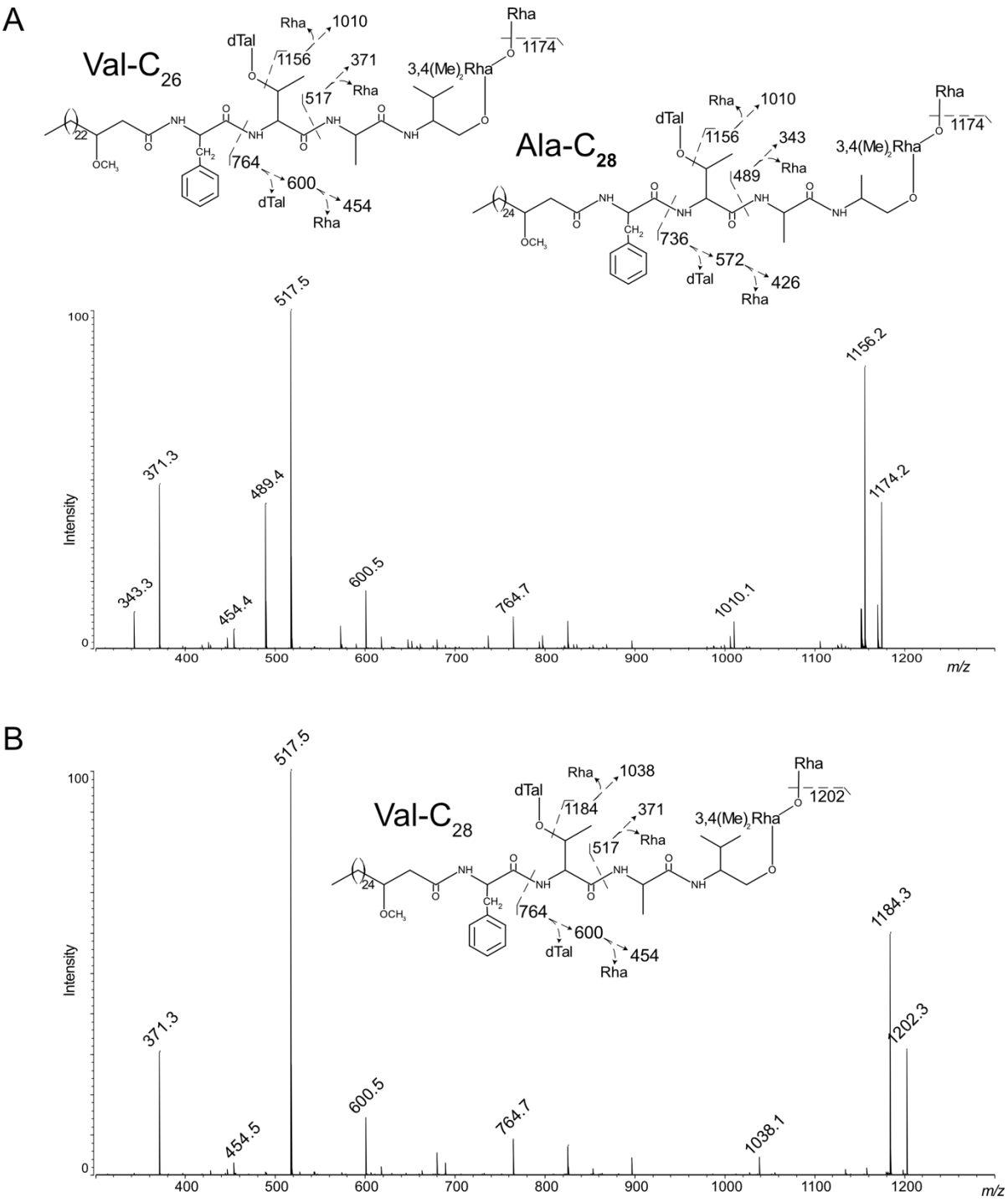


Figure S6: (A) MALDI-MS2 spectrum in positive mode of the selected parent ion at *m/z* 1320 of F1 purified fraction from Δ1,2,1725c,3448. Fragmentation patterns of two GPL-3 isomers Ala-C_{28} and Val-C_{26}. **(B)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1348 from Δ1,2,1725c,3448 mutant strain. Fragmentation pattern of GPL-3 Val-C28 is shown.

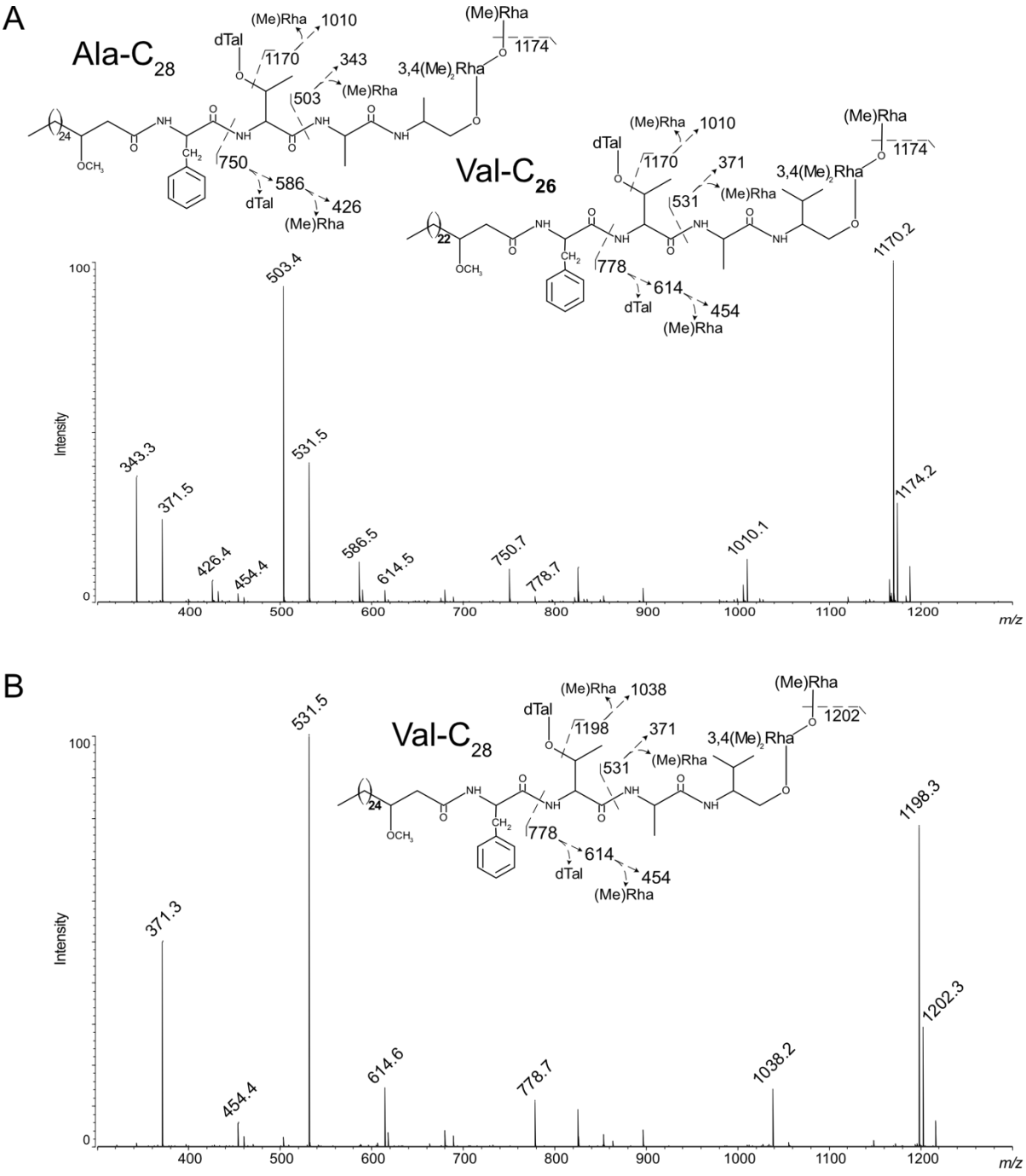


Figure S7: (A) MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1334 of F2 purified fraction from Δ1,2,1725c,3448. Fragmentation patterns of two GPL-3 isomers Ala-C_{28} and Val-C_{26} are shown **(B)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1362 of F2 purified fraction from Δ1,2,1725c,3448. Fragmentation pattern of GPL-3 Val-C_{28} is shown.

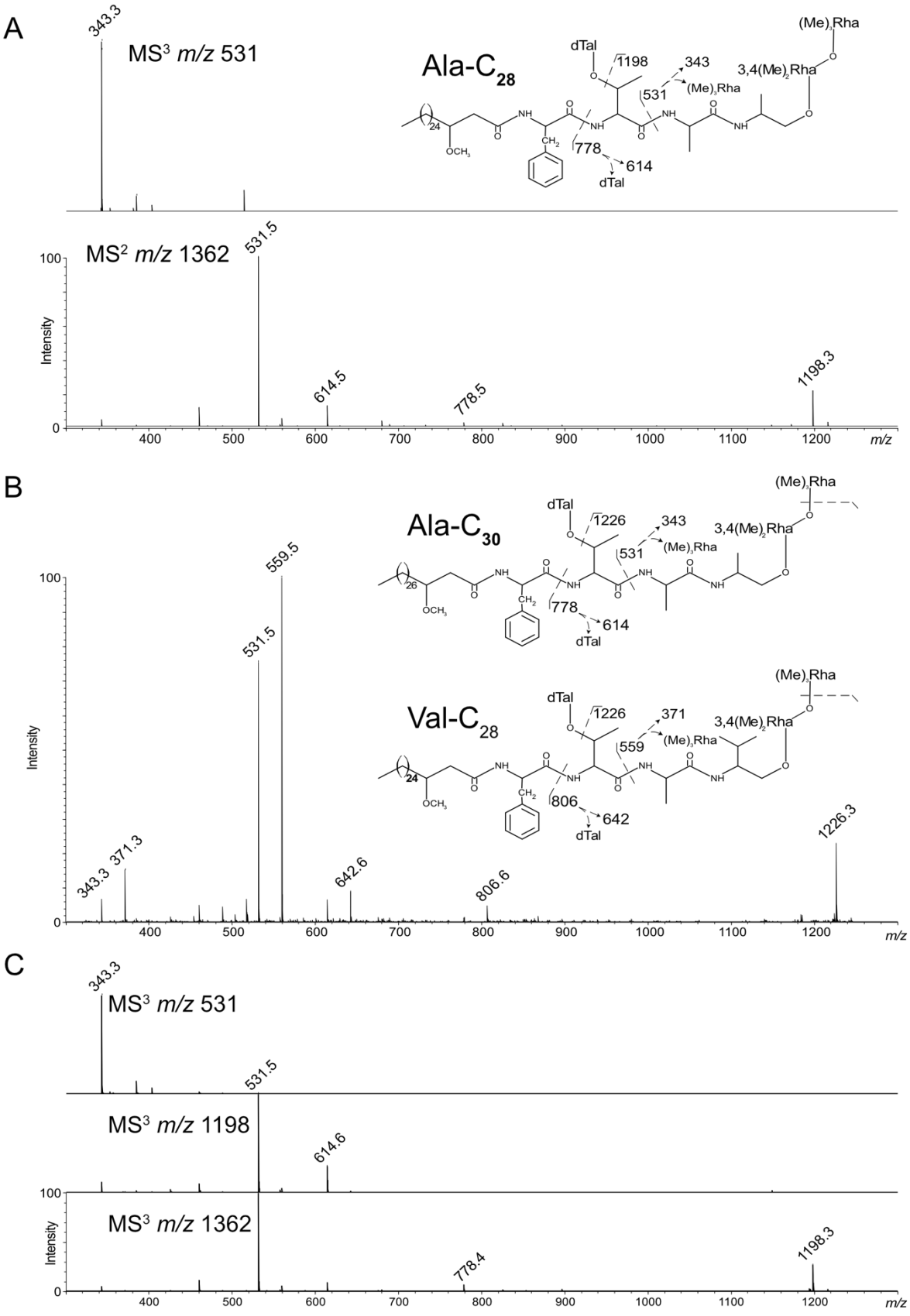


Figure S8: (A) MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1362, superimposed with MS^3 of fragment ion at *m/z* 531 of F3 purified fraction from Δ1,2,1725c,3448. Fragmentation pattern of GPL-3 Ala-C_{28} is shown. **(B)** MALDI-

MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1390 of F3 purified fraction from Δ1,2,1725c,3448. Fragmentation patterns of two GPL-3 isomers Ala-C_{30} and Val-C_{28} are shown. **(C)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1362, superimposed with MS^3 of fragment ions at *m/z* 1198 and 531 of F4 purified fraction from Δ1,2,1725c. These two fractions exhibited similar spectra.

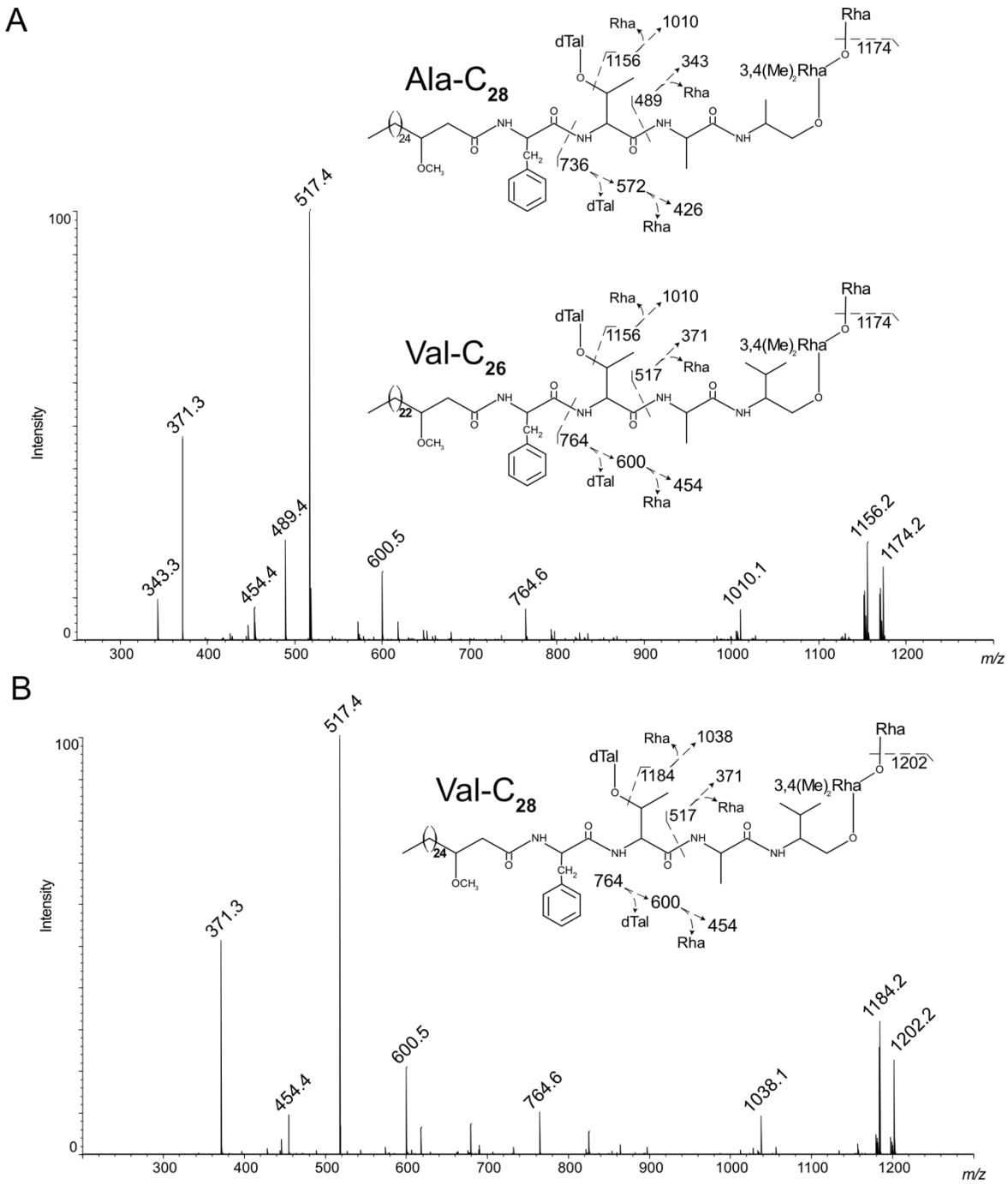


Figure S9: (A) MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1320 of F1 purified fraction from Δ1,2,1725c. Fragmentation patterns of two GPL-3 isomers Ala-C_{28} and Val-C_{26} are shown. **(B)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1348 of the same fraction. Fragmentation pattern of GPL-3 Val-C_{28} is shown.

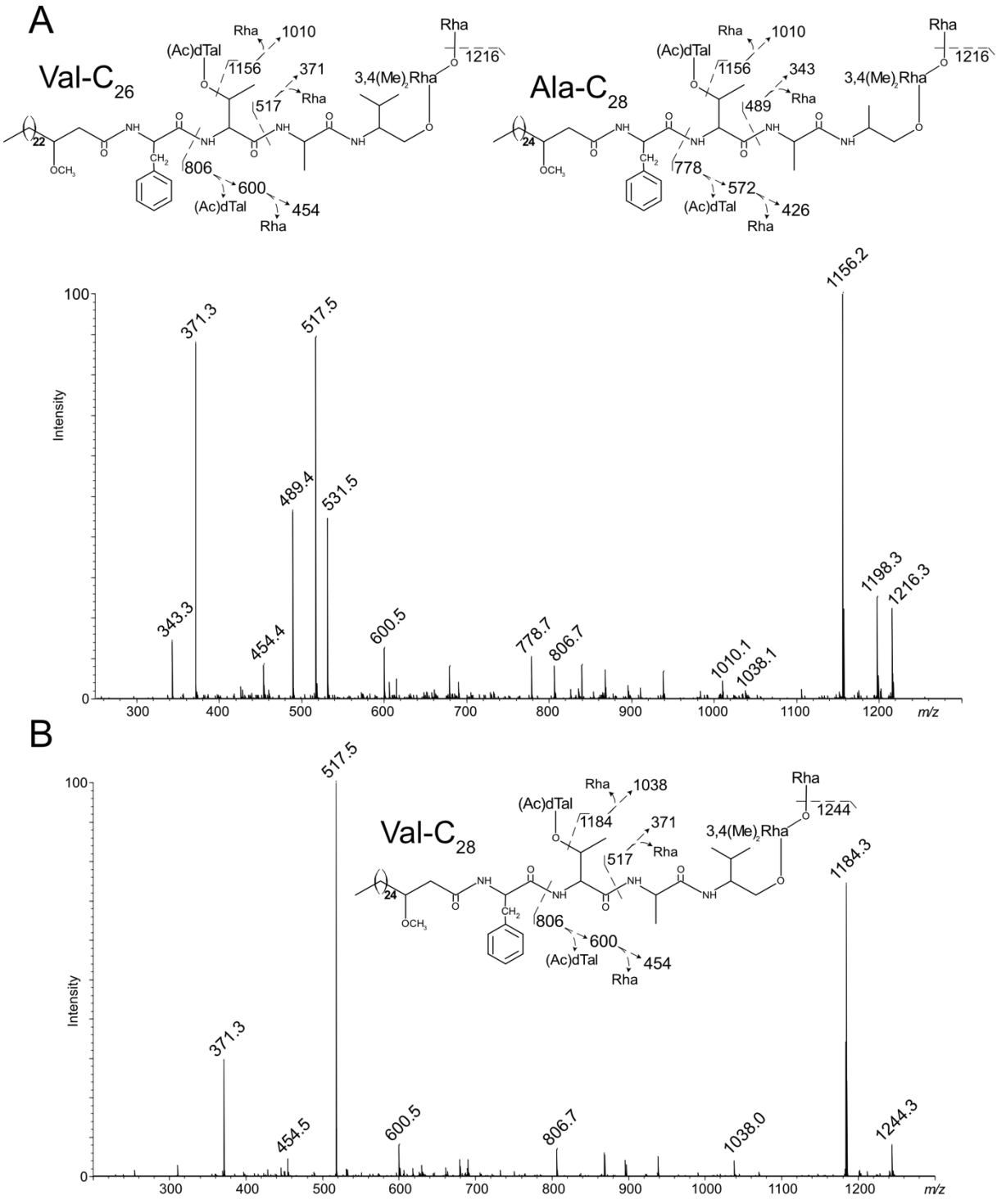


Figure S10: (A) MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1362 of F2 purified fraction from Δ1,2, 1725c. Fragmentation pattern of two GPL-3 isomers are shown. **(B)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1390 of the same fraction. Fragmentation pattern of mono-*O*-acetylated GPL-3 Val-C_{28}.

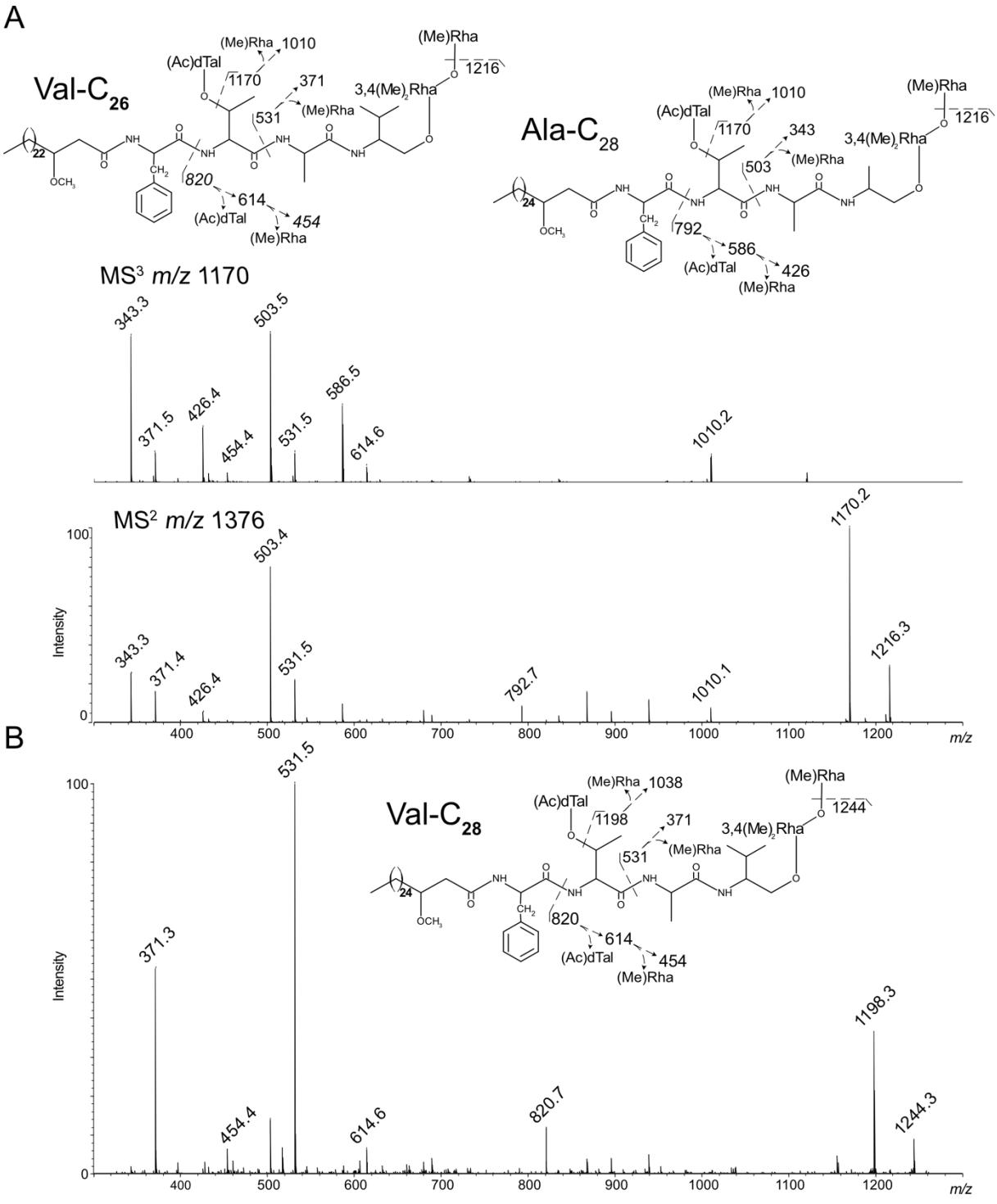


Figure S11: (A) MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1376 of F3 purified fraction from Δ1,2,1725c, superimposed with MALDI-MS^3 spectrum of fragment ion at *m/z* 1170, confirming the loss of mono-*O*-methyl Rha. Fragmentation patterns of two GPL-3 isomers Ala-C_{28} and Val-C_{26} are shown. **(B)** MALDI-MS^2 spectrum in positive mode of the selected parent ion at *m/z* 1404 of the same fraction. Fragmentation pattern of GPL-3 Val-C_{28} is shown.

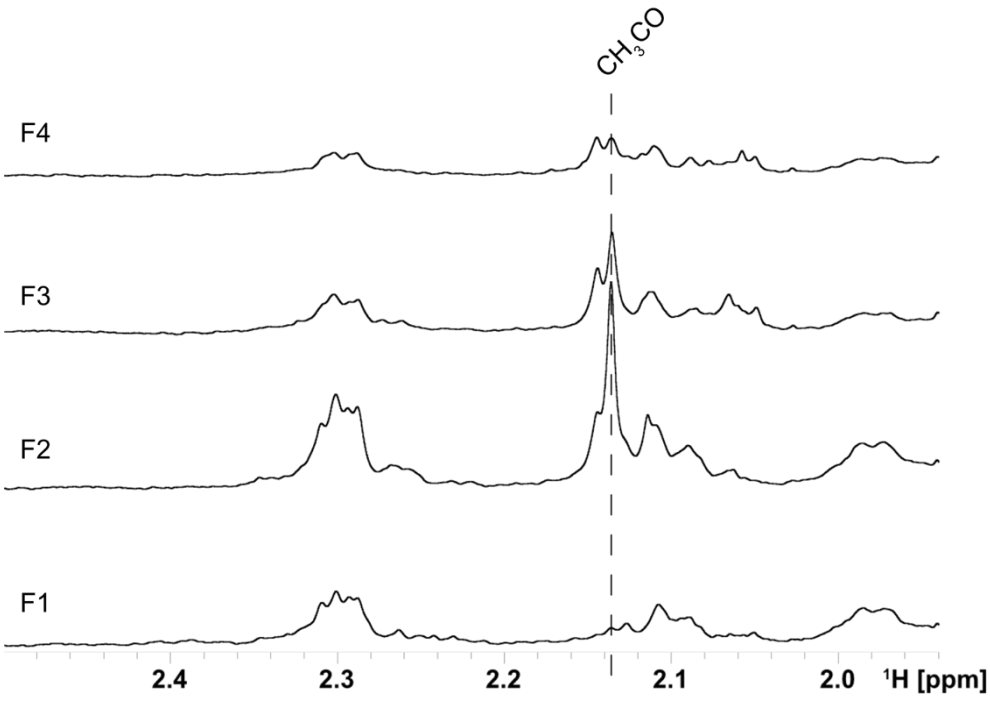


Figure S12: ^{1}H NMR spectra of F1 to F4 purified fractions from Δ1,2,1725c that show a clear signal at δ 2,14 ppm attributed to acetyl group in F2 and F3 fractions.

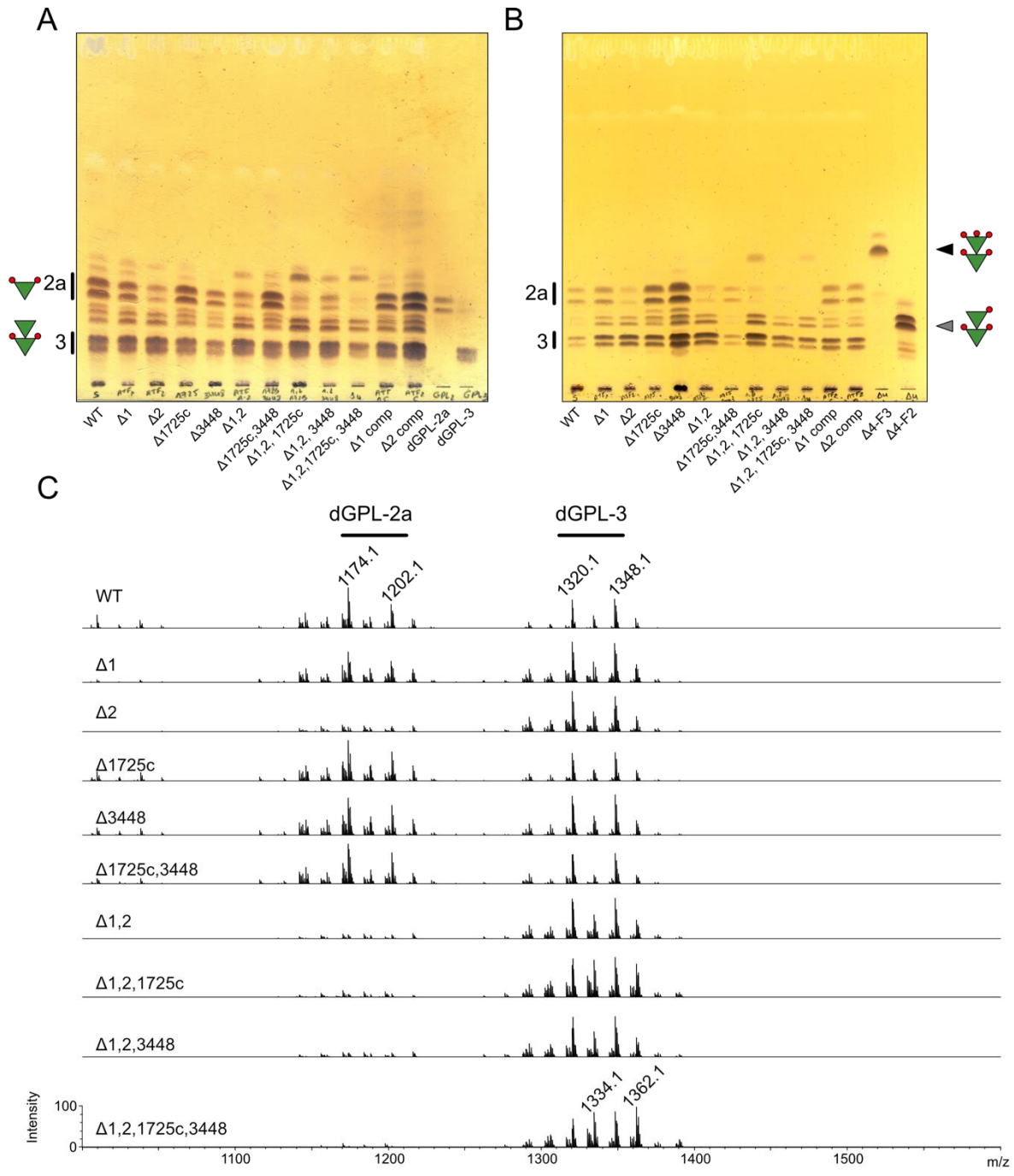


Figure S13: (A, B) TLC analysis of the saponified polar lipid fraction from WT and *atf* mutants. Butanol extracted GPLs were developed once with $CHCl_3$/MeOH/H_20 (90:10:1, v/v/v) and stained with orcinol. Purified GPL-2a and GPL-3 were submitted to the same treatment **(A)** as well as F2 and F3 fractions from Δ1,2, 1725c,3448 **(B)**. **(C)** MALDI-MS spectra in positive mode of the saponified GPL from WT and *atf* mutants. Previously characterized de-*O*-acetylated GPL-2a (dGPL-2a) and GPL-3 (dGPL-3) are shown on top of the spectra (1).

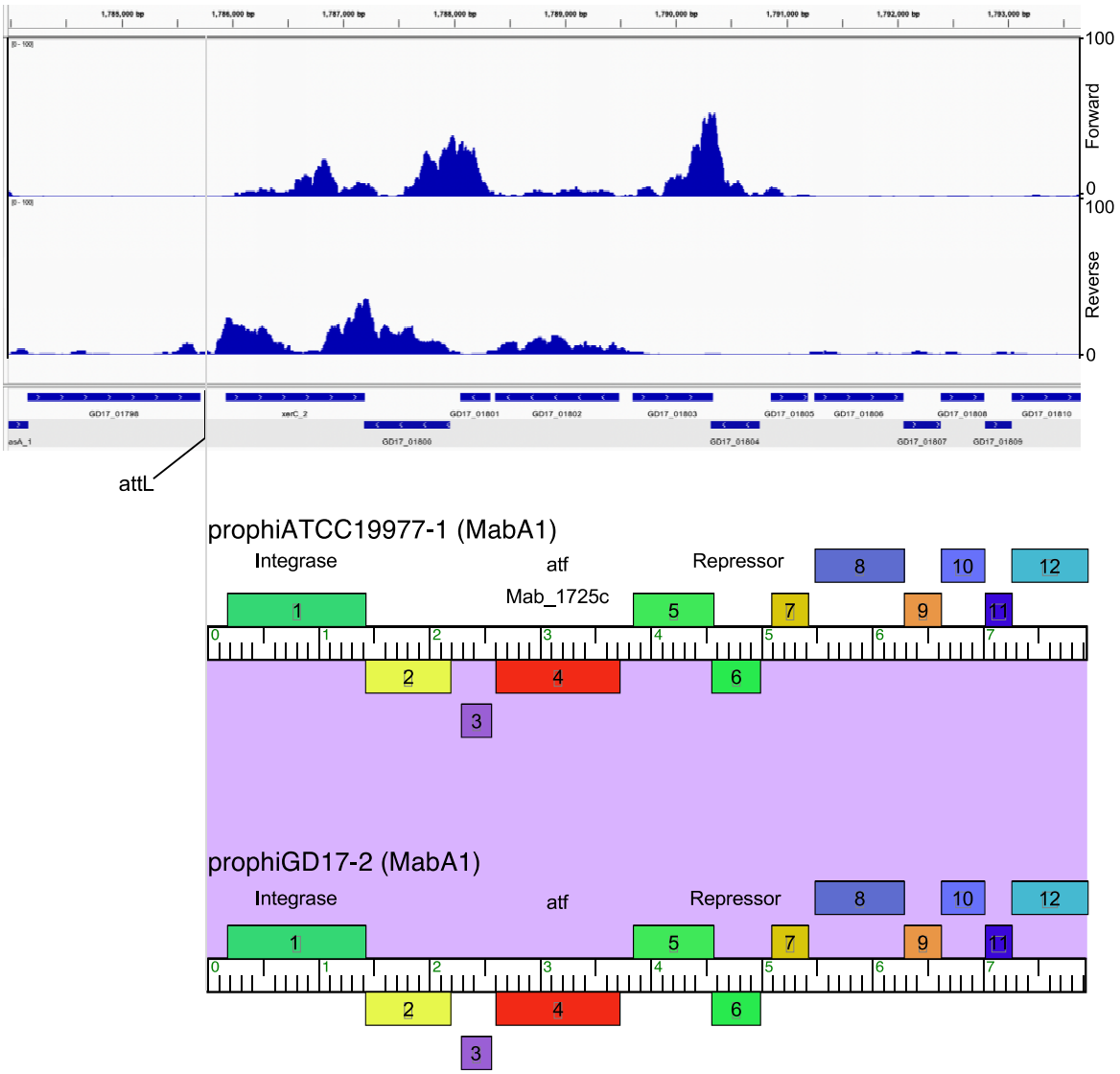


Figure S14: RNAseq analysis of prophage-encoded *atf* genes. Below is shown the left parts (~ 7 kbp) of the closely-related prophages prophiATCC19977-1 and prophiGD17-2, both members of Subcluster MabA1. The genes are shown as colored boxes above or below each genome maker, reflecting rightwards- and leftwards-transcription respectively. The purple shading between the genomes reflects BLASTN nucleotide sequence similarity, and the genomes are 100% identical in this segment. The genomes are aligned with the RNAseq data above for *M. abscessus* GD17, as reported previously (3), showing the numbers of reads mapping to the forward and reverse strands as indicated, from 0 to 100. The left ends of the prophages correspond to the left attachment site, *attL*, as indicated.

SUPPLEMENTARY TABLE

Table S1: Plasmids used in this study.

Plasmids			
pTEC27	*tdTomato* expressed under the control of a strong mycobacterial promoter by a multicopy *E. coli*/mycobacterial shuttle vector	Hyg	Addgene (plasmid 30182)
pUX1-*katG*	A pUX1 variant including *katG* gene of *M. tuberculosis* as a marker to counter-select in the presence of isoniazid and allowing to generate unmarked chromosomal alterations.	Kan, Hyg	(4)
pUX1-katG- *atf1*	pUX1-*katG* including the upstream and downstream sequences around *atf1*	Kan	This study
pUX1-katG- *atf2*	pUX1-*katG* including the upstream and downstream sequences around *atf2*	Kan	This study
pUX1-katG- *MAB-1725c*	pUX1-*katG* including the upstream and downstream sequences around *MAB-1725c*	Kan	This study
pUX1-katG- *MAB_3448*	pUX1-*katG* including the upstream and downstream sequences around *MAB-3448*	Kan	This study
pMV306	Integrative vector	Kan	(5)
pMV306-*atf1*	pMV306 enabling the expression of Atf1-HA in mycobacteria	Kan	This study
pMV306-*atf2*	pMV306 enabling the expression of Atf2-HA in mycobacteria	Kan	This study
pMV306-*MAB_1725c*	pMV306 enabling the expression of MAB-1725c-HA in mycobacteria	Kan	This study
pMV306-*MAB_3448*	pMV306 enabling the expression of MAB-3448-HA in mycobacteria	Kan	This study

Hyg, hygromycin; Kan, kanamycin

Table S2: Primers used in this study.

Number		5'→3' sequence	Restriction site	F (forward), R (reverse)
Cloning in pUX1-katG				
1	Atf1 U (F)	GAGA**TTAATTAA**AACACCACGCTGATGTCTGCGG	PacI	F
2	Atf1 U (R)	GAGA**CAATTG**GAGTGTCAAACCCTTAGCCTGCC	MfeI	R
3	Atf1 D (F)	GAGA**CAATTG**CCGATTCCTCTTGATCGCGATGCT	MfeI	F
4	Atf1 D (R)	GAGA**GCTAGC**GGATAGTGTCACCAAGGCCGCCGA	NheI	R
5	Atf2 U (F)	GAGA**TTAATTAA**CACAGCACAGTCGCTTTCGGAG	PacI	F
6	Atf2 U (R)	GAGA**CAATTG**GTTCAGCGCGTTGTTTCGTGAATCGA	MfeI	R
7	Atf2 D (F)	GAGA**CAATTG**CTGATCGAGAAACGCGCGCTG	MfeI	F
8	Atf2 D (R)	GAGA**GCTAGC**GAATCACGACCACCGAGTCGAAC	NheI	R
9	MAB_1725c U (F)	GAGA**TTAATTAA**ACTGCGATGGGAACCAGGTTCG	PacI	F
10	MAB_1725c U (R)	GAGA**CAATTG**GAGCGCGTTGTTGCGTGGGT	MfeI	R
11	MAB_1725c D (F)	GAGA**GAATTC**GAGAAACGCACTCTGGCGTTGAA	EcoRI	F
12	MAB_1725c D (R)	GAGA**GCTAGC**TAGTGAGCATGATTGCCGTGCGTTG	NheI	R
13	MAB_3448 U (F)	GAGA**TTAATTAA**ATGAGCTGATACCGAATCCACTGCTTG	PacI	F
14	MAB_3448 U (R)	GAGA**CAATTG**TGCGTTAAGTGCGTTGTTACGCGGA	MfeI	R
15	MAB_3448 D (F)	GAGA**CAATTG**GCACTCAAGAAACGGCTGCGC	MfeI	F
16	MAB_3448 D (R)	GAGA**TCTAGA**GATCGTTGTCATGCTGGTCTATTTCGGA	XbaI	R
Primers to verify double homologous recombination				
17	Atf1	GACGTGTATCGCAATTTCTGGACGTTCTTCT	-	F
18	Atf1	GAAATCAACTTAGCCGACTCAAGGTAGGTGCT	-	R
19	Atf2	GATTCATTGGATCCGCGCTGTCCAA	-	F
20	Atf2	CTGCATTTTCGGTGGAATCTGTCAAAAAATCTTATG	-	R
21	MAB_1725c	AACATACGGCAAACATCTGGGCATTACAAG	-	F
22	MAB_1725c	TCTAACCTCAACGAGCTGCTGAGCTAA	-	R
23	MAB_3448	CGCTCGAAGCTCAATAGGCCAGTAA	-	F
24	MAB_3448	CTGCACTGGGCAGTGCTGTTGT	-	R
Cloning in pMV306-KanR-Hsp60 promoter				
25	Atf1	GAGA**CAATTG**CCATGACCAAGGCGGGTGCGACC	MfeI	F
26	Atf1	GAGA**GTTAAC**CTAAGCGTAATCTGGAACATCGTATGGGT ACGCAGCATCGCGATCAAGAGGA	HpaI	R
27	Atf2	GAGA**GAATTC**CCATGAAGCTCGGTTCTGTATTCGATTCAC GA	EcoRI	F
28	Atf2	GAGA**GTTAAC**CTAAGCGTAATCTGGAACATCGTATGGGT AACTGTCCTCATGCGGCGACAGT	HpaI	R

29	MAB_1725c	GAGA**GAATTC**CCGTGCTCGGACGCGTATTCGAC	EcoRI	F
30	MAB_1725c	GAGA**GTTAAC**CTAAGCGTAATCTGGAACATCGTATGGGT AGACGGCCACCACTTGCTTC	HpaI	R
31	MAB_3448	GAGA**GAATTC**CCGTGCTGGGACGGGTATTTGATCC	EcoRI	F
32	MAB_3448	GAGA**GTCGAC**CTAAGCGTAATCTGGAACATCGTATGGGT ACCCCGCACCCGCTTTGACGCGCA	SalI	R

Table S3: Strains used in this study.

Name	Description/genotype	Resistance	Reference
M. abscessus Smooth (S)	*M. abscessus* sensu stricto, strain CIP104536^T, S morphotype	-	Laboratoire de Référence des Mycobactéries
M. abscessus Rough (R)	*M. abscessus* sensu stricto, strain CIP104536^T, R morphotype	-	Laboratoire de Référence des Mycobactéries
Δ 1	Unmarked deletion of *atf1* (*MAB_4106c*) in *M. abscessus*	-	This study
Δ 2	Unmarked deletion of *atf2* (*MAB_4110c*) in *M. abscessus*	-	This study
Δ MAB_1725c	Unmarked deletion of *MAB_1725c* in *M. abscessus*	-	This study
Δ MAB_3448	Unmarked deletion of *MAB_3448* in *M. abscessus*	-	This study
Δ 1, 2	Unmarked deletion of *atf1*(*MAB_4106c*) and *atf2* (*MAB_4110c*) in *M. abscessus*	-	This study
Δ MAB_1725c, MAB_3448	Unmarked deletion of *MAB_1725c* and *MAB_3448* in *M. abscessus*	-	This study
Δ 1, 2, MAB_1725c	Unmarked deletion of *atf1*, *atf2* and *MAB_1725c* in *M. abscessus*	-	This study
Δ 1, 2, MAB_3448	Unmarked deletion of *atf1*, *atf2* and *MAB_3448* in *M. abscessus*	-	This study
Δ4	Unmarked deletion of *atf1*, *atf2*, *MAB_1725c* and *MAB_3448* in *M. abscessus*	-	This study
Δ 1 + pMV306-*atf1*	Δ*atf1* carrying pMV306-*atf1*	KAN	This study
Δ 2 + pMV306-*atf2*	Δ*atf2* carrying pMV306-*atf2*	KAN	This study
Δ MAB_1725c + pMV306-*MAB_1725c*	Δ*MAB_1725c* carrying pMV306-*MAB_1725c*	KAN	This study
Δ MAB_3448 + pMV306-*MAB_3448*	Δ*MAB_3448* carrying pMV306-*MAB_3448*	KAN	This study

ADDITIONAL REFERENCES

1. Daher, W., Leclercq, L.-D., Viljoen, A., Karam, J., Dufrêne, Y. F., Guérardel, Y., and Kremer, L. (2020) O-Methylation of the Glycopeptidolipid Acyl Chain Defines Surface Hydrophobicity of Mycobacterium abscessus and Macrophage Invasion. *ACS Infect Dis*. **6**, 2756–2770
2. Daher, W., Leclercq, L.-D., Johansen, M. D., Hamela, C., Karam, J., Trivelli, X., Nigou, J., Guérardel, Y., and Kremer, L. (2022) Glycopeptidolipid glycosylation controls surface properties and pathogenicity in Mycobacterium abscessus. *Cell Chem Biol*. **29**, 910–924.e7
3. Dedrick, R. M., Aull, H. G., Jacobs-Sera, D., Garlena, R. A., Russell, D. A., Smith, B. E., Mahalingam, V., Abad, L., Gauthier, C. H., and Hatfull, G. F. (2021) The Prophage and Plasmid Mobilome as a Likely Driver of Mycobacterium abscessus Diversity. *mBio*. 10.1128/mBio.03441-20
4. Viljoen, A., Gutiérrez, A. V., Dupont, C., Ghigo, E., and Kremer, L. (2018) A Simple and Rapid Gene Disruption Strategy in Mycobacterium abscessus: On the Design and Application of Glycopeptidolipid Mutants. *Front Cell Infect Microbiol*. **8**, 69
5. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., and Hatfull, G. F. (1991) New use of BCG for recombinant vaccines. *Nature*. **351**, 456–460

III.E.6 IDENTIFICATION D'UN NOUVEAU GPL CHEZ *Mabs* R

Les GPLs du morphotype lisse ainsi que les DATs et les PIMs communs aux deux morphotypes sont des glycolipides libres majoritaires aisément détectables sur CCM après la révélation à l'orcinol-sulfurique. En outre, les GPLs sont les composés produisant les ions les plus intenses sur le spectre MALDI-TOF d'un extrait polaire total de *Mabs* S, suivies des PIM_2. Ces derniers sont également observés en mode négatif. A l'inverse, les DATs se détectent mieux sous forme purifiée.

Au fil de nos recherches de nouveaux glycolipides responsables de la virulence de la forme rugueuse, nous avons observé par spectrométrie de masse, de manière reproductible, sur de nombreux extraits polaires deux signaux de rapport *m/z* 1391 et 1750. Ces derniers ne correspondaient à aucun composé décrit dans la littérature. C'est pourquoi, nous avons entrepris d'obtenir en quantité suffisante des fractions purifiées et enrichies en ces deux ions afin de déterminer l'originalité de structure des entités moléculaires correspondantes.

III.E.6.a Purification orthogonale du GPL4690

III.E.6.a.i Purification sur colonne de silice

A partir de l'extrait polaire total de la souche rugueuse cultivée sur milieu LB, j'ai réalisé une préparation par chromatographie-flash sur colonne de silice vierge au moyen d'un gradient croissant de méthanol dans le chloroforme. Seule la fraction 25 éluant autour de 10% de méthanol présente une bande de coloration orangée en plus d'un dépôt qui se décolore après CCM et révélation à l'orcinol-sulfurique (Figure 56A). D'autre part, les spectres MALDI-TOF des fractions 20, 25 et 30 présentaient l'ion de rapport *m/z* 1750 très intense (Figure 56B). Enfin, la comparaison des fractions 21 à 31 regroupées avec l'extrait polaire total de *Mabs* R confirme la présence d'une bande de R_f identique moins intense (Figure 56C). Les résultats des expériences de spectrométrie de masse FT-ICR qui suivent ont été obtenues sur ce type d'échantillon. Néanmoins, les analyses RMN préliminaires ayant révélé la présence majoritaire d'un glycérophospholipide, probablement de la cardiolipine, j'ai donc poursuivi la purification afin d'éliminer ce contaminant de l'échantillon.

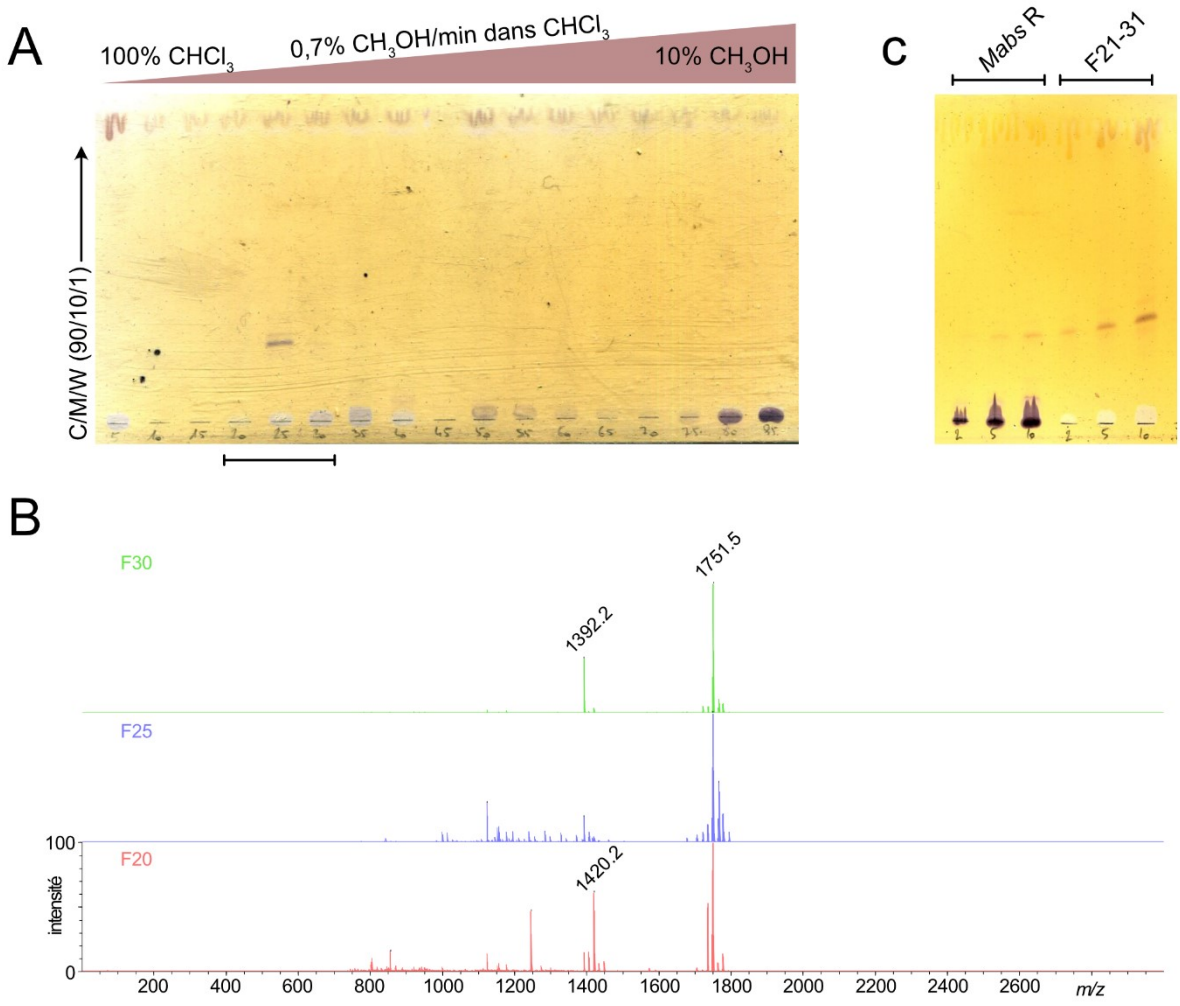


Figure 56 : 1ère purification du GPL4690. Une séparation sur colonne de silice met en évidence la fraction 25 (A) révélée par l'orcinol-sulfurique. Cette même fraction révèle l'ion *m/z* 1750 sur le spectre MALDI-TOF ainsi que dans les fractions 20 à 30 (B). Ces dernières ont été rassemblées puis comparées à l'extrait polaire total dont il est issu par CCM (C). La zone des dépôts de 2,5 et 10µL est décoloré en blanc.

III.E.6.a.ii Purification sur colonne de silice greffée C_{18}

Pour éliminer le glycérophopsholipide, j'ai procédé à une seconde purification des fractions 21 à 31 sur colonne de silice greffée C_{18} au moyen d'un gradient croissant de méthanol dans l'eau à 0,1% de TFA. Les fractions 60, 65 et 70 éluant à 100% de méthanol ont été analysées par MALDI-TOF afin de repérer les ions de rapport *m/z* 1391 et 1750 (Figure 57A). Ce dernier était le plus intense dans la fraction 65. Une révélation à l'orcinol-sulfurique après CCM des fractions 58 à 68 laissent apparaitre deux bandes orangées dont une particulièrement intense pour les fractions 63 et 64 (Figure 57B). Finalement, chaque fraction positive à la CCM et à l'ion 1750 par MALDI-TOF a été analysée individuellement par RMN monodimensionnelle ^1H (Figure 57C), et les fractions 64 et 65 possédant les

spectres parfaitement superposables ont été rassemblées pour atteindre une quantité et une pureté suffisante à l'analyse RMN multidimensionnelle qui suit.

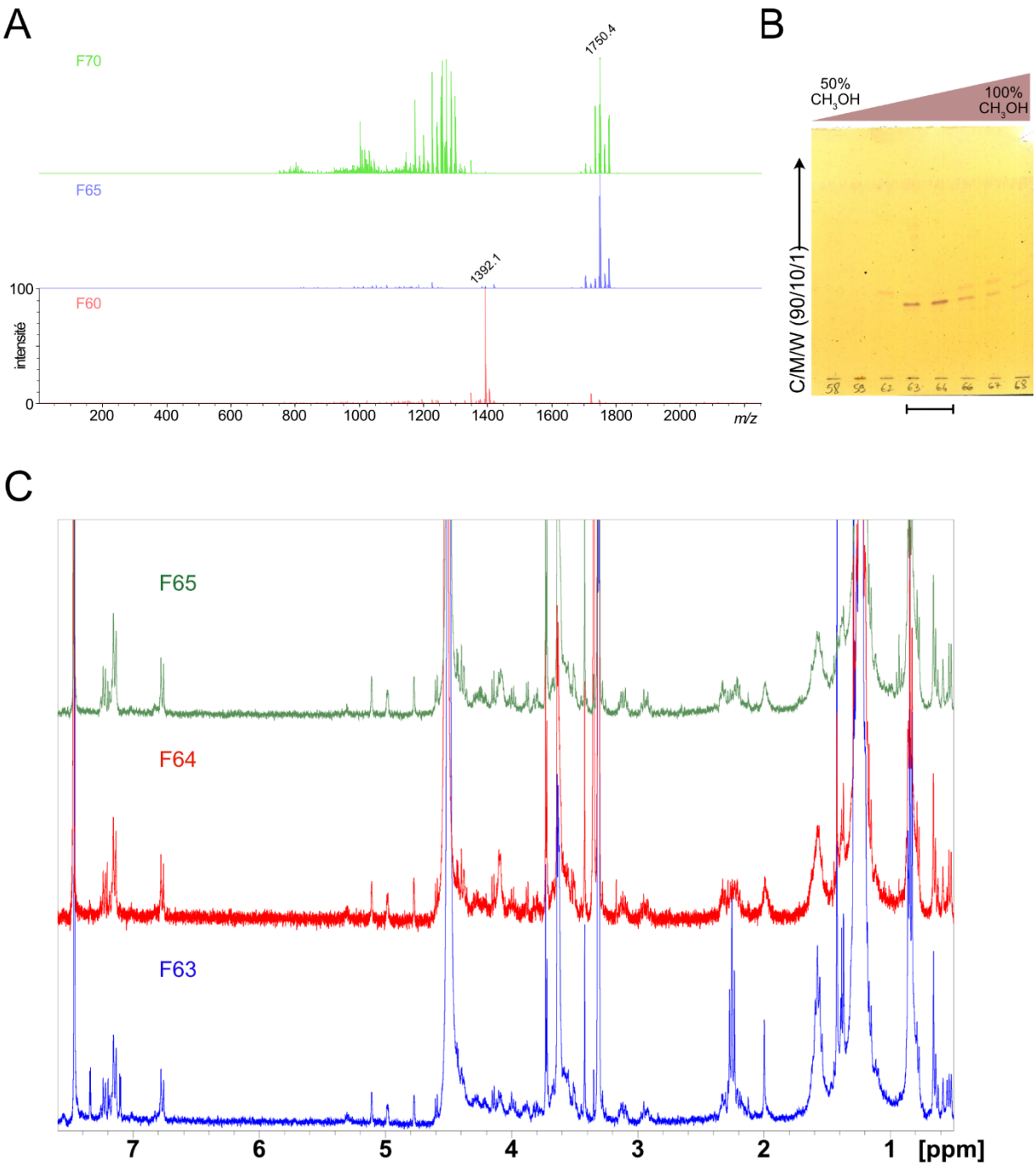


Figure 57 : 2ème purification du GPL4690. La purification de la fraction précédente F21-31 sur colonne de silice greffée C_{18} permet l'observation de l'ion *m/z* 1750 seul dans la fraction 65 (A) alors que l'ion *m/z* 1391 est identifié dans la fraction 60. Le contrôle CCM indique des bandes orangées intenses dans les fractions 63, 64, 66 et 67 encadrant la fraction 65 (B). Les spectres RMN ^1H révèlent que les fractions 64 et 65 sont plus pures que la 63 et parfaitement superposables.

III.E.6.b Analyse FT-ICR-MS (Fourier Transform Ion Cyclotron Resonance-Mass spectrometry)

Afin d'évaluer une formule brute exacte pour le composé correspondant à l'ion de rapport *m/z* 1750, des spectres de masse ont été acquis sur un système FT-ICR via les modes d'ionisation MALDI et électrospray à la plateforme COBRA de Rouen par le Dr LAVANANT et le Dr SCHMITT-ALONSO. Les six spectres acquis ont permis de déterminer une masse monoisotopique exacte moyenne de 1750,065792 Da avec un écart type de 2,23 10^{-4} Da pour l'adduit $[M+Na]^+$. A partir de cette valeur, j'ai établi une liste de 8 formules brutes avec une erreur de mesure maximale estimée à 1 ppm *via* une requête à partir de la formule brute $C_{0\text{-}100}H_{0\text{-}200}N_{0\text{-}20}O_{0\text{-}30}$ sur l'outil en ligne https://ms.epfl.ch/applications/theoretical-calculations/ :

Tableau 18 : Formule brute du GPL4690. Prédictions des formules brutes possibles pour la masse monoisotopique exacte 1750,065792 Da mesurée par ionisation MALDI et analyse FT-ICR d'une fraction uniquement purifiée sur silice. La formule soulignée est celle retenue après analyse RMN.

Formula	Exact mass	Sodium adduct	Insaturations	Δ (mDa)	Δ (ppm)
C75H146N20O25	1727.0768	1750.0660	13	-0.229	-0.13
C90H150N8O24	**1727.0763**	**1750.0655**	**20**	**0.274**	**0.16**
C89H25N11O30	1727.0769	1750.0661	83	-0.307	-0.18
C91H146N12O20	1727.0776	1750.0669	25	-1.064	-0.61
C74H150N16O29	1727.0755	1750.0647	8	1.108	0.63
C89H154N4O28	1727.0750	1750.0642	15	1.611	0.92
C87H142N18O18	1727.0749	1750.0642	26	1.621	0.93
C90H21N15O26	1727.0782	1750.0674	88	-1.645	-0.94

Parmi les formules proposées (Tableau 18), deux formules possèdent plus de quatre vingt insaturations, les rendant fortement improbables. La formule brute $C_{90}H_{150}N_8O_{24}$ possède une erreur entre la valeur expérimentale et la valeur théorique inférieure à 0,13 ppm. De plus, le profil isotopique théorique de $C_{90}H_{150}N_8O_{24}Na^+$ est similaire à plus de 97% avec les profils expérimentaux des six spectres (Figure 58A).

Tableau 19 : Formule brute du dérivé du GPL4690. Prédictions des formules brutes possibles pour la masse monoisotopique exacte 1391,830005 Da mesurée par ionisation MALDI et analyse FT-ICR d'une fraction uniquement purifiée sur silice. La formule soulignée est celle retenue après analyse MALDI-TOF.

Formula	Exact mass	Sodium adduct	Insaturations	Δ (mDa)	Δ (ppm)
C71H116N8O18	**1368.8408**	**1391.8300**	**18**	**0.025**	**0.02**
C100H108N2O2	1368.8411	1391.8303	48	-0.296	-0.21
C58H124N6O29	1368.8413	1391.8305	0	-0.488	-0.35
C87H116O13	1368.8416	1391.8308	30	-0.809	-0.58
C70H120N4O22	1368.8394	1391.8286	13	1.363	0.98

Un second adduit $[M+Na]^+$ a été mesuré plus intensément dans la même fraction à une masse exacte moyenne de 1391,830005 Da avec un écart type de 1,25 10^{-4} Da. A partir de la même requête que précédemment, la formule brute proposée $C_{71}H_{116}N_8O_{18}$ est la plus probable parmi les 5 possibles. Elle présente une erreur de 0,02 ppm la plus faible entre la valeur expérimentale et la valeur théorique. De plus, il est remarquable que le nombre d'atomes d'azote soit identique avec la formule brute précédente, $C_{90}H_{150}N_8O_{24}$. La similarité du profil isotopique est de nouveau supérieure à 97%. D'autres ions de masses variant de 1363,799035 Da à 1433,876968 Da ont été observés et interprétés comme des variations du nombre de groupements méthylènes accompagnant l'ion 1391,830005 Da (Figure 58B).

L'utilisation de l'ionisation électrospray associé à une calibration externe aboutit sur la même fraction à observer un ion intense de rapport *m/z* 886,527506 Da qui est un adduit $[M+2Na]^{2+}$ comme le montre le décalage de masse d'environ 0,5 Da sur le profil isotopique résolu (Figure 58C). Une nouvelle requête produit la même liste de résultats que le tableau 18, affichant en seconde position la formule brute la plus probable $C_{90}H_{150}N_8O_{24}$ avec une erreur inférieure à 1 ppm. Celle-ci a été confirmée par l'analyse RMN qui suit.

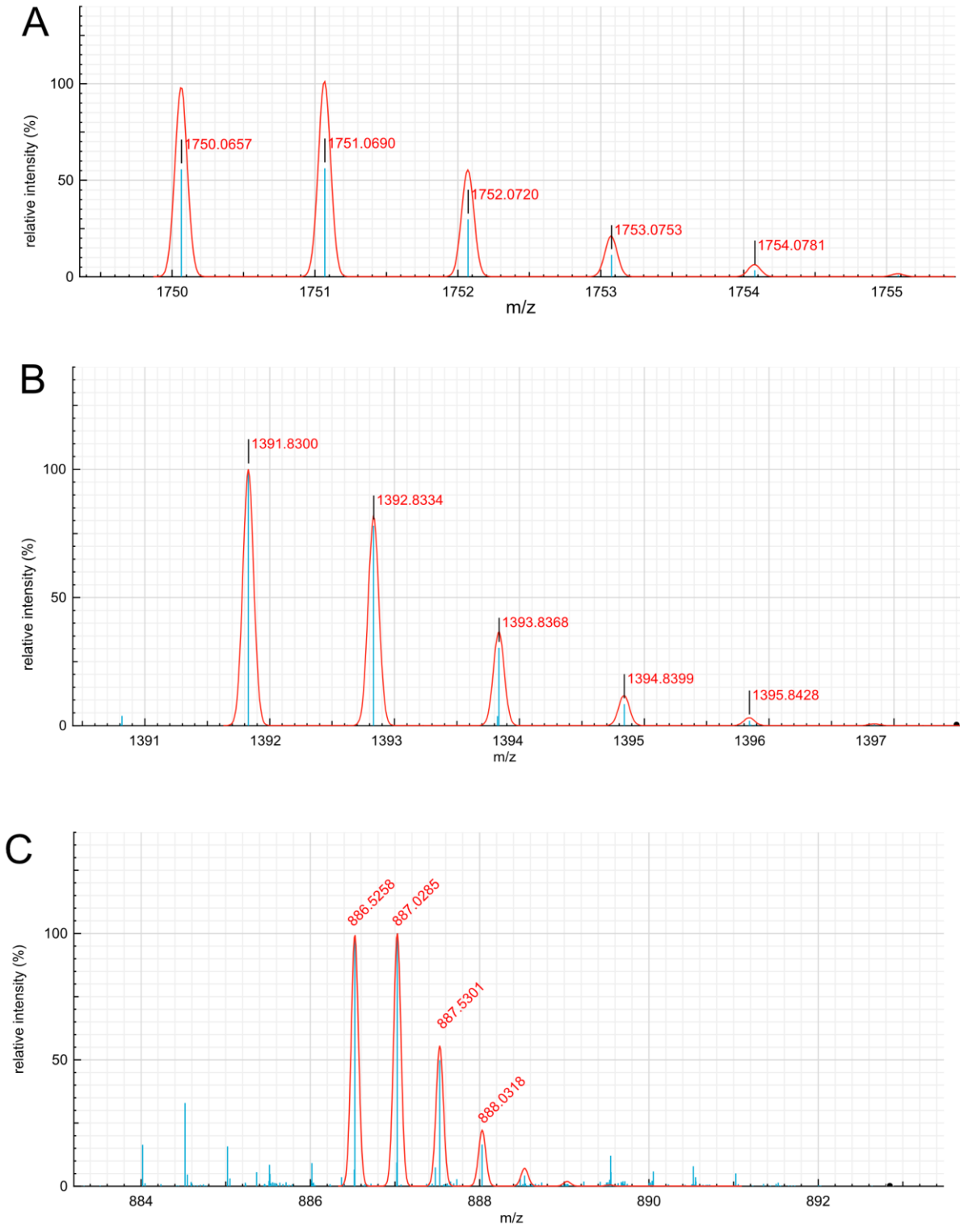


Figure 58 : Analyse FT-ICR du GPL4690 et de son dérivé. Les spectres (A) et (B) de la fraction uniquement purifié sur silice contenant le GPL4690 et son dérivé résultent de l'ionisation MALDI et le spectre (C) a été mesuré en électrospray. Sur chaque spectre, le tracé bleu est la mesure expérimentale alors que le tracé en rouge correspond au profil isotopique théorique de la formule brute sélectionnée et soulignée dans les tableaux 18 et 19.

III.E.6.c Analyse RMN du GPL4690

III.E.6.c.i Spectres dans les solvants deutérés

Identification des acides aminés

Les spectres $^{1}H/^{13}C$ HSQC et $^{1}H/^{1}H$ COSY n'affichent plus de signaux relatifs à un glycérophospholipide majoritaire, ce qui confirme que les produits ont atteint un niveau de pureté suffisant à leur analyse (Figure 59). Celle-ci a permis d'identifier huit acides aminés (Figure 59A) : trois résidus de thréonine, un d'alanine, un de phénylalanine, un d'isoleucine, un de tyrosine et un de leucine dont l'ensemble des paramètres RMN a été reporté dans le tableau 20. Les trois signaux $^{1}H/^{13}C$ à 4,15/59,82 ppm, 4,43/58,74 ppm et 4,6/57,35 appartiennent à des résidus de thréonine dont on observe la corrélation avec le proton de la position β, lui-même corrélé avec le proton de la position γ sur le spectre $^{1}H/^{1}H$ COSY (Figure 59B). Les trois groupements méthyles ont été respectivement identifiés à 1,218/19,85 ppm, 1,154/19,53 ppm et 1,185/15,55 ppm. Les positions α et β d'un résidu d'alanine dont les protons corrèlent entre eux sur le spectre $^{1}H/^{1}H$ COSY sont observés à 4,376/49,29 ppm et 1,238/16,85 ppm sur le spectre $^{1}H/^{13}C$ HSQC.

Les positions α quasiment confondus d'une phénylalanine et d'une tyrosine à 4,387/56,16 ppm et 4,373/56,38 ppm portent chacune un proton qui est corrélé sur le spectre $^{1}H/^{1}H$ COSY avec deux protons non équivalents du groupement méthylène β. Sur le spectre $^{1}H/^{13}C$ HSQC, les quatre signaux sont observés respectivement à 3,119-2,944/37,14 ppm et 3,093-2,924/35,96 ppm (Figure 59A).

De plus, la présence de signaux dans la zone aromatique confirme l'identification de ces résidus. Les positions δ, ε et ζ de la phénylalanine sont observés à 7,15/129,54 ppm, 7,23/128,89 ppm et 7,17/127,31 ppm ce dernier étant moins intense que les deux autres (Figure 60A). La position ε corrèle effectivement sur le spectre $^{1}H/^{1}H$ COSY avec les deux autres positions (Figure 60B). Le carbone quaternaire de la position γ a été déterminé à ^{13}C à 136,9 ppm par $^{1}H/^{13}C$ HMBC.

La position ε de la tyrosine est davantage blindée à 6,77/114,19 ppm et elle est corrélée sur le spectre $^{1}H/^{1}H$ COSY avec la position δ à 7,14/130,48 ppm (Figure 60A-B). Enfin, un groupement méthoxy à 3,728/55,31 ppm présente une forte corrélation sur le spectre $^{1}H/^{13}C$ HMBC avec le carbone quaternaire ζ qui est mis en évidence par deux autres corrélations avec les positions précédemment décrites δ et ε (Figure 60C). De même, ce groupement possède une corrélation spatiale ^{1}H sur le spectre $^{1}H/^{1}H$ NOESY avec le proton de la position ε (Figure 60D). Ces éléments suggèrent que la fonction phénol du résidu de tyrosine est méthoxylée.

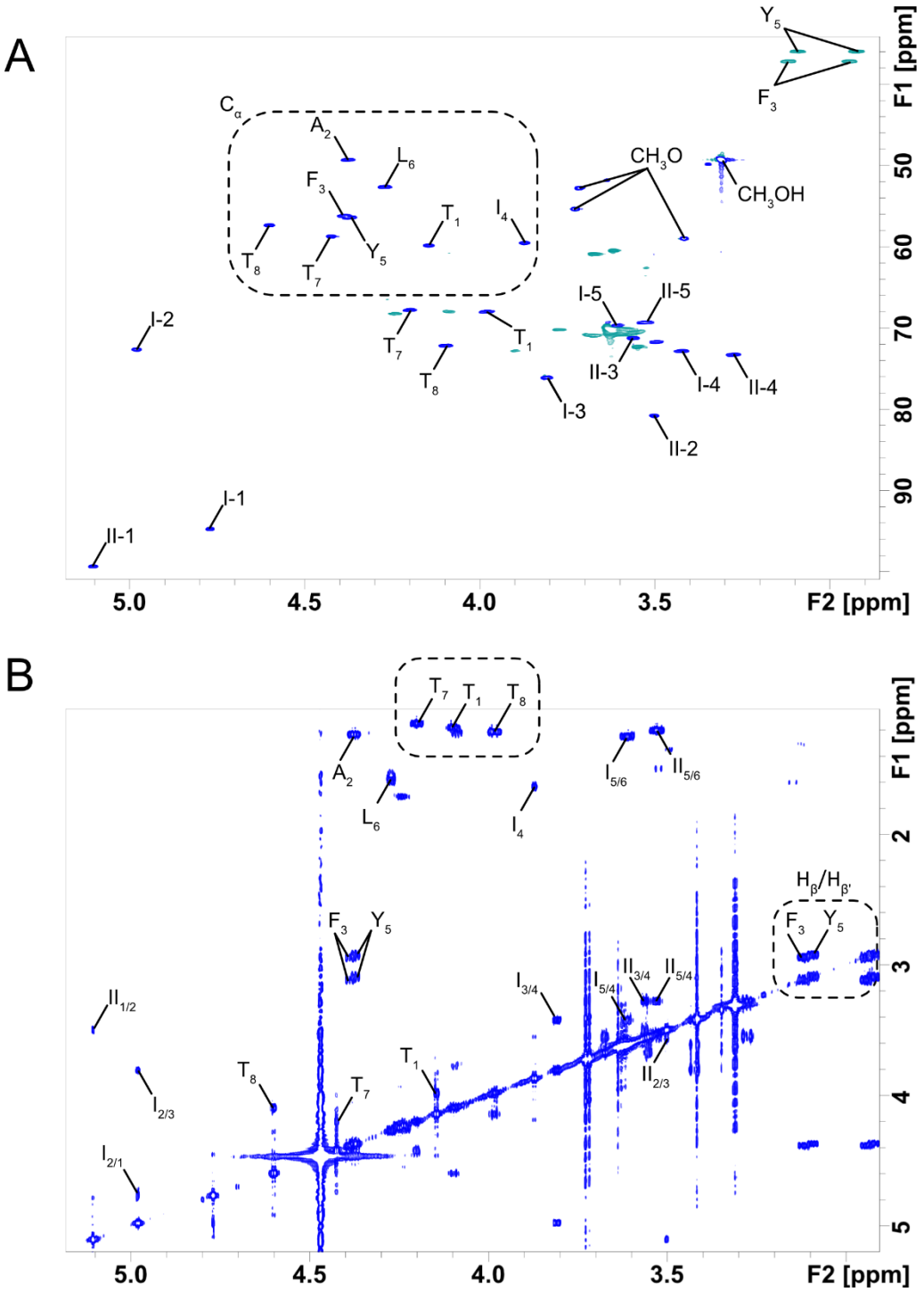


Figure 59 : Analyse RMN du GPL4690. Spectres des fractions 64-65 mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K, après calibration sur le méthanol. Les acides aminés sont indiqués dans l'ordre croissant avec le code à une lettre du N-terminal vers le C-terminal, et les positions α sont regroupées dans le cadre pointillé sur le spectre $^1H/^{13}C$ HSQC (A). Les trois signaux méthoxy appartiennent à la même molécule et seuls les signaux CH_2 de la phénylalanine et de la tyrosine ont été attribués. Sur le spectre $^1H/^1H$ COSY (B), les 2 zones en pointillés indiquent les corrélations H_β/H_γ des thréonines et $H_\beta/H_{\beta'}$ de la phénylalanine et de la tyrosine. Les corrélations des 2 deoxyhexoses I et II sont également affichées.

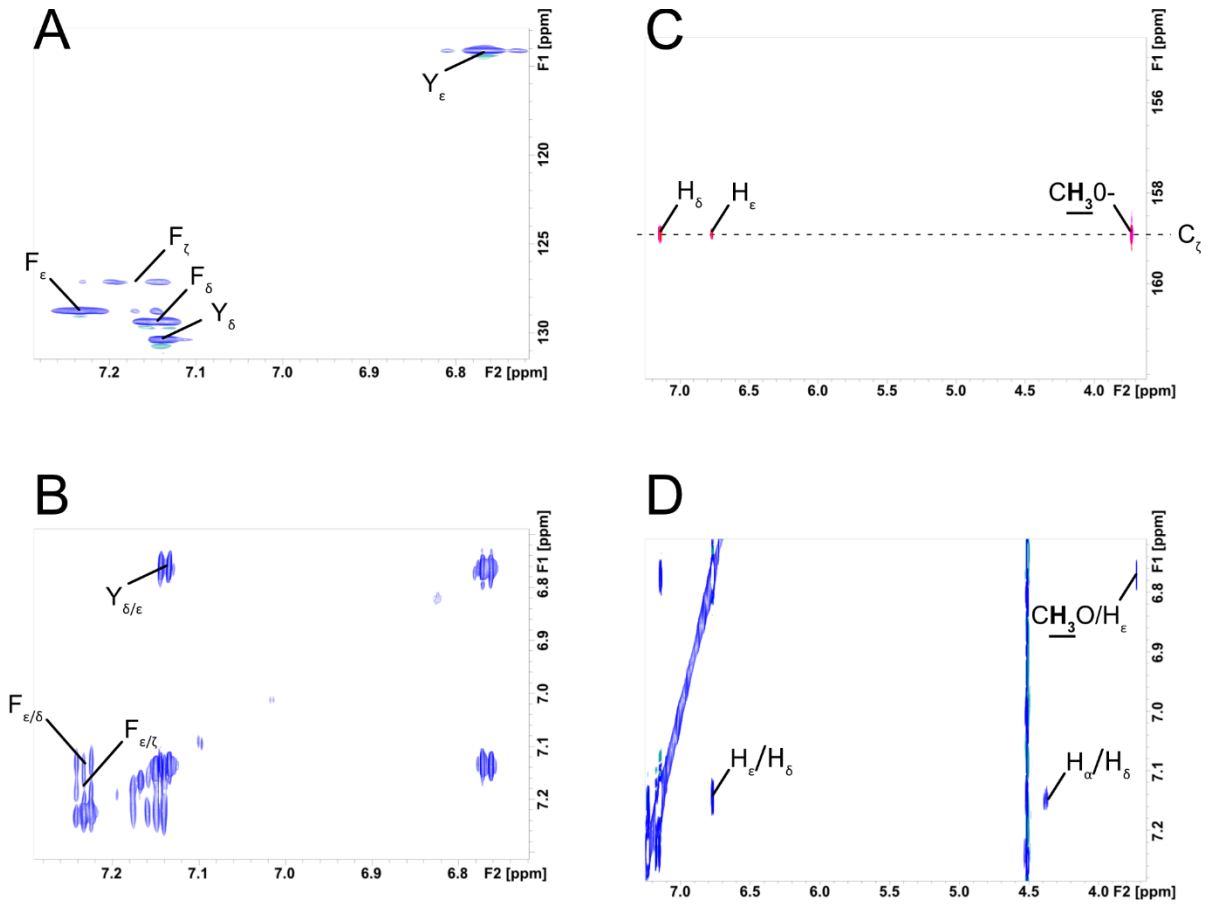


Figure 60 : Signaux aromatiques des acides aminés du GPL4690. Spectres des fractions 64-65 mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K, après calibration sur le méthanol. A gauche, les signaux et les corrélations des noyaux aromatiques de la phénylalanine et de la tyrosine sont visibles sur les zooms du spectre $^1H/^{13}C$ HSQC (A) et $^1H/^1H$ COSY (B) de la Figure 57. A droite, les spectres HMBC (C) et $^1H/^1H$ NOESY (D) montrent les corrélations des protons du groupement méthoxy avec le carbone ζ et le proton ε du noyau aromatique de la tyrosine justifiant la liaison de ce groupement.

La position α de la leucine est observée à 4,27/52,62 ppm, corrélée aux deux protons non équivalents du CH_2 à 1,54-1,58/40,48 ppm (Figure 61A & B). Le proton α de l'isoleucine est identifié à 3,87/59,5 ppm. Il est corrélé avec le CH de la position β à 1,63/35,86 ppm (Figure 61A). Ce dernier est lui-même couplé aux protons des positions γ_1 du CH_3 à 0,52/15,5 ppm d'une part, et γ_2 du CH_2 à 0,814-0,87/24,26 ppm d'autre part.

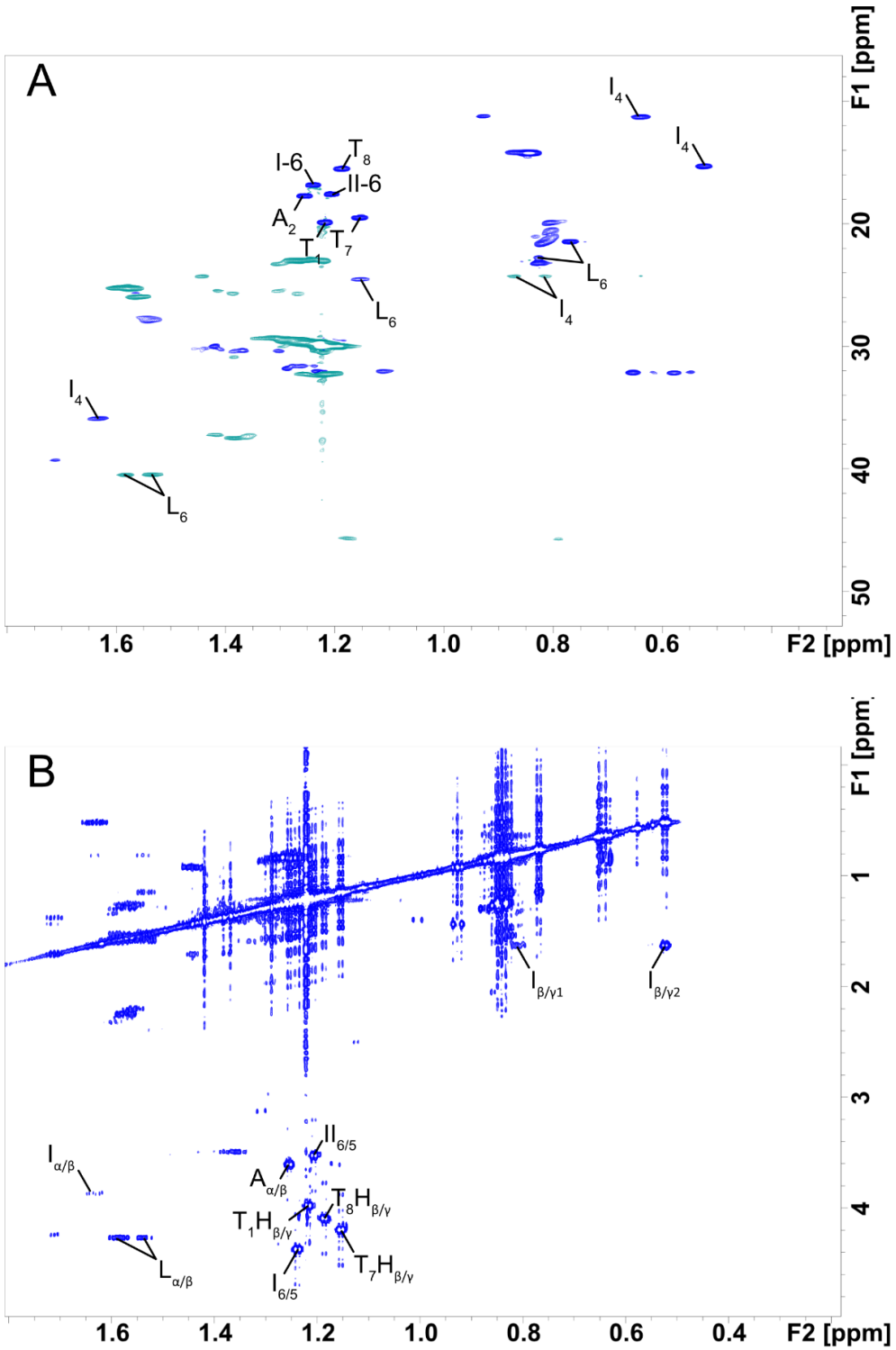


Figure 61 : Signaux aliphatiques des acides aminés du GPL4690. Spectres des fractions 64-65 mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K, après calibration sur le méthanol. Les signaux $^1H/^{13}C$ (A) ainsi que les corrélations protons (B) des groupements méthylènes en vert, et méthyles en bleu, sont visibles sur les zooms du spectre $^1H/^{13}C$ HSQC (A) et $^1H/^1H$ COSY (B) de la Figure 59. pour les acides aminés aliphatiques, les thréonines et les déoxyhexoses.

Tableau 20 : Signaux RMN des acides aminés du GPL4690. Déplacements chimiques ^{1}H et ^{13}C des 8 acides aminés identifiés dans les fractions 64-65. 2 groupements méthoxy de la dernière colonne sont couplés par HMBC avec le C_ζ de la tyrosine et le carbonyle de la thréonine (signaux soulignés).

Acides aminés	δ (ppm)	NH	CO	CH_α	CH_β	$CH_{\gamma1}$	$CH_{\gamma2}$	$CH_{\delta1}$	$CH_{\delta2}$	CH_ε	CH_ζ	CH_3-O
Thr-1	^{1}H	7,64	-	4,150 *(d;6)*	3,985 *(m)*	1,218 *(d;6,2)*	-	-	-	-	-	-
	^{13}C	-	171,95	59,82	67,99	19,85	-	-	-	-	-	-
Ala-2	^{1}H	8,19	-	4,376	1,238 *(d;7)*	-	-	-	-	-	-	-
	^{13}C	-	173,99	49,29	16,85	-	-	-	-	-	-	-
Phe-3	^{1}H	8,06	-	4,387	3,119/2,944	-	-	7,15	-	7,23 *(Dd;7,6)*	7,17	-
	^{13}C	-	173,61	56,16	37,14	136,9	-	129,54	-	128,89	127,31	-
Ile-4	^{1}H	7,46	-	3,870 *(d;6,2)*	1,63	0,870/0,814	0,52 *(d;6,9)*	0,64 *(t;7,4)*	-	-	-	-
	^{13}C	-	173,15	59,499	35,86	24,26	15,5	11,2	-	-	-	-
Tyr-5	^{1}H	7,85	-	4,373	3,093/2,924	-	-	7,14	-	6,77 *(d;8,7)*	-	3,728
	^{13}C	-	172,69	56,38	35,96	n.d.	-	130,48	-	114,19	158,89	55,31
Leu-6	^{1}H	7,83	-	4,27 *(Dd;4,5/10,7)*	1,58/1,154	1,153	-	0,77 *(d;6,5)*	0,825 *(d;6,7)*	-	-	-
	^{13}C	-	174,17	52,62	40,48	24,51	-	21,5	23,18	-	-	-
Thr-7	^{1}H	7,73	-	4,428 *(d;3,6)*	4,2 *(m)*	1,154 *(d;6,3)*	-	-	-	-	-	-
	^{13}C	-	171,3	58,74	67,77	19,53	-	-	-	-	-	-
Thr-8	^{1}H	8,03	-	4,6 *(d;5,8)*	4,1 *(m)*	1,185 *(d;6,1)*	-	-	-	-	-	3,717
	^{13}C	-	171,11	57,35	72,17	15,55	-	-	-	-	-	52,718

Séquençage du peptide

A partir du spectre $^{1}H/^{13}C$ HMBC, dans la zone de 171 à 176 ppm correspondant au déplacement chimique ^{13}C des carbones quaternaires des groupements carbonyles, plusieurs corrélations apparaissent notamment avec les protons portés par les positions α des acides aminés précédemment identifiés (Figure 62A). Le point de départ est un signal $^{1}H/^{13}C$ à 3,717/52,72 ppm assigné à une fonction méthoxy. Ses protons sont corrélés avec le carbone à 171,11 ppm, lui-même corrélé avec le proton β à 4,1 ppm et le proton α à 4,6 ppm de l'un des résidus de thréonine, noté T_8 (Figure 62A). A partir de ce côté C-terminal, il est possible d'observer une corrélation plus faible du proton α de T_8 avec le carbone à 171,3 ppm. Le proton α à 4,428 ppm d'un 2ème résidu de thréonine T_7 est également corrélé à ce carbone plus intensément, ainsi qu'a un 3ème carbone à 174,17 ppm. Ce dernier possède une corrélation plus intense avec le proton de la position α à 4,27 ppm de la leucine L_6 qui est aussi corrélé au 4ème carbone à 172,69 ppm. Celui-ci est ensuite corrélé avec le proton β à 3,093 ppm et le proton α à 4,373 ppm appartenant au résidu de tyrosine Y_5.

La corrélation suivante entre le proton α de Y_5 et le 5ème carbone à 173,15 ppm est beaucoup moins intense qu'avec le proton α de l'isoleucine I_4 à 3,87 ppm. Ce dernier est corrélé avec un 6ème carbone à 173,61 ppm qui l'est aussi avec les protons β et α de la phénylalanine F_3 à 3,119 ppm et 4,387 ppm. Malgré la proximité des corrélations, l'avant dernier carbone à 173,99 ppm permet de relier le proton α de F_3 avec celui de l'alanine A_2 à 4,376 ppm. Il corrèle enfin avec le dernier carbone à 171,95 ppm qui possède 2 corrélations supplémentaires avec les protons α et β du résidu de

thréonine T_1 côté N-terminal à 4,15 ppm et 3,985 ppm. Le proton α de T_1 étant corrélé avec un dernier carbone quaternaire à 175,29 ppm, il semble que la position N-terminal soit acylée. Comme le montre l'enchainement des corrélations entre les protons portés par les positions α et les carbones quaternaires n et n-1 de la chaine peptidique, la séquence déduite du spectre $^{1}H/^{13}C$ HMBC est donc Thr-Ala-Phe-Ile-(OMe)Tyr-Leu-Thr-Thr, notée $T_1A_2F_3I_4Y_5L_6T_7T_8$ (Figure 62B).

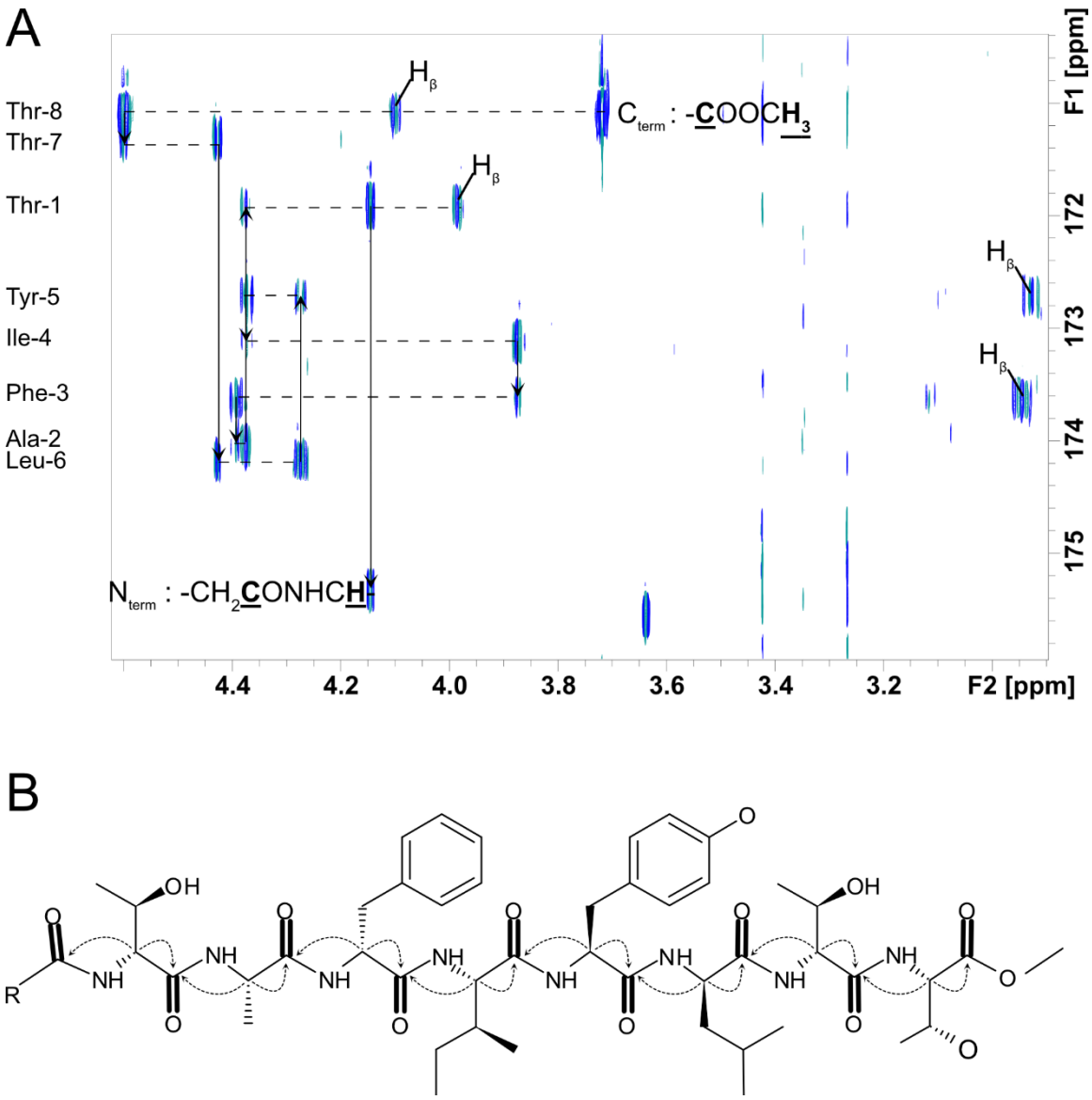


Figure 62 : Séquence peptidique du GPL4690 déterminée par RMN. Spectre $^{1}H/^{13}C$ HMBC des fractions 64-65 mesuré dans le $CDCl_3/CD_3OD$ (2:1) à 293K, après alignement sur le spectre HSQC. Les pointillées sont alignés avec les signaux des groupements carbonyles de chaque acide aminé, en partant du côté C-terminal vers le côté N-terminal. La corrélation de chaque proton α avec son groupement carbonyle et celui de l'acide aminé précédent conduit à la séquence peptidique $T_1A_2F_3I_4Y_5L_6T_7T_8$ (B) dans laquelle les carbonyles sont en gras.

Identification des monosaccharides

Sur le spectre $^{1}H/^{13}C$ HSQC, nous avons également détectés les signaux correspondant aux positions anomériques de deux monosaccharides. Le signal I-1 du premier monosaccharide à 4,77/94,73 ppm (Figure 59A) est corrélé sur le spectre $^{1}H/^{1}H$ COSY avec I-2 à 4,98/72,6 ppm (Figure 59B). Le fort déblindage de I-2 suggère que cette position est acylée. Le spectre $^{1}H/^{13}C$ HMBC montre d'ailleurs une corrélation commune avec la même fonction carbonyle entre le groupement CH_2 à 2,35-2,31/34,5 ppm et I-2. A partir de cette dernière, les corrélations $^{1}H/^{1}H$ COSY se succèdent avec I-3, I-4 et I-5 à 3,81/76,1, 3,42/72,82 et 3,612/69,7 ppm (Figure 59A & B). En outre, le proton I-5 étant corrélé avec ceux d'un groupement CH_3 à 1,255/17,7 ppm, nous en déduisons que ce monosaccharide est un déoxyhexose. Enfin, les constantes de couplages $^{3}J_{H1-H2}$ et $^{3}J_{H2-H3}$ inférieures à 4 Hz ainsi que $^{3}J_{H3-H4}$ et $^{3}J_{H4-H5}$ d'environ 10 Hz, démontrent une configuration manno, ce qui identifie ce premier déoxyhexose comme un rhamnose. Finalement, la valeur de 170 Hz de la constante de couplage $^{1}J_{C-H}$ de la position anomérique implique une configuration α.

Tableau 21 : Signaux RMN des rhamnoses du GPL4690. Déplacements chimiques et constantes de couplage attribués aux 6 positions des 2 rhamnoses Rha I et Rha II caractérisés dans les fractions 64-65 rassemblés.

Monosaccharides	δ (ppm)	Positions						Functions	
		1	2	3	4	5	6	$-CH_2-CO$	CH_3-O-
Rha-I	^{1}H	4,77	4,98	3,81	3,42	3,612	1,255	2,35/2,31	-
	m	*d*	*Dd*	*Dd*	*T*	*n.d.*	*d*	-	-
	$^{3}J_{Hn/Hn+1}$	*1,5*	*3,3*	*9,7*	*9,7*	*n.d.*	*6,1*	-	-
	^{13}C	94,73	72,6	76,1	72,82	69,7	17,7	34,5	-
	$^{1}J_{H/C}$	*170*	-	-	-	-	-	-	-
Rha-II	^{1}H	5,1	3,502	3,563	3,277	3,529	1,206	-	3,42
	m	*d*	*Dd*	*Dd*	*T*	*n.d.*	*d*	-	-
	$^{3}J_{Hn/Hn+1}$	*1,2*	*3,5*	*9,7*	*9,6*	*n.d.*	*6,1*	-	-
	^{13}C	99,35	80,77	71,21	73,25	69,38	17,6	-	58,96
	$^{1}J_{H/C}$	*169*	-	-	-	-	-	-	-

De la même manière, le second monosaccharide II dont la position anomérique II-1 est observée à 5,1/99,35 ppm et a été identifié comme un résidu α-rhamnopyranosyl. Néanmoins, la position II-2 ne semble pas acylée mais le fort déblindage du signal carbone à 3,502/80,77 ppm suggère une modification par un groupement méthoxy dont le signal à 3,42/58,96 ppm corrèle effectivement avec II-2 sur le spectre HMBC (Figure 63A).

Liaisons glycosidiques

Si l'on superpose le spectre HMBC avec le spectre HSQC, une corrélation réciproque apparait entre la position anomérique I-1 et la position β à 4,1/72,17 ppm du résidu de thréonine T_8 (Figure 63A). La même superposition met en évidence une 2ème corrélation réciproque entre la position II-1 du 2ème résidu α rhamnopyranosyl et la position I-3 du 1er.

Les spectres ^{1}H/^{1}H COSY et ^{1}H/^{1}H NOESY superposés révèlent des corrélations spatiales entre la position β à 4,1/72,17 ppm du résidu de thréonine T_8 et les positions I-1 et I-5 du 1er résidu α rhamnopyranosyl (Figure 63B). Les protons du groupement γ-méthyle du résidu de thréonine sont également corrélés avec la position I-1. D'autre part, la position anomérique II-1 du 2ème résidu α rhamnopyranosyl est corrélée non seulement avec la position I-3 mais aussi avec le groupement méthoxy à 3,42/58,96 ppm porté par la position II-2.

L'ensemble de ces observations implique que le motif glycannique 2-*O*-Me-α-rhamnopyranosyl(1-3)2-*O*-acyl-α-rhamnopyranose est relié au dernier résidu de thréonine T_8 par une liaison glycosidique avec son groupement β-hydroxyle.

Autres observations

Le spectre ^{1}H/^{13}C HSQC présente également les signaux caractéristiques avec les déplacements chimiques ^{1}H/^{13}C caractéristiques d'une fonction éthylénique à 5,296/30,09 ppm associé à son groupement méthylène à 1,971/27,46 ppm, d'une fonction alcool secondaire à 3,494/71,78 ppm associé à un groupement méthylène autour de 1,38 ppm, d'au moins 4 groupements CH_2 à 2,32-2,21/34,5-36,3 ppm en α de groupements carbonyles d'après le spectre HMBC.

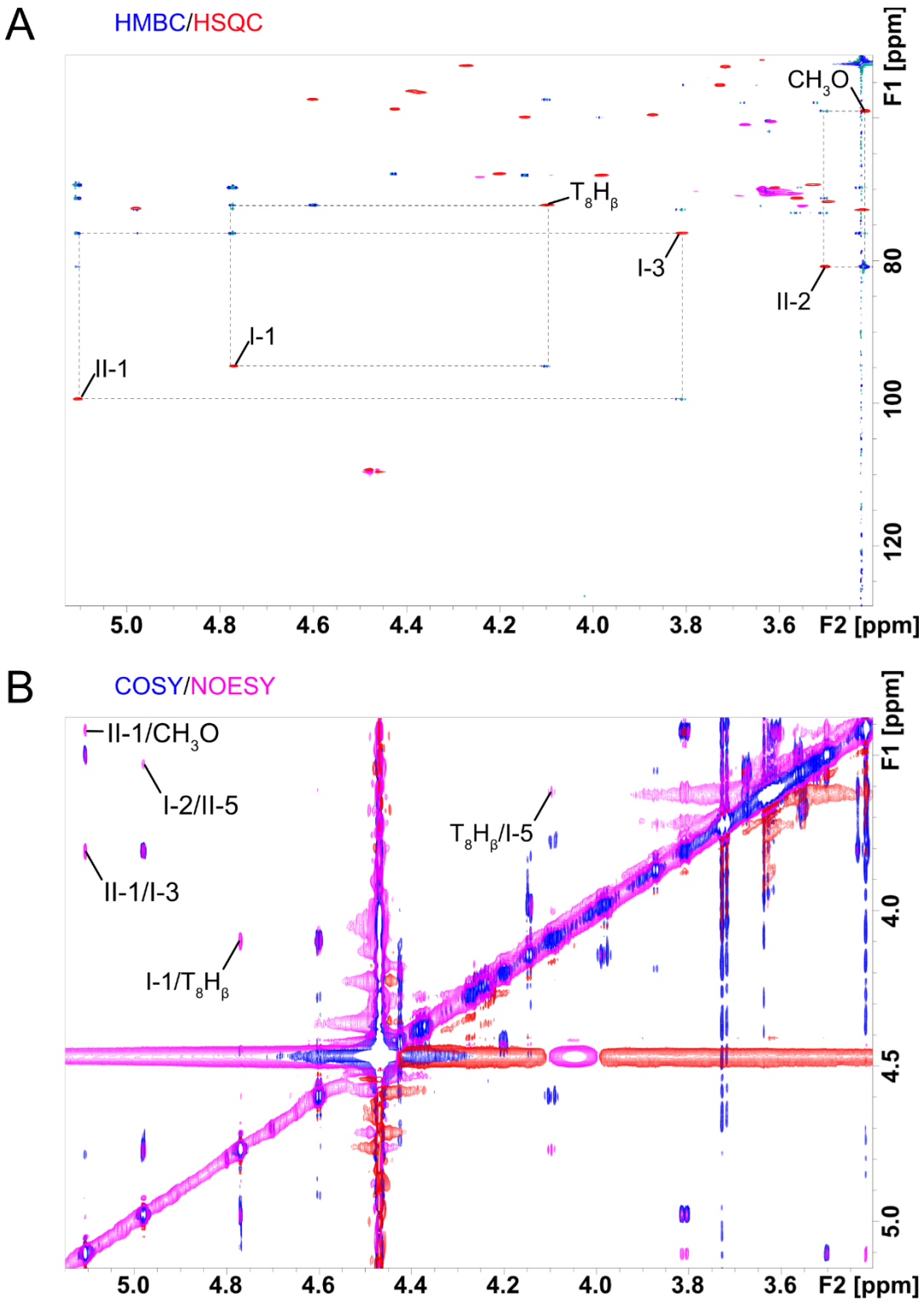


Figure 63 : Liaison des rhamnoses au peptide. Spectres des fractions 64-65 mesurés dans le $CDCl_3/CD_3OD$ (2:1) à 293K, après calibration sur le méthanol. Les expériences $^1H/^{13}C$ HSQC et HMBC superposés (A) mettent en évidence les corrélations protons/carbones entre les positions Rha I-1 et β du résidu T_8, la position Rha II-1 et Rha I-3 ainsi que la fonction méthoxy avec Rha II-2. Les spectres $^1H/^1H$ COSY et $^1H/^1H$ NOESY superposés (B) confirment la proximité spatiale des protons Rha I-1 avec T_8 ainsi que de Rha II-1 avec Rha I-3, confirmant l'origine des liaisons glycosidiques.

III.E.6.c.ii Spectres dans les solvants protiques

L'utilisation d'un mélange de solvants organiques contenant une fraction de méthanol protoné permet de favoriser la détection des protons amides échangeables qui ne sont pas visibles dans les solvants deutérés. Cela permet non seulement de vérifier la nature des acides aminés précédemment identifiés mais également de déterminer la séquence peptidique. Nous avons tout d'abord réalisé des spectres ^{1}H à différentes températures (293K, 283K et 275K) afin de repérer les protons associés à des fonctions amides autour d'un déplacement chimique de 8 ppm et de choisir la température optimale qui s'est révélée être 293K (Figure 64).

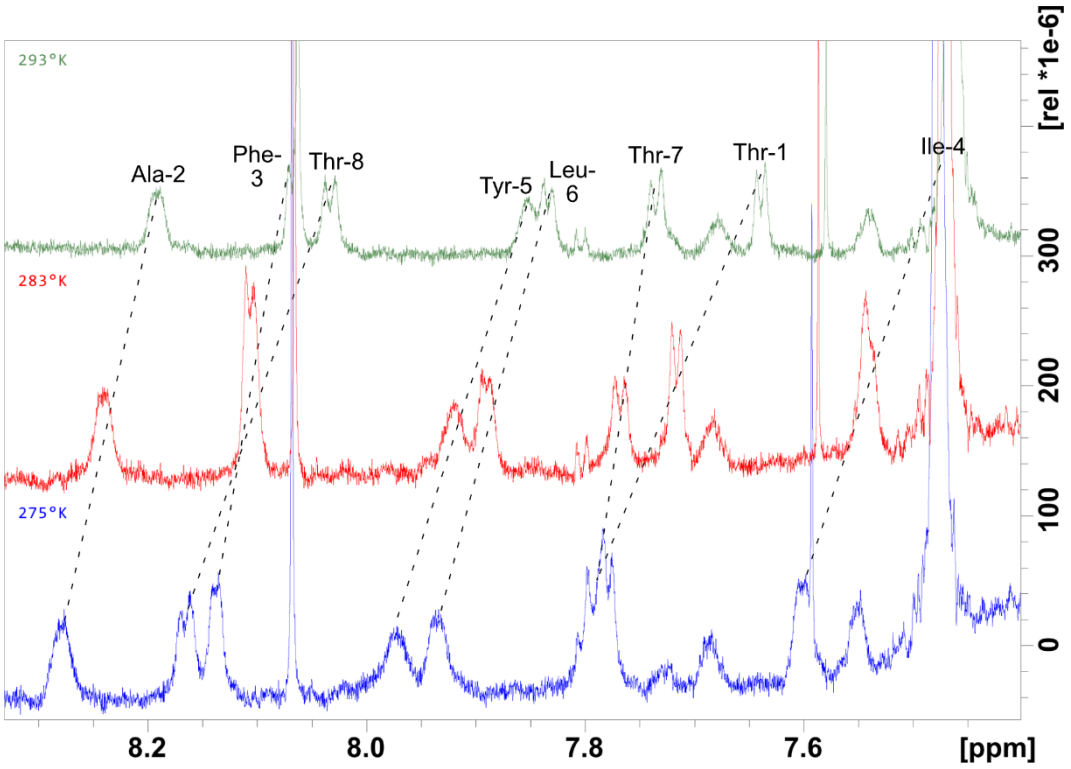


Figure 64 : Influence de la température dans les solvants protiques. Spectres RMN ^{1}H des fractions 64-65 mesurés à différentes températures dans $CDCl_3/CD_3OH$, après calibration sur le méthanol. La zone dans laquelle sont détectés les protons amides est agrandie. Le proton amide de l'isoleucine est le plus blindé à droite du spectre.

Séquence peptidique

La fenêtre spectrale 7.4-8.2/0.5-4.6 ppm du spectre $^{1}H/^{1}H$ TOCSY montre (en bleu) les corrélations entre les protons amides de chaque acide aminé avec tous les protons des autres positions pour les acides aminés non aromatiques (Figure 65A). Elle permet ainsi de confirmer l'ensemble du système de spin et la nature de chaque acide aminé.

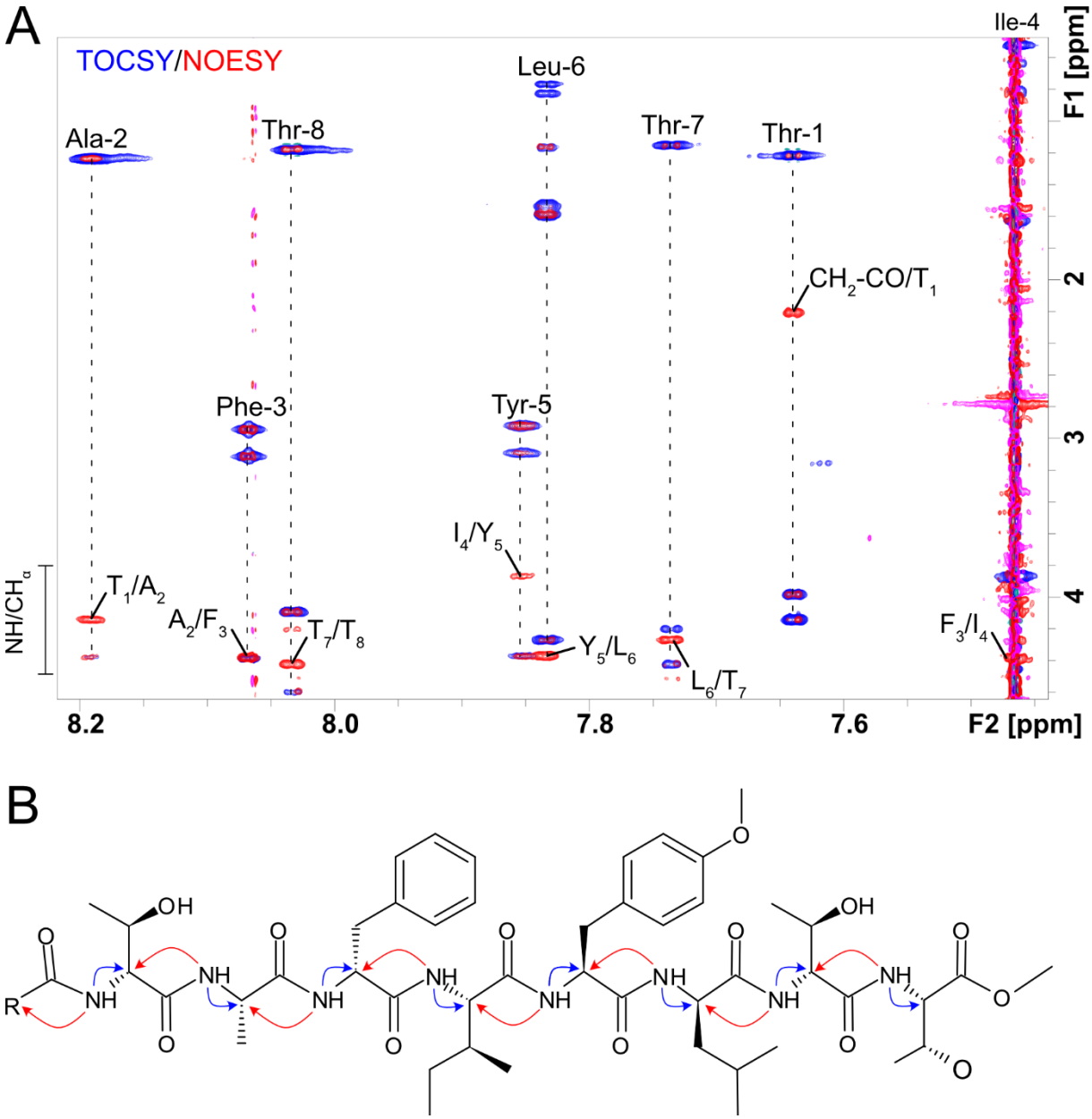


Figure 65 : Séquence peptidique du GPL4690. Spectres des fractions 64-65 mesurés dans $CDCl_3/CD_3OH$ à 293K, après calibration sur le méthanol. La superposition (A) des spectres TOCSY (bleu) et NOESY (rouge) confirme l'identité des acides aminés ainsi que leur enchainement (B) par la succession des corrélations scalaire (bleu) et spatiale (rouge) de chaque proton amide respectivement avec les protons α du résidu n et n-1.

La superposition du spectre $^1H/^1H$ NOESY sur le spectre $^1H/^1H$ TOCSY met en évidence les corrélations spatiales associées à la proximité du proton amide du résidu n avec le proton α du résidu n-1 qui permettent de séquencer le peptide (Figure 65A & B). A l'extrémité N-terminale, le proton amide du résidu T_1 présente une corrélation avec un groupement CH_2 du fait de l'acylation par un lipide. De l'autre côté, le proton α de T_1 est corrélé au proton amide de A_2. Les 6 autres corrélations

NOESY qui sont détectées dans la fenêtre 8,2-7,4/3,8-4,6 ppm permettent de séquencer le peptide vers l'extrémité C-terminale. Dans le cas de la corrélation entre le proton amide de F_3 et le proton α de A_2 à 4,376 ppm, elle est superposée à la corrélation TOCSY avec son propre proton α à 4,386 ppm. La corrélation suivante entre le proton amide de I_4 est le proton α de F_3 est masquée partiellement par le bruit de fond. Finalement, les 4 corrélations NOESY suivantes sont plus évidentes et relient dans l'ordre I_4 avec Y_5, puis L_6, T_7 et T_8. Enfin, la zone du spectre $^1H/^1H$ NOESY entre 8,2 et 7,6 ppm fait apparaitre les corrélations entre les protons amides du résidu d'alanine A_2 avec la thréonine T_1 et la phénylalanine F_3 de part et d'autre. En outre, les corrélations du proton amide de la thréonine T_7 avec celui de T_8 et de la leucine L_6 sont également visibles (Figure 66A).

Fonctions alcools

L'utilisation de solvants protiques permet aussi d'observer de nouvelles corrélations sur les spectres $^1H/^1H$ TOCSY et NOESY avec des protons alcools (Figure 66B). C'est le cas pour le résidu de thréonine T_1 dont les protons α, β et γ sont corrélés avec le proton de la fonction alcool à 5,159 ppm, ainsi que pour les positions β et γ du résidu de thréonine T_7 qui sont corrélés avec le proton de la fonction alcool à 5,257 ppm, confirmant ainsi que les groupements hydroxyles de ces deux thréonines sont restées libres. Le proton lié à la fonction alcool à 5,29 ppm possède des corrélations plus ou moins intenses avec les protons des positions 2 à 6 du 1er résidu α,rhamnopyranosyl. A contrario, les protons alcooliques des positions 3 et 4 du 2ème résidu n'ont pas été caractérisés.

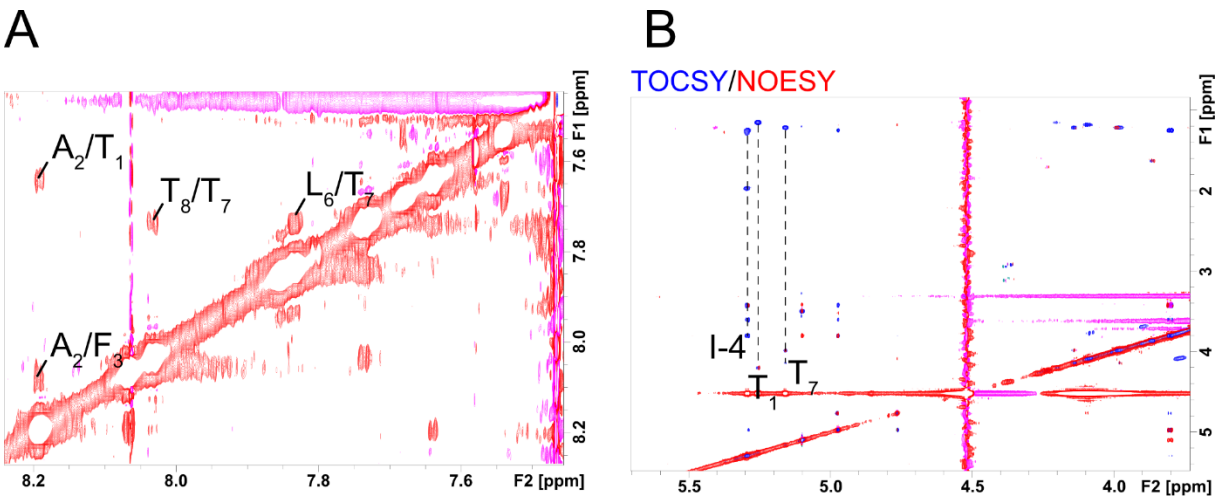
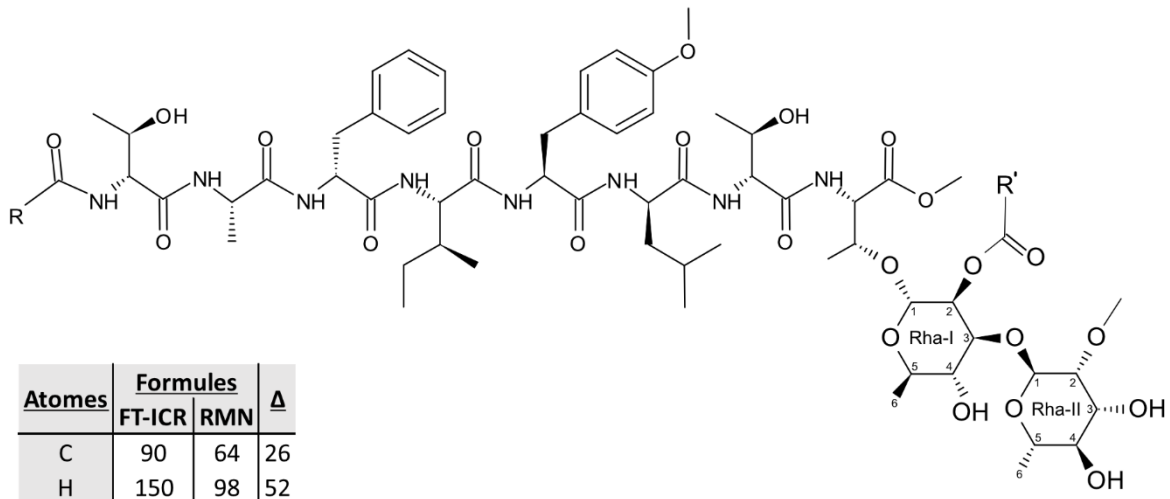


Figure 66 : Attribution des signaux du GPL4690 dans les solvants protiques. Spectres des fractions 64-65 mesurés dans dans $CDCl_3/CD_3OH$ à 293K, après calibration sur le méthanol. Le zoom du spectre NOESY sur la zone des protons amides confirment la position de A_2 entre T_1 et F_3, ainsi que celle de T_7 entre L_6 et T_8. L'apparition de corrélations nouvelles (lignes verticales en pointillés) sur la superposition (B) des spectres TOCSY (bleu) et NOESY (rouge) est liée à la présence de protons sur les fonctions alcools de T_7, T_1 et de la position Rha I-4.

III.E.6.c.iii Conclusion

L'analyse RMN complète des fractions 64 et 65 rassemblés pour lesquelles l'ion de rapport m/z 1750 est le plus intense, a permis d'identifier huit acides aminés et deux monosaccharides reliés entre eux. L'identification des acides aminés suggère la présence de para-*O*-méthyl tyrosine qui est un acide aminé original non décrit à ce jour. Le séquençage de l'octapeptide notée $T_1A_2F_3I_4Y_5L_6T_7T_8$ a été réalisé par HMBC dans les solvants deutérés, et par superposition $^1H/^1H$ TOCSY/NOESY dans les solvants protiques. De plus, les mêmes spectres RMN ont mis en évidence les deux liaisons glycosidiques entre le groupement hydroxyle de T_8 et le premier rhamnose et entre sa position I-3 et celle II-1 du second rhamnose. Enfin, deux acylations côté N-terminal et sur la position 2 du premier rhamnose ont été caractérisées. D'après la formule développée de la figure 67, nous avons donc attribué les 20 insaturations ainsi que les 8 atomes d'azote, mais nous n'avons attribué que 23 atomes d'oxygène sur les 24 prédits dans la formule brute $C_{90}H_{150}N_8O_{24}$ prédite par l'analyse FT-ICR. En plus de l'atome d'oxygène, il demeure 26 carbones et 52 hydrogènes contenus dans les chaines lipidiques des 2 groupements acyles.



Atomes	Formules		Δ
	FT-ICR	RMN	
C	90	64	26
H	150	98	52
N	8	8	0
O	24	23	1

Figure 67 : Structure intermédiaire du GPL4690 déduite des analyses FT-ICR et RMN. Les configurations D/L des positions α des acides aminés ont été déduites de l'analyse *in-silico* des gènes *MAB_4690c* et *MAB_4691c* (chapitre IV.E.7.a.i). D'après sa structure tripartite associant un noyau peptidique N-acylé à des rhamnoses méthoxylés ou acylés, ce nouveau composé appartient à la famille des GPLs. D'après le reste des atomes à attribuer (Δ), les chaines alkyles R et R' seraient des lipides saturés dont l'un est hydroxylé.

III.E.6.d Analyse par MALDI Q-IT TOF

III.E.6.d.i Spectres MS des fractions natives

Les spectres en mode positif des fractions 65 et 60 doublement purifiée, corroborent les résultats de l'analyse FT-ICR. En effet, ils génèrent chacun un ion intense à *m/z* 1750,243 et 1391,935 u.m.a. correspondant à des adduits $[M+Na]^+$ accompagnés d'ions secondaires M+/- 14 u.m.a., témoignant d'une faible dispersion lipidique.

III.E.6.d.ii Spectres de fragmentation MS^n des fractions natives

La fragmentation de l'ion parent *m/z* 1750 produit un grand nombre d'ions d'intensité variable selon l'énergie de collision choisie. L'ion de rapport *m/z* 1706 est plus intense à faible énergie alors que l'ion de rapport *m/z* 1662 est plus intense à forte énergie (Figure 68A). Ces deux fragments résultent de la perte par les résidus libres de thréonine de la chaine latérale CH_3CHO. En effet, cet évènement de fragmentation a fréquemment été observé sur les GPLs dépourvu de 6 déoxy-talose favorisant les fragments M-44 u.m.a. De même, des pertes de molécule d'eau M-18 u.m.a. multiplient les fragments secondaires.

Afin de simplifier l'interprétation, j'ai donc sélectionné l'ion fragment de rapport *m/z* 1662. Sur le spectre MS^3 obtenu, le nombre d'ions fragments est réduit (Figure 68B). Parmi eux, cinq ions fragments dont quatre très intenses ont été attribués aux fragments a_2 jusqu'à a_6 dont la différence de masse correspond à une phénylalanine, deux résidus de leucine ou d'isoleucine et un acide aminé de 177 Da compatible avec la masse théorique du para-*O*-méthyl tyrosine. Les ions fragments y_4 à y_7 se révèlent moins intenses que les fragments a. Néanmoins, les différences de masse constatées dans cette deuxième série sont en accord avec la présence d'un enchainement leucine/isoleucine, phénylalanine et alanine. D'autre part, un ion de rapport *m/z* 1502 apparait très peu intense. Ce fragment M-160 u.m.a. est compatible avec la perte d'un *O*-méthyl déoxyhexose. L'ion *m/z* 1140 correspond à une perte de 522 u.m.a. expliqué par le départ du motif glycannique 2-*O*-Me-α-rhamnopyranosyl(1-3)2-*O*-hyrdoxylauryl-α-rhamnopyranose. Dans ce cas, le résidu acyle est compatible avec un acide hydroxy laurique (Figure 68B).

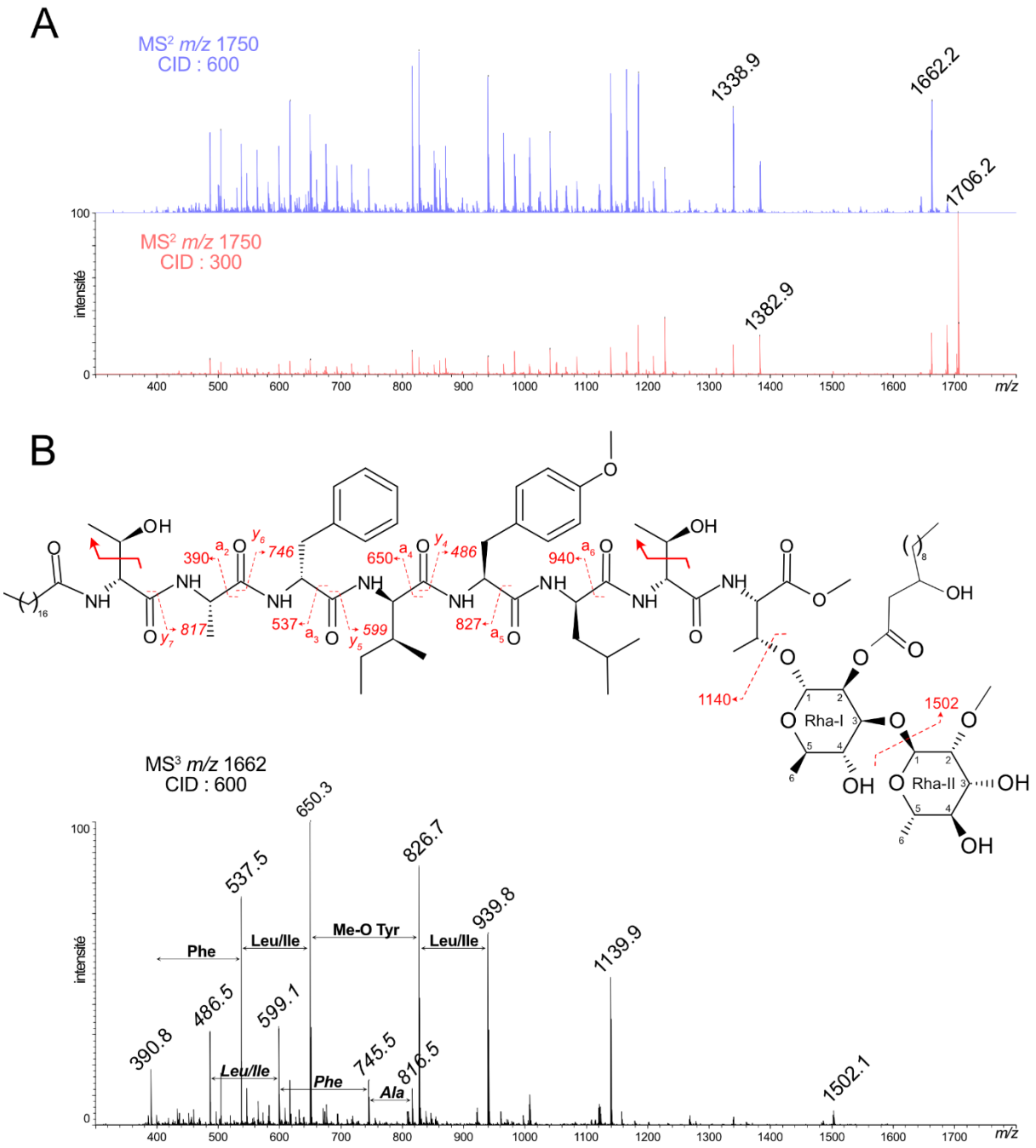


Figure 68 : Analyse MALDI-TOF du GPL4690. La fragmentation en mode CID de l'ion parent *m/z* 1750 avec des énergies croissantes produit un ion *m/z* 1706 M-44 puis *m/z* 1662 M-88 u.m.a.. suite au départ des chaines latérales des 2 thréonines (A). Le spectre MS^3 du parent *m/z* 1662 (B) révèle un fragment M-160 u.m.a.. compatible avec le départ d'un *O*-méthyl déoxyhexose, ainsi que des fragments a et y provenant du noyau peptidique dont les deux séries sont en accord avec la séquence déduite de la RMN. Les ions *m/z* 1382 et 1338 (A) correspondent au fragment y_7 avant la perte du motif glycannique et d'une chaine latérale de thréonine.

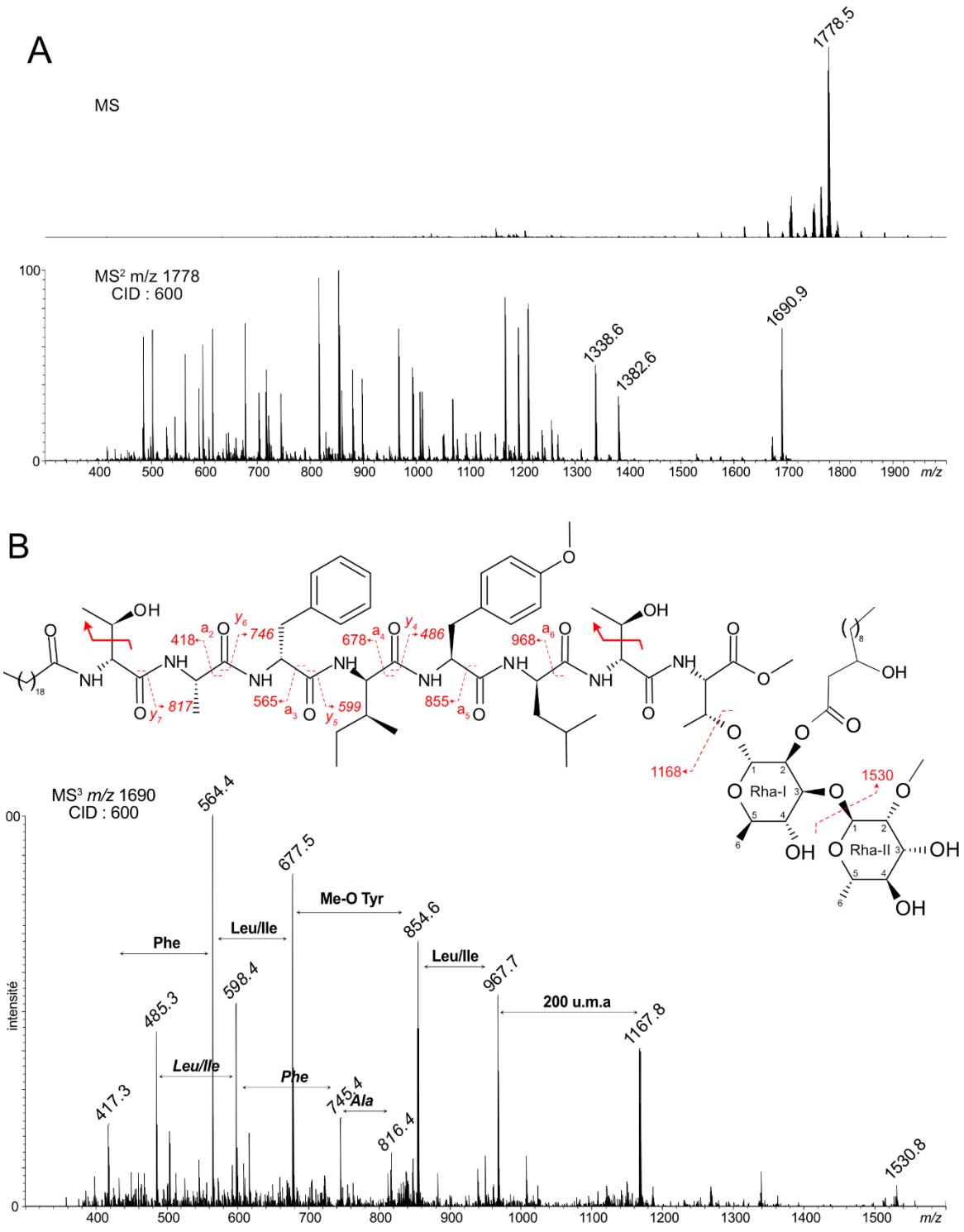


Figure 69 : Analyse MALDI-TOF d'un variant lipidique du GPL4690. La fragmentation de l'ion parent *m/z* 1778 avec l'énergie de collision CID 600 produit un ion *m/z* 1690 M-88 u.m.a.. suite au départ des chaines latérales des 2 thréonines (A). Le spectre MS^3 du parent *m/z* 1690 (B) révèle un fragment M-160 u.m.a.. compatible avec le départ d'un *O*-méthyl déoxyhexose, ainsi que des fragments a (+28 u.m.a.. par rapport à la figure 68) et y provenant du noyau peptidique dont les deux séries sont en accord avec la séquence déduite de la RMN. Les ions *m/z* 1382 et 1338 (A) correspondent au fragment y_7 avant la perte du motif glycannique et d'une chaine latérale de thréonine.

Des expériences MS^2 puis MS^3 sur l'ion à m/z 1778, détecté dans la fraction 68 suivante lors de la double purification, ont révélé les mêmes mécanismes de fragmentation concernant la perte rapide des chaines latérales de thréonine, d'un *O*-méthyl déoxyhexose et la présence des fragments a et y confirmant la séquence peptidique (Figure 69A & B). De plus, la comparaison des spectres MS^2 de 1750 et 1778 (Figure 68A et 69A) met en évidence deux ions à m/z 1382 et 1338 correspondant au fragment y_7 avant la perte du motif glycannique et de la chaine latérale de thréonine dont les spectres MS^3 sont identiques.

Dans le cas de la fraction 60, Les mêmes expériences démontrent que l'ion parent *m/z* 1391 produit successivement l'ion fragment *m/z* 1348 puis 1304 (Figure 70A), confirmant la présence de deux résidus thréonine. En outre, la sélection de cet ion *m/z* 1304 et sa fragmentation MS^3 révèle un spectre avec les mêmes fragments a et y que ceux observés dans le spectre MS^3 1750>1662 de la fraction 65, supposant un noyau peptidique commun (Figure 68B et 70B). A nouveau, le spectre MS^3 1391>1140 de la fraction 60 présentent les mêmes fragments a et y. C'est pourquoi, l'ion fils *m/z* 1140 résulte de la perte par l'ion parent *m/z* 1304 d'un déoxyhexose de 164 u.m.a. relié à une thréonine.

III.E.6.d.iii Spectres des fractions après traitement alcalin

Dans le but de libérer le lipide supposé être acylé sur le premier résidu de rhamnose *via* une liaison ester, tout en préservant la liaison amide côté N-terminal, j'ai choisi de tester l'hydrolyse alcaline douce. Cette méthodologie avait notamment été appliquée avec succès dans l'étude des GPLs de *Mabs* pour démontrer la présence des deux groupements acétyles.

Le spectre d'une fraction purifiée uniquement sur silice puis traitée avec la soude 0,1 M a produit un ion majoritaire de rapport *m/z* 1228 accompagné de deux ions moins intenses *m/z* 1246 et 1260 alors que les ions 1391 et 1750 initialement détectés dans la fraction native ont complétement disparu (Figure 71A). De même, les spectres de la fraction polaire totale et de la fraction F67 doublement purifiée sur silice vierge puis sur silice greffée C18 (Figure 57B) présentent des massifs autour des ions *m/z* 1228 et 1246.

III.E.6.d.iv Spectres MS^n des fractions après traitement alcalin

La fragmentation de l'ion *m/z* 1228 génère un fragment intense à *m/z* 1184 suivi d'un second à m/z 1140. Le spectre MS^3 de l'ion parent *m/z* 1184 produit l'ion fragment *m/z* 1140. Les pertes successives M-44 u.m.a. suggèrent encore le départ des chaines latérales CH_3CHO des résidus de thréonine. D'autre part, les fragments obtenus sur le spectre MS^3 de l'ion parent *m/z* 1184 mettent en évidence des résidus isoleucine/leucine, alanine, phénylalanine et thréonine (Figure 71B).

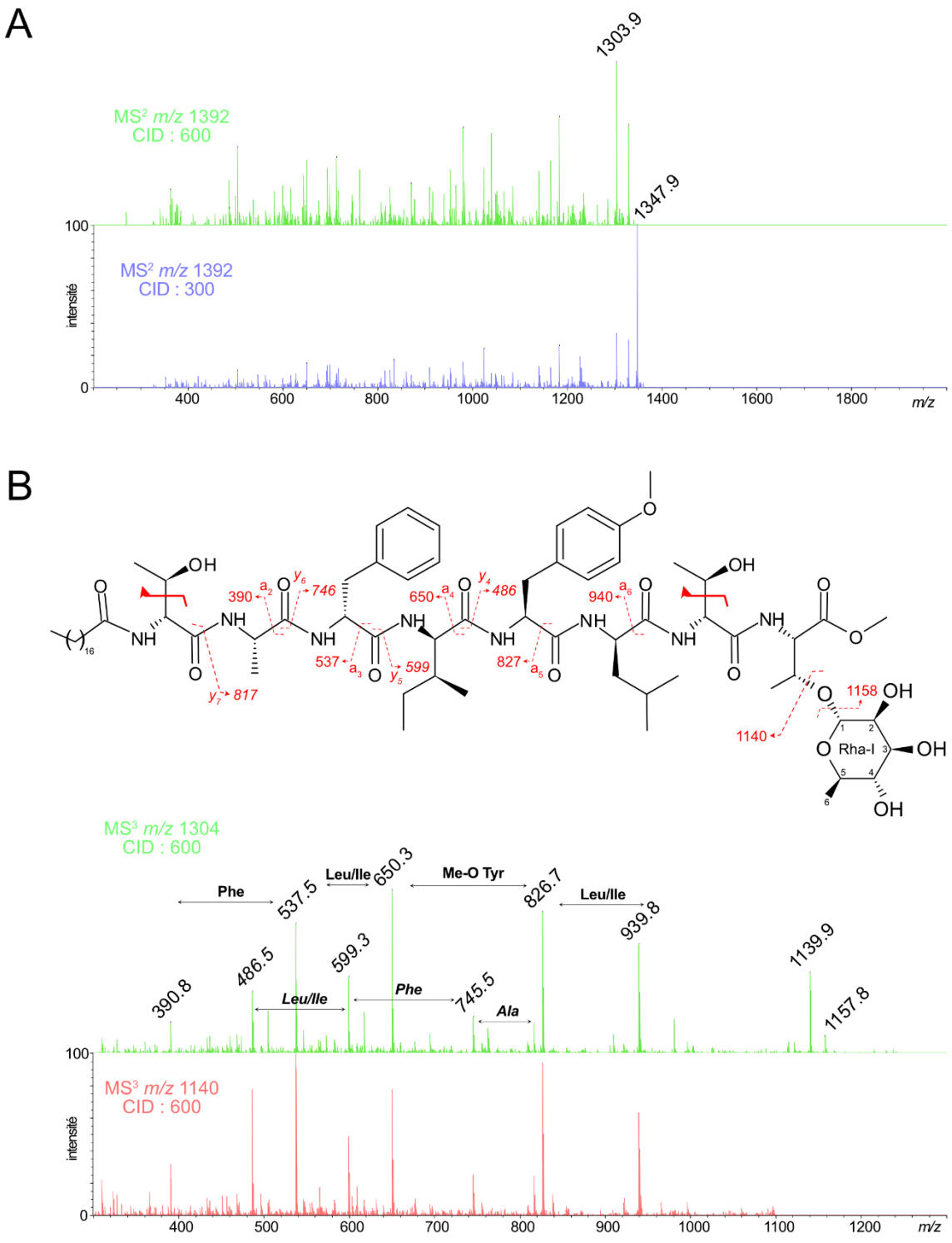


Figure 70 : Analyse MALDI-TOF d'un dérivé du GPL4690. La fragmentation en mode CID (A) de l'ion parent *m/z* 1391 avec des énergies croissantes produit un ion *m/z* 1348 M-44 u.m.a. (bleu) puis *m/z* 1304 M-88 u.m.a..(vert) suite au départ des chaines latérales des 2 thréonines. Le spectre MS^3 du parent *m/z* 1304 (B) révèle des fragments a et y provenant du noyau peptidique dont les deux séries sont en accord avec la séquence déduite de la RMN ainsi qu'un fragment M-164 u.m.a.. à *m/z* 1140 compatible avec le départ d'un déoxyhexose (vert). En outre le spectre MS3 de cet ion produit les mêmes fragments a et y (rouge).

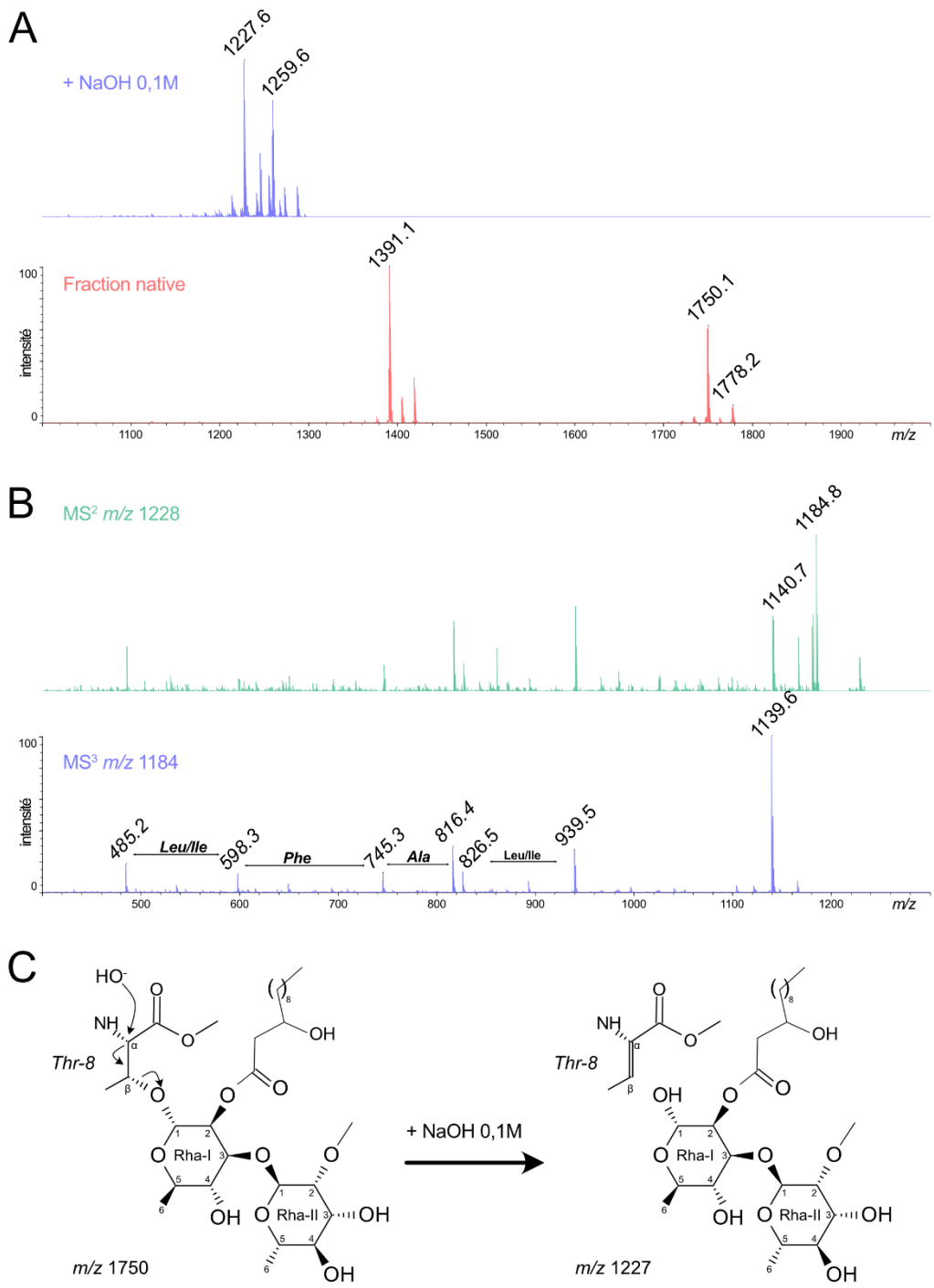


Figure 71 : Traitement alcalin du GPL4690 et de ses dérivés. Spectres MALDI-TOF (A) de fractions purifiées uniquement sur silice, contenant le GPL4690 et ses dérivés, avant (rouge) et après (bleu) traitement avec 0,1 M NaOH. Les ions *m/z* 1391, 1750 et 1778 disparaissent au profit d'un massif dans lequel on retrouve l'ion *m/z* 1228. La fragmentation de ce dernier (B, vert) produit successivement les ions *m/z* 1184 et 1140, issues de la perte des chaines latérales des thréonines. Les fragments y_4 à y_7

précédemment décrits sont finalement visibles dans le spectre MS^3 de l'ion *m/z* 1184. La réaction de β-élimination illustrée sur l'ion *m/z* 1750 (C) entraine le départ du motif glycannique.

III.E.6.d.v Conclusion

L'interprétation des spectres de fragmentation d'au moins 3 fractions natives et une fraction traitée à la soude montre non seulement qu'elles contiennent des composés dont le noyau peptidique est commun mais aussi que sa séquence est en accord avec celle déduite de l'analyse RMN.

D'autre part, les expériences MS^n réalisées sur l'ion à *m/z* 1750 suggèrent que le lipide N-acylé est un acide stéarique alors que celui acylé sur la position 2 du rhamnose est un acide hydroxylaurique, contenant notamment le dernier atome d'oxygène prédit par l'analyse FT-ICR. Seule la position de ce groupement hydroxyle n'a pas pu être déterminée dans la structure du GPL4690 (Figure 69B).

En outre, les résultats de fragmentation pour l'ion à *m/z* 1391 suggèrent que ce dérivé du GPL4690 ne contient qu'un unique deoxyhexose porté par le même noyau lipopeptidique. En accord avec la formule brute $C_{71}H_{116}N_8O_{18}$ déduite de l'analyse FT-ICR et les 2 insaturations manquantes, nous proposons que ce dérivé du GPL 4690 a perdu le 2-*O*-méthyl rhamnose terminal ainsi que le groupement hydroxylauryl (Figure 70B).

Enfin le traitement alcalin, initialement choisi pour déduire la nature des lipides, a entrainé une réaction de β-élimination décrite à la figure 71C en accord avec la présence d'une déhydrothréonine en position 8 du lipopeptide. De plus, ce résultat renforce la présence du motif glycannique 2-*O*-Me-α-rhamnopyranosyl(1-3)2-*O*-hydroxylauryl-α-rhamnopyranose dans le GPL4690 et son dérivé, même si son analyse direct n'a pas été réalisé.

III.E.6.e Détermination de la composition en monosaccharides et en lipides

Seule la composition monosaccharidique d'une fraction doublement purifiée produisant l'ion à *m/z* 1750 et dont la pureté a été vérifiée par un spectre ^1H RMN a été déterminée. Après hydrolyse totale de cette fraction, suivie d'une dérivation par la méthode des itols acétates, j'ai analysé l'échantillon par GC-MS. Sur la base des temps de rétention et des spectres EI-MS ces analyses ont confirmé la présence d'un résidu de rhamnose à Tr = 19.78 min et d'un dérivé de deoxyhexose méthylé à Tr = 17,95 min (Figure 72A-C). Ce dernier a été identifié sur la base des temps de rétention de dérivés *O*-méthylés du rhamnose provenant de *Bacillus cereus* et du spectre EI-MS produisant trois ions de rapport *m/z* 117, 129 et 275 qui démontrent qu'il s'agit de 2-*O*-méthyl rhamnose (Figure 72B). La composition en monosaccharides de la fraction répondant à l'ion m/z 1391 restera à déterminer.

D'autre part, seule la présence d'acide stéarique sous forme d'esters méthyliques d'acides gras a pu être confirmée après méthanolyse acide et GC-MS. D'autres composés demeurent à identifier.

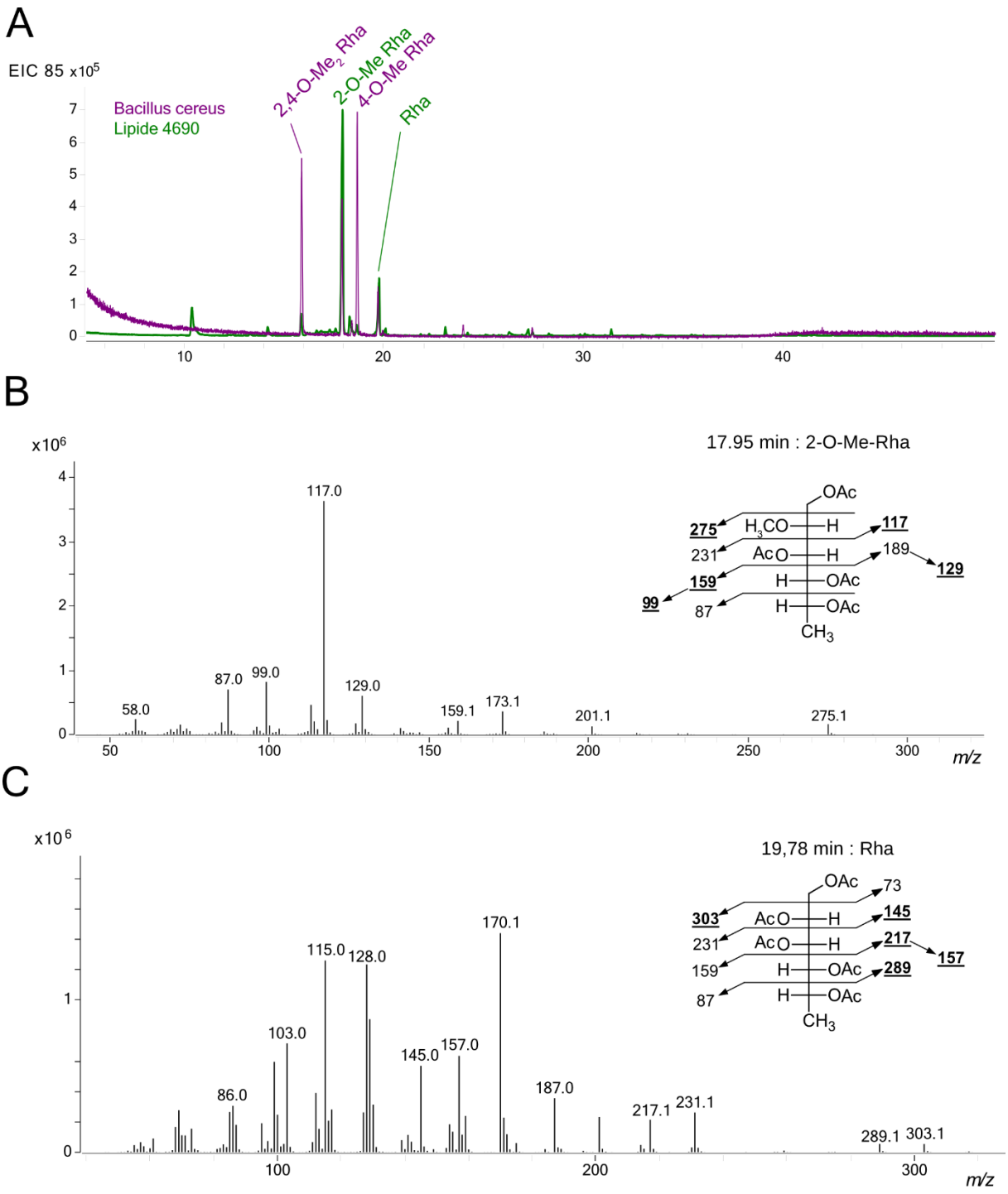


Figure 72 : Composition en monosaccharides du GPL4690. Après dérivation, les itols-acétates ont été séparées en GC-MS et les déoxyhexoses ont été repérés grâce à l'ion *m/z* 85 (A). Ceux du GPL4690 apparaissent en vert et ceux de *Bacillus cereus* en violet. 2 pics chromatographiques à Tr = 17,95 min et 19,78 min sont communs aux 2 échantillons. D'après leur spectre, le 1er est un 2-*O*-méthyl rhamnose (B) et le second est un rhamnose (C).

III.E.6.f Conclusion sur les structures identifiées

L'ensemble des analyses menées sur les fractions purifiées a montré l'existence d'au moins trois variantes structurales d'un nouveau GPL, nommé GPL4690.

L'analyse la plus complète a été menée sur le GPL4690 correspondant à l'ion m/z 1750. En effet, sa formule brute la plus probable $C_{90}H_{150}N_8O_{24}$ déduite de l'analyse FT-ICR révèle 20 insaturations toutes caractérisées par l'analyse RMN. De même, cette technique a permis de déterminer la séquence octapeptidique $T_1A_2F_3I_4Y_5L_6T_7T_8$ dans laquelle la para-O-méthyl tyrosine représente un acide aminé jamais décrit. En outre, la RMN a mis en évidence la liaison glycosidique entre la thréonine T_8 et le motif 2-*O*-Me-α-rhamnopyranosyl(1-3)2-*O*-acyl-α-rhamnopyranose. La présence des 2 déoxyhexoses a été confirmée par la composition en monosaccharides obtenue en GC-MS mais leur configuration D/L n'a pas été étudiée. Enfin, l'interprétation des spectres MS et MS^n ainsi que l'ensemble des atomes restant à attribuer montre qu'un acide stéarique C_{18} est acylé du côté N terminal alors qu'un acide hydroxylaurique est estérifié sur la position 2 du 1er rhamnose. D'après sa fragmentation, l'ion à m/z 1778 possèderait un acide eicosanoïque C_{20} côté N-terminal.

Par déduction, un second GPL correspondant à l'ion m/z 1391 a été caractérisée. Son mécanisme de fragmentation MALDI-TOF s'est révélé fortement similaire au GPL4690 supposant un noyau lipopeptidique commun portant un unique déoxyhexose non modifié. Considérant la nouvelle formule brute $C_{71}H_{116}N_8O_{18}$ issue de l'analyse FT-ICR, nous proposons que l'absence du 2-*O*-Me rhamnose et de l'acide hydroxylaurique explique la différence d'insaturations et d'atomes avec la formule brute précédente $C_{90}H_{150}N_8O_{24}$ (Figure 73).



Figure 73 : Structures identifiées du GPL4690. Suite à toutes les analyses structurales réalisées en FT-ICR, RMN, MALDI-TOF et GC-MS, nous avons pu déterminé les 2 variantes dirhamnosylées (à gauche) et monorhamnosylée (à droite) du GPL4690 avec le rappel de leurs formules brutes et du rapport *m/z* de l'adduit $[M+Na]^+$. Les configurations D/L demeurent à déterminer ainsi que la position du groupement hydroxyle dans l'acide hydroxylaurique.

III.E.7 RECHERCHE D'UNE CORRELATION AVEC LE GENE *MAB_4690c*

Compte tenu des structures chimiques précédemment identifiés et de la similarité entre le locus de biosynthèse des GPLs et celui auquel appartient le gène *MAB_4690c*, l'objectif des résultats qui suivent est d'apporter des éléments démontrant que les deux variantes du GPL4690 résultent de l'expression des gènes de ce locus (Figure 15C). Dans un premier temps, je présenterai quelques informations issues des analyses bioinformatiques sur les gènes de ce locus. Ensuite, je terminerai sur l'analyse de mutants comparativement à la souche sauvage.

III.E.7.a Analyse bioinformatique des gènes autour de *MAB_4690c*

D'après une recherche sur la base de données des génomes mycobactériens https://mycobrowser.epfl.ch, le gène *MAB_4690c* de la souche ATCC_19977 possède 8151 paires de base (bp) et code une probable peptide synthase non ribosomale (NRPS) de même que le gène *MAB_4691* qui s'étend sur 24327 bp. Au voisinage de ces deux gènes, les gènes *MAB_4703* et *MAB_4704* codent des protéines de transport de type *MmpL*, et le gène *MAB-4705* celle de type *MmpS*. Les gènes *MAB_4685* et *MAB_4686* ainsi que *MAB_4694* et *MAB_4695* ont été annotés comme de potentielles glycosyltransférases. Enfin, *MAB_4696* et *MAB_4700* coderaient chacune une méthyl-transférase d'après le serveur www.uniprot.org. C'est pourquoi l'organisation de ce locus rappelle celui de la biosynthèse des GPLs (Figures 11 & 15C). J'ai donc évalué la fonction supposée de chacun des gènes au moyen d'outils bioinformatiques de prédiction et d'alignement.

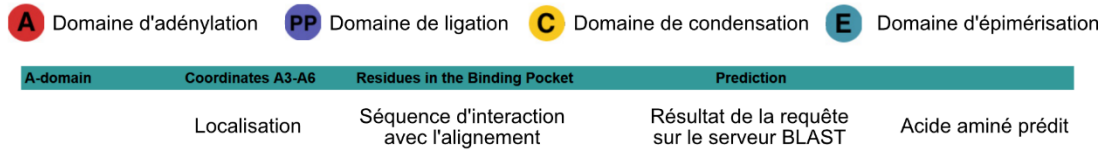
III.E.7.a.i Prédiction des substrats ajoutés par *MAB_4690c* et *MAB_4691c*

L'analyse du gène *MAB_4690c* sur le site http://nrps.igs.umaryland.edu (Figure 74A) suppose qu'il code pour un enchainement de 2716 acides aminés. La chaine polypeptidique est organisée en huit domaines distincts qui se regroupent en deux modules. Dans le premier module, les trois domaines A, PP et C participent à l'activation et à la condensation d'un résidu de thréonine, alors que la présence du domaine E dans le deuxième module pourrait être responsable de l'épimérisation d'un second résidu de thréonine (Figure 74B).

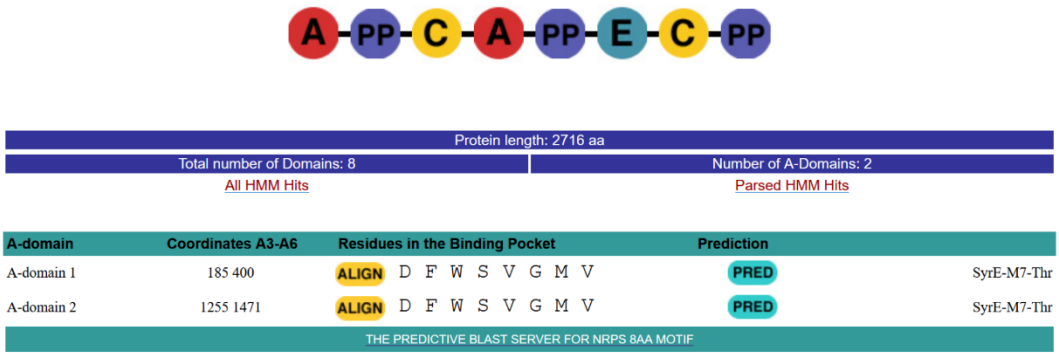
Le gène voisin *MAB_4691* étant plus grand, il code six modules distincts pour une longueur de 8108 acides aminés. Les 22 domaines prédits participent à l'ajout d'acides aminés aliphatiques de type leucine, phénylalanine et valine ainsi que d'une thréonine. A nouveau, la présence de trois épimérases implique que les acides aminés en position 1, 3 et 6 soient modifiés vers la configuration D. Enfin, la

prédiction du substrat potentiellement ajouté n'a pas été possible pour le troisième module puisqu'aucun équivalent connu n'a été retrouvé pour l'alignement (Figure 74C).

A



B



C

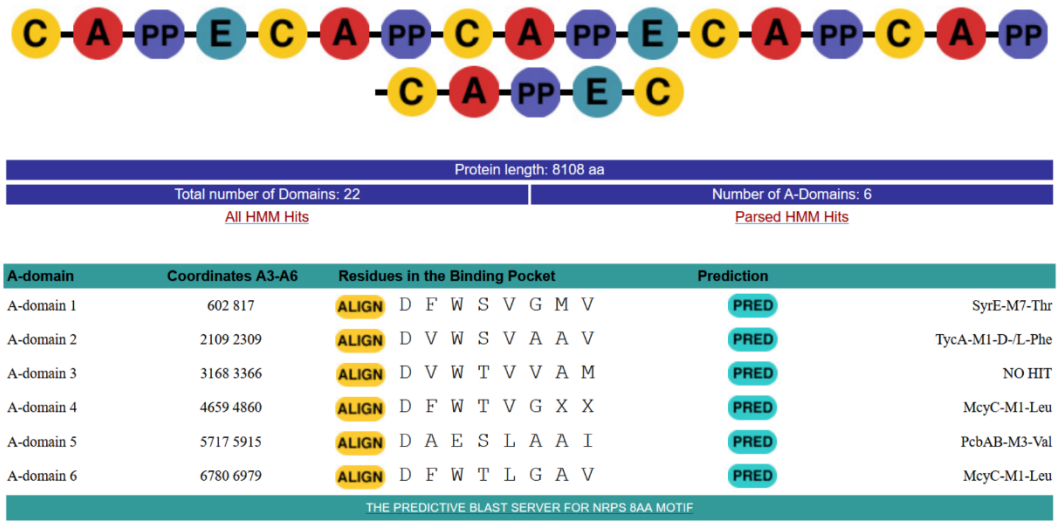


Figure 74 : Prédiction de substrats ajoutés par les NRPS. (A) Légende indiquant le rôle des domaines protéiques et les informations d'alignement. (B) Le produit de *MAB_4690c* ajoute deux résidus de thréonine alors que celui de *MAB_4691* relie cinq acides aminés dont celui ajouté par le domaine 3 qui n'est pas prédit.

Ensemble, les deux gènes pourraient être responsables de la production d'un octapeptide contenant majoritairement des résidus de thréonine et d'acides aminés aliphatiques dont 4 dans une configuration D. Le GPL4690 sera donc recherché dans les mutants délétés pour ces deux gènes.

III.E.7.a.ii Recherche des rhamnosyltransférases potentielles

Les gènes *MAB_4694* et *MAB_4695* ont été annotés comme des glycosyltransférases. Leurs séquences présumées transcrites ont donc été alignées avec celles des gènes codant les trois glycosyltranférases *gtf1*, *gtf2* et *gtf3* impliquées dans la biosynthèse des GPLs. L'objectif est de distinguer ces deux gènes de *MAB_4685* et *MAB_4686* qui codent également pour des glycosyltransférases afin de poursuivre l'étude fonctionnelle par une approche de génétique inverse. L'utilisation de l'outil MUSCLE afin de procéder à un alignement multiple global sur les sept séquences montre une forte identité entre les deux gènes ainsi qu'avec les gènes *gtfs* du locus de biosynthèse des GPLs qui varie de 54 à 64% (Tableau 22). Par contre, le pourcentage d'identité avec les gènes *MAB_4685* et *MAB_4686* oscille entre 15 et 32%. Si l'on exclut ses deux gènes, l'analyse des cinq séquences restantes sur le serveur PROMALS confirme une identité stricte des résidus pour plus de 150 acides aminés. Néanmoins, les pourcentages constatés d'identité ne permettent de prédire la nature exacte du monosaccharide transféré. C'est pourquoi, l'analyse du contenu en glycolipide et de la composition en monosaccharides des mutants Δ*MAB_4694* et Δ*MAB_4695* sera nécessaire.

Tableau 22 : Recherche des rhamnosyltransférases impliquées dans la biosynthèse du GPL4690. Alignement des séquences protéiques des gènes *gtf1*, *gtf2* et *gtf3* codant les glycosyltranférases du locus de biosynthèse des GPLs avec quatre glycosyltranférases au voisinage des gènes *MAB_4690c* et *MAB_4691c*. Les protéines issus de *MAB_4694* et *MAB_4695* présentent l'identité la plus forte avec les Gtf1-3 et seront délétés pour des études ultérieures.

	MAB_4694	***MAB_4695***	***Gtf1***	***Gtf2***	***Gtf3***	***MAB_4686***	***MAB_4685***
MAB_4694	100.00	64.36	59.95	55.18	55.37	30.85	15.79
MAB_4695	64.36	100.00	57.44	54.81	55.56	28.36	16.56
Gtf1	59.95	57.44	100.00	56.87	60.38	30.85	17.88
Gtf2	55.18	54.81	56.87	100.00	60.77	30.85	20.00
Gtf3	55.37	55.56	60.38	60.77	100.00	32.34	16.22
MAB_4686	30.85	28.36	30.85	30.85	32.34	100.00	26.72
MAB_4685	15.79	16.56	17.88	20.00	16.22	26.72	100.00

III.E.7.a.iii Alignement avec les autres gènes du locus de biosynthèse des GPLs

A l'inverse, la comparaison de *MAB_4696 et MAB_4700* avec les gènes codant les méthyltransférases *Fmt*, *Rmt2*, *Rmt3* et *Rmt4* dont les fonctions sont respectivement l'addition d'un groupement O-méthyle sur le lipide et sur les positions 2, 3 et 4 du rhamnose révèlent un taux de similarité oscillant entre 15 et 25% (Tableau 23), mais les protéines de transport MmpL retrouvées dans les 2 locus possèdent 58,9 à 67,4% d'identité (Tableau 24). La fonction du gène *MAB_4696* resterait donc à définir puisque deux groupements méthyles sur la tyrosine et le rhamnose sont présents dans les structures identifiées.

Tableau 23 : Alignement des protéines impliqués dans la méthylation des GPLs. Les méthyltransférases codées par les gènes *MAB_4696c* et *MAB_4700c* possèdent peu de similarité avec les méthyltransférases responsables de la modification du lipide (Fmt) ou du rhamnose (Rmt2-4).

	Rmt2	***Fmt***	***MAB_4696c***	***MAB_4700c***	***Rmt3***	***Rmt4***
Rmt2	100.00	17.87	16.75	18.05	12.06	14.14
Fmt	17.87	100.00	26.27	17.37	14.29	14.22
MAB_4696c	16.75	26.27	100.00	15.88	14.29	12.29
MAB_4700c	18.05	17.37	15.88	100.00	15.56	15.88
Rmt3	12.06	14.29	14.29	15.56	100.00	24.90
Rmt4	14.14	14.22	12.29	15.88	24.90	100.00

Tableau 24 : Alignements des protéines impliqués dans le transport des GPLs. Les gènes codant les protéines de transport des GPLs (*MmpL4*) possèdent une forte similarité avec celles potentiellement impliquées dans la prise en charge du GPL4690 et codées par *MAB_4703c* et *MAB_4704c*.

	MmpL4a	***MAB_4703c***	***MAB_4704c***	***MmpL4b***
MmpL4a	100.00	63.92	58.88	62.84
MAB_4703c	63.92	100.00	64.98	62.83
MAB_4704c	58.88	64.98	100.00	67.40
MmpL4b	62.84	62.83	67.40	100.00

III.E.7.a.iv Conclusion

L'analyse bioinformatique des gènes *MAB_4690c* et *MAB_4691c* prédit un noyau riche en thréonines et parmi huit acides aminés dont la structure est proche de celle identifiée pour le GPL4690 dans le chapitre IV.E.6.f. D'autre part, la comparaison des glycosyltransférases avec Gtf1, Gtf2 et Gtf3 dont les fonctions respectives ont été démontrées dans la biosynthèse des GPLs (Figure 55 et chapitre IV.E.5.d) montre que les probables rhamnosyltransférases codées par les gènes *MAB_4695c* et *MAB_4696c* ont l'identité la plus élevée. De même, la présence de 2 méthyltransférases codées par *MAB_4696c* et *MAB_4700c* soutiendrait les O-méthylations constatées de latyrosine et du second rhamnose. Enfin, la présence des gènes codant les protéines de transport MmpL dont la forte identité avec les MmpL4 responsable du transport des GPLs est un dernier argument soutenant l'implication du locus au voisinage du gène *MAB_4690c* dans la biosynthèse et le transport du GPL4690.

La fonction supposée de chaque gène pourra se confirmer par l'étude du contenu glycolipidique des mutants correspondants. C'est pourquoi, nous avons commencé par l'analyse des mutants *ΔMAB_4690c* dont la virulence était fortement diminuée chez la souris (Figure 15A), et *ΔMAB_4691c* qui code la seconde NRPS.

III.E.7.b Production du GPL4690 par *Mabs* et les mutants NRPS

Au début du projet thèse, nous avions envisagé l'utilisation d'un mutant *ΔMAB_0855* délété du gène *MmpL8* impliqué dans le transport du GDND[158] pour faciliter la purification du GPL4690. Finalement, cette stratégie a été abandonnée et ce mutant s'est révélé être une souche contrôle pour la comparaison avec la souche sauvage *Mabs* R et le mutant *ΔMAB_4690c* dans l'analyse CCM ainsi que lors du couplage chromatographie flash avec le MALDI-TOF. Enfin, ce mutant ainsi que celui délété du gène *MAB_4691* ont tous les deux fait l'objet des analyses préliminaires LC-MS.

III.E.7.b.i Séparation CCM

Après culture sur un milieu 7H9OADC et préparation des extraits polaires totaux de *Mabs* R, des mutants *Δ_MAB0855* et *Δ_MAB4690* ainsi que des souches complémentées, j'ai réalisé une séparation CCM suivie d'une coloration avec l'orcinol-sulfurique, l'acide phosphomolybdique ou l'acide sulfurique. Les profils obtenus diffèrent surtout pour le mutant *ΔMAB_4690c* avec une bande de R_f 0,29 dont l'intensité diminue voire s'annule selon la méthode de révélation (Figure 75). La difficulté de mise en évidence par CCM d'un glycolipide faiblement exprimé, lorsque le milieu de culture favorise la production de glycolipides majoritaires, m'a amené à rechercher l'ion *m/z* 1750 sur des extraits polaires totaux de *Mabs* R préalablement séparés.

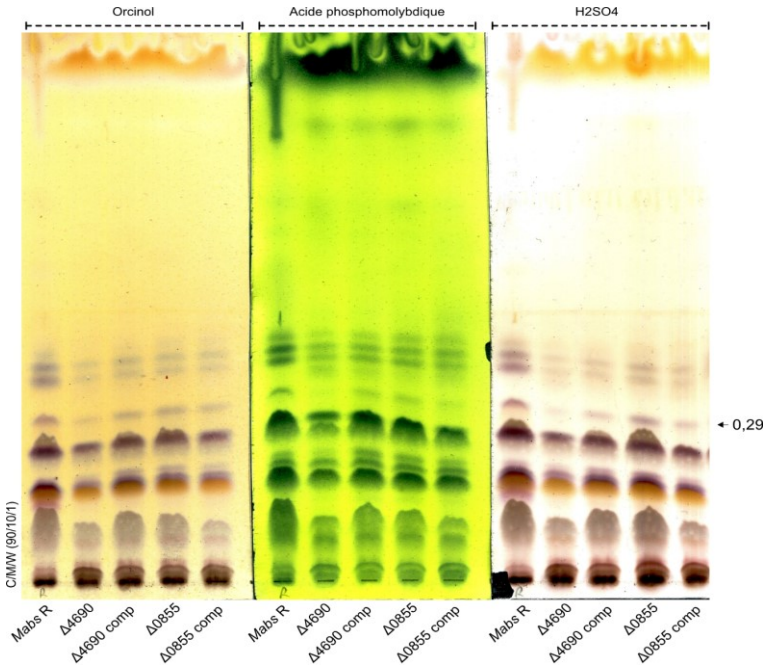


Figure 75 : Analyse CCM de *Mabs* R et de mutants dans la synthèse ou le transport de glycolipides. Parmi les 3 révélations, le PMA met en évidence la disparition d'une bande de R_f 0,29 qui demeure détectable à l'orcinol-sulfurique. La richesse du profil de glycolipides et le seuil de détection nous a conduits à tester d'autres méthodologies sur les mutants *Δ_MAB4690* et *Δ_MAB0855*.

III.E.7.b.ii Séparation par chromatographie-flash puis détection MALDI-TOF

Les lipides de l'extrait polaire total de *Mabs* R ont été séparés sur colonne de silice avec un gradient de méthanol jusqu'à 10% dans le chloroforme et analysés par CCM et spectrométrie de masse MALDI-TOF. Comme le montre la figure 76A, les fractions successives 23, 24 et 25 qui éluent durant un plateau à 10% de méthanol, n'ont pas le même profil de coloration après la révélation à l'orcinol-sulfurique et seules les spectres de masse des fractions 23 et 24 présentent l'ion de rapport *m/z* 1750 comme le plus intense (Figure 76B). La fraction 27 révèle une bande violette intense et son spectre de masse les ions caractéristiques à *m/z* 841 à 897 du 3,4 DAT. La fraction 33 correspond au 2,3 DAT.

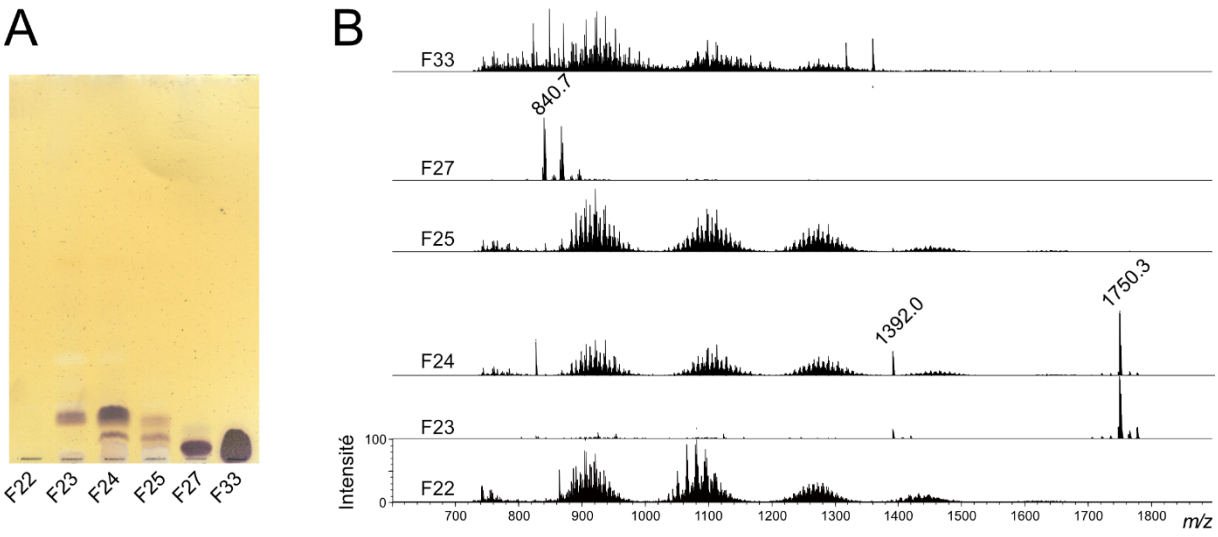


Figure 76 : Couplage flash/MALDI-TOF pour la recherche du GPL4690. Des fractions successives 22 à 33 ont été purifiées à parti d'un extrait polaire total de *Mabs* R, puis leur profil CCM a été révélé par l'orcinol-sulfurique (A). Les spectres MALDI-TOF (B) montrent l'ion *m/z* 1750 dans les fractions 23 et 24, au profil de coloration très différent.La fraction 27 présente les signaux *m/z* 841 et 869 attribués aux 3,4 DAT dans le chapitre IV.E.3.a.

De façon similaire, j'ai préparé les extraits polaires totaux des deux mutants *ΔMAB_0855* et *ΔMAB_4690c* ainsi que du mutant *ΔMAB_4690c* complémenté. Ensuite, j'ai séparé les trois extraits avec un gradient strictement identique par chromatographie-flash. Toutes les fractions obtenues ont été analysées par CCM et spectrométrie de masse MALDI-TOF (Figure 77). Alors que l'orcinol-sulfurique met en évidence une bande faiblement colorée dans la fraction 16 du mutant *ΔMAB_0855* (Figure 77A), cette dernière produit un ion *m/z* 1750 qui n'est pas détecté dans les fractions correspondantes du mutant *ΔMAB_4690c* et de son complémenté (Figure 77B). A la place, on observe une série d'ions peu intenses décalés de 74 u.m.a. correspondant probablement à un polymère non caractérisé ici. A l'inverse, la fraction 21 qui présente une bande de coloration violette pour les 3

souches (Figure 77A), possède bien les ions *m/z* 841 et 869 spécifiques du 3,4 DAT sur les 3 spectres MALDI-TOF.

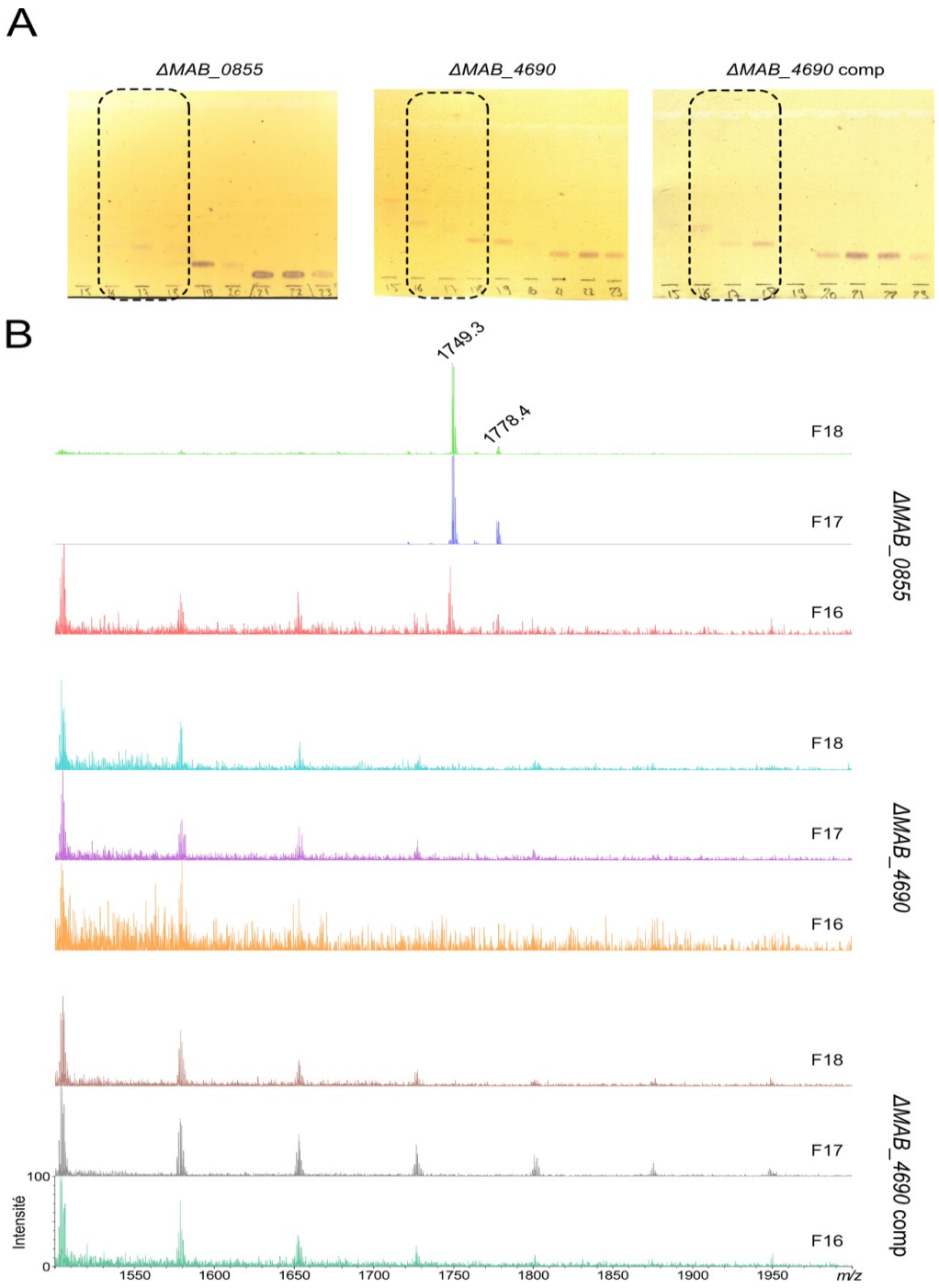


Figure 77 : Test de complémentation du GPL4690 par couplage flash/MALDI-TOF. Trois fractions successives 16 à 18 ont été purifiées à partir des extraits polaires totaux de trois mutants, puis leur profil CCM a été révélé par l'orcinol-sulfurique (A). Les spectres MALDI-TOF (B) montrent l'ion *m/z* 1750 dans les fractions 17 et 18 du mutant *ΔMAB_0855* contrôle, mais pas dans les mutants *ΔMAB_4690c* et son complémenté. Le 3,4 DAT a été détecté dans toutes les fractions 21 à un niveau équivalent.

Si la délétion du gène *MAB_4690c* entraine bien la disparition de l'ion de rapport *m/z* 1750 par rapport à la souche rugueuse *Mabs* R, la complémentation n'a pas restauré la production de ce dernier à un niveau détectable par l'analyse MALDI-TOF. Comparativement, la délétion du gène *MAB_0855* n'entraine pas la disparition de l'ion *m/z* 1750 dans des conditions de culture, de préparation et d'analyse strictement identiques avec le mutant Δ*MAB_4690c*.

III.E.7.b.iii Analyses LC-MS

L'utilisation du couplage LC-MS a été utile non seulement pour confirmer des informations structurales sur le motif glycannique, mais également pour quantifier relativement la présence du GPL4690 parmi les souches sauvages *Mabs* R ou *Mabs* S ainsi que chez deux mutants *Δ_MAB4690c* et *Δ_MAB4691c*.

Tout d'abord, j'ai développé une méthode afin de séparer le GPL4690 à *m/z* 1750 et les deux dérivés également individualisés à *m/z* 1778 et 1391 précédemment purifiées dans les fractions 60 à 68 (Figures 57B et 68-70). Une extraction de chacun de ces ions à partir du signal ionique total (TIC) montre que la fraction 66 contenant le GPL4690 présente un pic chromatographique intense à Tr = 28,1 min (Figure 78). La fraction 68 présente bien l'ion *m/z* 1778 qui élue à Tr = 32,2 min correspondant à la forme C_{20}, et la fraction 60 avec l'ion *m/z* 1391 à Tr = 13,1 min est la forme mono-rhamnosylée.

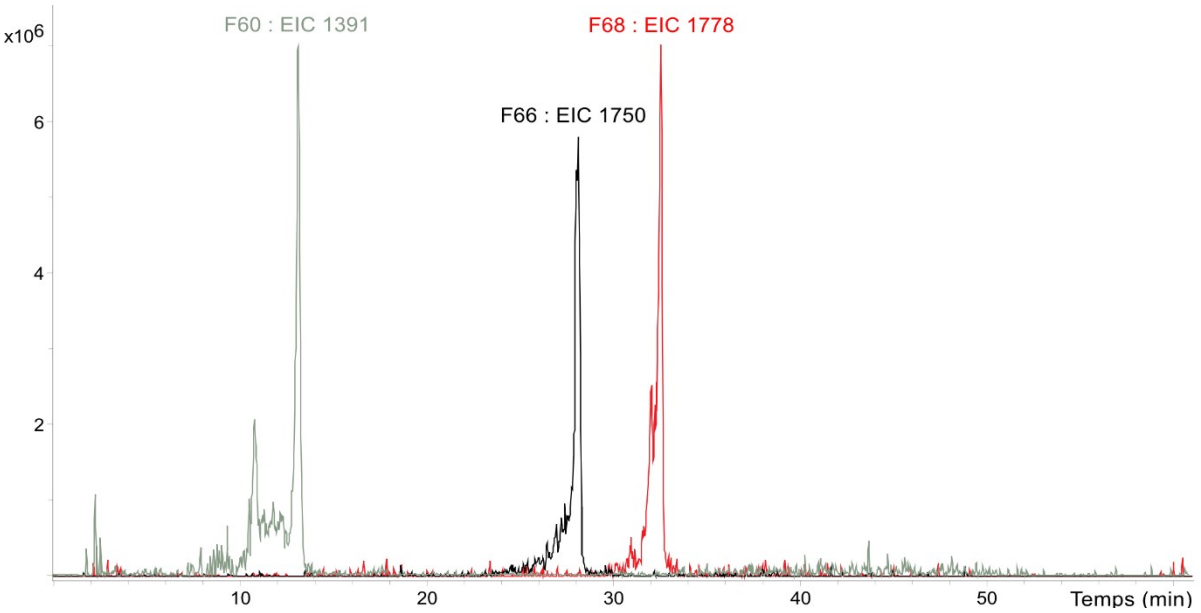


Figure 78 : Analyse LC-MS du GPL4690 et de ses dérivés. Les fractions doublement purifiées sur silice vierge puis silice greffée C_{18} (Figure 57B) ont été injectées dans les mêmes conditions de séparation HPLC et d'acquistion afin de repérer les temps de rétention du GPL4690 Tr = 28,1 min (noir) de sa forme C_{20} à Tr = 32,2 min (rouge) et de la forme mono-rhamnosylée à Tr = 13,1 min (vert). Chaque tracé est une extraction de l'ion *m/z* correspondant à partir du courant ionique total.

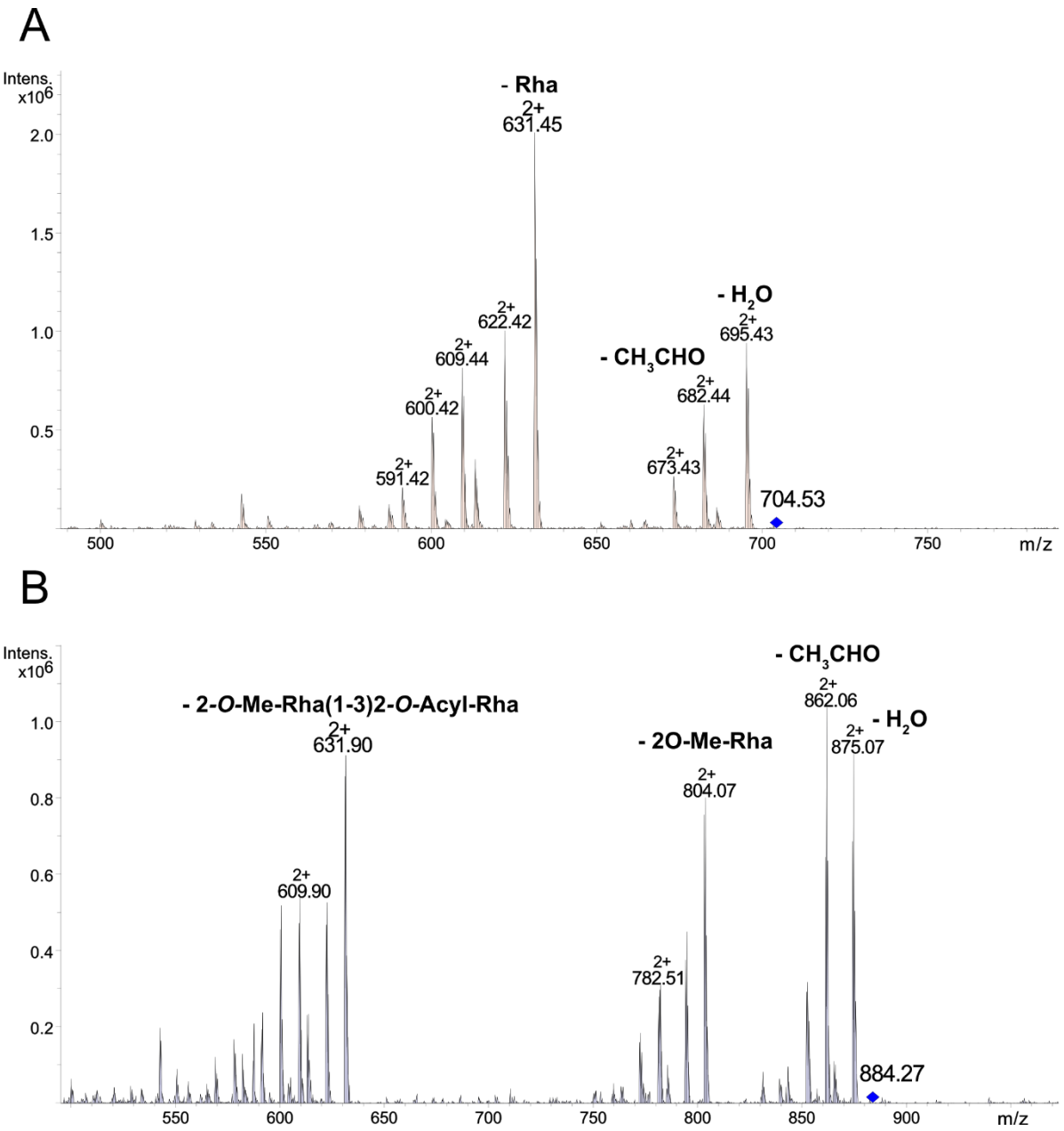


Figure 79 : Confirmation du motif glycannique du GPL4690 par LC-MS. A partir de la fraction 60, l'ion dichargé 704,5 a été sélectionné puis fragmenté en mode CID pour montrer la perte du déoxyhexose *via* la détection de l'ion 631,5 (A). De même, dans la fraction 66, le spectre MS^2 de l'ion dichargé 884,3 confirme la rupture possible des deux liaisons glycosidiques avec la perte du 2-*O*-méthyl déoxyhexose, puis celle du 2-*O*-hydroxydodecanoyl-rhamnose(B). Des pertes de molécules d'eau et de chaines latérales de thréonine s'observent également.

D'autre part, on note également sur les spectres de masse la présence d'adduits dichargés plus intenses que l'adduit sodium monochargé qui correspondent tantôt au $[M+2Na]^{2+}$ ou au $[M+H+K]^{2+}$. Dans le cas de la forme mono-rhamnosylée de la fraction 60, l'adduit $[M+H+K]^{2+}$ à *m/z* 704,5 produit sur le spectre MS^2 un fragment dichargé 631,5 correspondant bien à la perte du déoxyhexose (Figure

79A). Dans le cas du GPL4690 retrouvé dans la fraction 66, le spectre de l'ion parent *m/z* 884,5 génère un ion fils *m/z* 804,5 résultant de la perte du 2-*O*-Me Rha, et à nouveau le fragment *m/z* 631,5 résultant de la rupture de la deuxième liaison glycosidique (Figure 79B). On remarque ici que l'ionisation électrospray suivi d'une fragmentation CID à plus faible énergie met davantage en évidence la partie glycannique.

Ensuite, j'ai injecté les extraits polaires totaux des deux morphotypes, rugueux et lisse, de la souche sauvage ainsi que des mutants *Mabs* R *Δ_MAB4690c* et *Δ_MAB4691c*, afin de détecter les proportions relatives du GPL4690 *via* la recherche en continu de l'ion *m/z* 1750. Comme présenté dans la figure 80, l'intensité du pic chromatographique dans la souche sauvage *Mabs* R cultivé sur le milieu LB est au moins 1000 fois supérieure à l'intensité du pic dans la souche sauvage *Mabs* S alors qu'il est absent dans les deux mutants, suggérant que la mutation d'un des des deux gènes est suffisante pour entrainer la disparition du GPL4690. Dans ce cas, l'analyse des mutants complémentés n'a pas été effectuée à ce stade de l'étude.

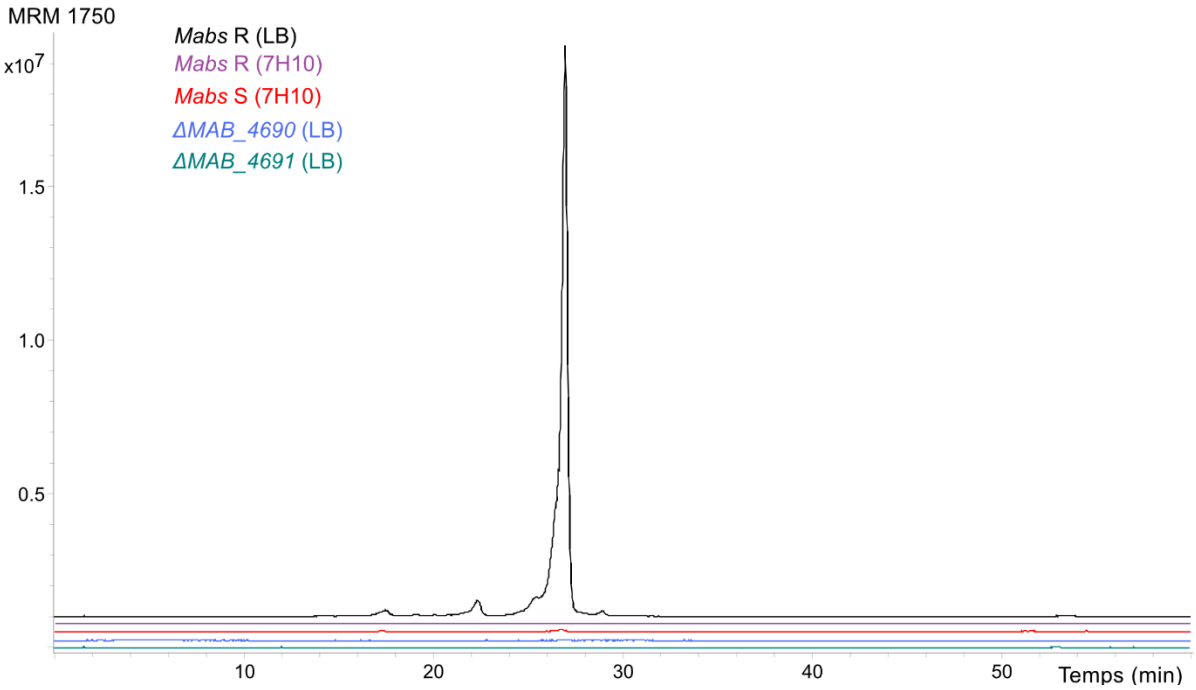


Figure 80 : Recherche LC-MS du GPL4690. La mesure continue de l'intensité de fragmentation de l'ion 1750 révèle que ce composé est le plus abondant dans la souche *Mabs* R cultivé sur milieu LB comparativement aux mutants *Δ_MAB4690c* et *Δ_MAB4691c,* à la souche *Mabs* S ou à la culture sur le milieu 7H10.

III.E.7.c Conclusion

D'après les analyses bioinformatiques menées sur le gène *MAB_4690c* et les gènes à son voisinage, nous avons établi que la fonction probable de ce locus (Figure 15C) est compatible avec la synthèse des variantes structurales du GPL4690 précédemment décrites (Figure 73). L'étude in-silico nous permet aussi de supposer que les résidus T_1, F_3, L_6 et T_7 sont potentiellement sous la forme D.

De plus, nous avons démontré au moyen de deux méthodologies analytiques que la délétion du gène *MAB_4690c* entraine bien la disparition de l'ion *m/z* 1750 dont la structure complète a été résolue dans le chapitre IV.E.6.f. Il en est de même lorsque le gène *MAB_4691c* est délété.

Les différences observées de niveau d'expression selon la souche et les conditions de culture pourront être quantifiées plus précisément grâce aux solutions développées, qui serviront également à la recherche approfondie d'intermédiaires de biosynthèse chez les mutants de glycosyltransférases et de méthyltransférases prévues dans la suite du projet ANR Sunlive.

IV DISCUSSION

L'objectif principal de mes travaux de thèse était la description exhaustive du contenu glycolipidique de l'enveloppe de la souche ATCC19977 de *Mycobacterium abscessus* pour expliquer leur implication dans les caractéristiques morphologiques et physicochimiques de sa paroi, en particulier son hydrophobicité, et dans la pathogénicité de cette espèce. En effet, des travaux de Jankute et collaborateurs sur les SGL de *Mtb* et les LOSs de *Mkan* suggèrent que l'hydrophobicité croissante de l'enveloppe s'explique par la perte des glycolipides les plus polaires, et qu'elle s'accompagnerait d'une possibilité accrue de transmission sous forme d'aérosol et d'une pathogénicité plus marquée[78].

Nous avons ainsi identifié dans la fraction apolaire Ac_2TMM qui est un nouveau glycolipide accompagnant le TDM. Ce dérivé de tréhalose et d'acide mycolique présente deux fonctions alcools secondaires acylées en position 2 et 3, au lieu d'une fonction alcool primaire en position 6 comme dans le TDM. De cette manière, il arbore un comportement plus apolaire lors de la séparation CCM et apporte un surplus d'hydrophobicité. Ce composé présente des caractéristiques structurales proches du 2,3 DAT qui s'est également accumulé en grande quantité dans la fraction polaire lorsque *Mabs* a été cultivée sur le milieu 7H10OADC. Dans ces deux composés, des lipides courts et polyméthylbranchés sont estérifiés en position 2 et 3 de la même unité glycosidique. Pour l'Ac_2TMM, l'analyse des spectres de fragmentation MALDI-TOF suggère la présence d'acides gras C_9 ou C_{12}, et l'analyse RMN met en évidence les ramifications successives de la chaine alkyle. C'est également le cas du 2,3 DAT qui possède des spectres RMN très similaires à l'Ac_2TMM dans la zone plus blindée, et dans lequel l'étude de la composition lipidique révèle la présence d'acide 2,4 diméthyl heptanoïque et 2,4,6 triméthyl nonanoïque, respectivement C_9 et C_{12}. Si un lien métabolique formel entre ces deux glycolipides n'a pas pu être établi, il est probable que le 2,3 DAT présent en grande quantité, entre en compétition avec les substrats naturels de l'une des 3 mycolyltransférases de type Ag85 qui produit alors l'Ac_2TMM retrouvé dans la fraction apolaire. Parmi les substrats naturels, le TMM de *Mabs* n'a pas pu être isolé en quantité suffisante pour une détermination structurale complète contrairement au TDM retrouvé dans la mycomembrane.

Le TDM présente une structure symétrique identique à celui de toutes les mycobactéries, mais contient uniquement des acides mycoliques non oxygénés. Bien que les formes éthyléniques *cis* et *trans* soient majoritaires, une faible proportion de groupements *cis*-cyclopropyl a été détectée. De façon surprenante, Rao et collaborateurs ont montré que la cylopropanation favorise la croissance de

Mtb, l'inflammation et la formation de granulomes chez la souris[393], alors que le passage vers une stéréochimie *trans* atténue l'effet proinflammatoire et l'apparition des granulomes[394]. La présence de groupements *cis*-cyclopropyl augmenterait donc la virulence de *Mabs*, et leur caractérisation dans l'Ac_2TMM, le TDM et les acides mycoliques méthylestérifiés à partir d'un culot cellulaire confirme le lien métabolique entre ces 3 composés ainsi qu'avec le complexe mAGP.

S'agissant des DAT, l'originalité du 2,3 DAT réside également dans la faible longueur des chaines d'acides gras comparativement aux dérivés mycocérosiques, mycosanoïques, mycolipanoïque ou mycolipènoïque de *Mtb* comptant plus de 28 carbones. D'après la littérature, ces derniers sont produits grâce à la mycocerosic acid synthase (mas) et au substrat méthylmalonyl-coA caractérisé notamment chez *Mbov*[395]. Néanmoins, des études réalisées sur un mutant *Δmas* montre qu'une 2ème enzyme, nommée short mycocerosic acid synthase (smas)[396], est impliquée dans l'élongation *in-vitro* de précurseur dont la longueur de chaine fut diminuée jusqu'à 12 carbones. S'il n'est pas impossible que cet enzyme utilise comme précurseur le methylmalonyl-coA, cela n'explique pas pourquoi son activité conduirait à des chaines plus courtes chez *Mabs*. En outre, la présence de branchements méthyles est un point commun avec le PDIM ou le SL-1, reconnus pour être partiellement responsables de la virulence de *Mtb*[397,398,282].

A nombre de carbones égal, il a été démontré que le branchement de la chaine alkyle diminuait l'hydrophobicité par rapport à une chaine linéaire[399]. Cela explique le comportement chromatographique plus apolaire du 3,4 DAT par rapport au 2,3 DAT en CCM, puisqu'il contient des acides gras linéaires saturés ou insaturés C_{16} et C_{18}. Dans le cas de ce deuxième DAT, l'originalité réside davantage dans la position des groupements acyles en position 3 et 4. De plus, la caractérisation d'un second dérivé métabolique apparenté au 3,4 DAT issu d'une souche clinique GD163 devenue lisse, montre que la proportion des acyl-tréhaloses pourrait influencer la morphologie des colonies lorsque les GPLs sont absents de l'enveloppe. Une acylation supplémentaire sur la position 6 de la même unité de glucose ou un allongement de la chaine lipidique confère cette fois une hydrophobicité plus forte. Néanmoins, l'impact de la structure du lipide, du nombre et de la position des acylations ainsi que de la quantité de l'ensemble des acyl-tréhaloses identifiés dans les fractions polaires et apolaires sur l'hydrophobicité globale n'a pas été directement évalué. Il faudrait pour cela identifier les voies métaboliques et les mécanismes de régulation de leur production.

L'étude de la fraction polaire a également mis en évidence la présence des PIMs qui sont essentielles à la croissance des mycobactéries. La recherche des intermédiaires de leur biosynthèse

par la spectrométrie de masse a révélé que leur précurseur phosphatidyl myo-inositol contenait un acide palmitique et un acide tuberculostéarique. D'autre part, la caractérisation des $AcPIM_2$ a révélé que le deuxième résidu de monosaccharide relié à la position 6 du *myo*-inositol est un 2-*O*-méthyl mannose, retrouvé également en grande quantité dans la composition monosaccharidique de la fraction polaire total. Conjointement à notre travail, ceux de Palcekova et collaborateurs démontraient une substitution identique sur une fraction déacylée de PIM_2[212]. En outre, l'enzyme responsable de l'ajout de ce groupement O-méthyl ainsi que son rôle n'ont pas été déterminé, même s'il est probable qu'il puisse influer sur le processus inflammatoire via la liaison avec le récepteur à mannose ou avec les récepteurs de la famille TLR.

L'ensemble des glycolipides précédemment décrits a été visualisé sur les profils CCM des deux morphotypes, contrairement aux GPLs qui se retrouvent en grande quantité dans la forme lisse de *Mabs*. L'étude exhaustive de cette famille de composés a pu se faire grâce à l'analyse du contenu en GPLs de mutants aux gènes du locus de leur biosynthèse.

En premier lieu, nous avons montré que la délétion du gène *fmt* entrainait la perte du groupement β-méthoxy sur le lipide N-acylé. Cette seule modification structurale discrète des GPLs induit une plus faible hydrophobicité de l'enveloppe, une diminution des forces d'adhésion de surface de cette dernière, une réduction de la phagocytose de *Mabs* S par les macrophages mais ne modifie pas le morphotype lisse[390]. D'autre part, à travers l'étude des mutants simples et multiples des gènes codant les trois glycosyltransférases *gtf1*, *gtf2* et *gtf3*, nous avons souligné l'implication de la glycosylation (Figure 55) dans divers rôles biologiques.

Tableau 25 : Rôles biologiques des GPLs de *Mabs*. D'après nos résultats, les modifications du GPL de *Mabs* ont des rôles biologiques au début du cycle infectieuxmais ils ne sont pas directement liés à la virulence.

Modification	Rôles biologiques
β méthoxylation	1. Hydrophobicité de l'enveloppe 2. Forces d'adhésion 3. Phagocytose
Glycosylation	1. Hydrophilicité de l'enveloppe 2. Formation de biofilm, motilité 3. Initiation (*gtf2*) 4. Phagocytose (*gtf3*)
Acétylation	1. Morphologie des colonies 2. Régulation de la glycosylation 3. Régulation de la méthylation

Tout d'abord, si la délétion seule du gène *gtf1* n'a pas empêché la production des GPLs puisque les GPL-1b et les GPL-2b ont été caractérisés, l'absence du 6 déoxy-talose relié au résidu de thréonine était associée à un basculement vers la forme rugueuse, suggérant l'implication de ce déoxyhexose dans le changement de morphologie. Ce résultat s'est confirmé avec le mutant *Δgtf1,3* également rugueux qui synthétisait toujours le GPL-1b. A l'inverse, en accord avec la composition en GPLs et le morphotype constaté chez *Msmeg*[400], la délétion du gène *gtf3* n'a pas entrainé de changement de morphotype bien que le GPL-3 disparaissait et que le GPL-2a s'accumulait. Afin de confirmer le rôle de ce gène dans la phagocytose de *Mabs* par les macrophages, nous avons démontré non seulement que les billes fluorescentes recouvertes de GPL-3 sont davantage phagocytées par les macrophages THP-1 comparativement aux autres glycoformes, mais aussi que la surexpression du gène *gtf3* chez *Mabs* S entraine une augmentation de la phagocytose du bacille. Parmi les récepteurs impliqués dans ce processus, l'intervention du récepteur à mannose se justifierait par la reconnaissance du rhamnose terminal porté par le GPL-3 dont les hydroxyles en position 2, 3 et 4 sont dans la même configuration que le mannose. Ces résultats confortent ceux obtenus avec la souche ATCC 607 de *Msmeg* par Villeneuve puisqu'elle a démontré que les GPLs triglycosylés étaient capables d'interférer efficacement avec la phagocytose des bacilles de *Msmeg*, *Mkan* et *Mtb*[345]. Même si le rhamnose terminal est modifié en position 2 par un groupement succinyl, le récepteur à mannose a également été incriminé[346].

Dans le cas du gène *gtf2*, sa délétion a perturbé l'ensemble de la glycosylation car aucun GPL n'a été détecté dans ce mutant qui présentait également un morphotype rugueux. Son activité enzymatique dans l'ajout du premier résidu de rhamnose sur l'alaninol a plutôt été mis en évidence par sa seule expression dans le double mutant *Δgtf1,3* qui aboutit à l'accumulation du GPL-1b, ainsi que par l'assignation des activités des gènes *gtf1* et *gtf3* dans l'ajout du 6 déoxy-talose et du rhamnose terminal. De façon plus surprenante, lorsque le gène *gtf2* n'ajoute pas le premier rhamnose, nous n'avons pas mis en évidence d'accumulation aussi importante que pour le GPL-1b de la seconde forme mono-glycosylée qui porte uniquement le 6 déoxy-talose relié à la thréonine. Lors des étapes de la glycosylation, cela démontre que le gène *gtf2* initie la glycosylation des GPLs, et que l'action de *gtf1* nécessite la présence du premier rhamnose. Un enchainement similaire des deux premières étapes de glycosylation par les orthologues respectifs de *gtf2* puis *gtf1*, nommés *GtfB* puis *GtfA*, avait été proposé chez *Mav* par Eckstein en 2003[392]. Entre ces deux étapes, l'auteur indiquait que les méthyltransférases codées par *MtfC* et *MtfD* ajoutaient les groupements *O*-méthyl en position 3 et 4, dont les orthologues annotés chez *Mabs* sont les gènes *MAB_4105c* et *MAB_4108c*. N'ayant pas

étudié le rôle de ces deux gènes directement chez *Mabs*, nous postulons que leur activité enzymatique est bien la même. De cette manière, la reconnaissance du motif 3,4 di-O-méthyl rhamnose pourrait favoriser l'utilisation du susbtrat GPL-1b par le produit de *gtf1*. En outre, l'analyse approfondie du mutant *gtf2* a permis d'identifier le lipopeptide non glycosylée dont la structure a peut présenter un valinol au lieu de l'alaninol du côté C-terminal. Conjointement, les travaux de Wiersma ont montré la présence des mêmes résidus accompagnés de leucinol dans les GPLs de souches cultivées sur un milieu synthétique SCFM2 ou celui mimant les conditions d'infection[349]. Ces résultats signifient que le produit de *gtf2* est suffisamment flexible pour agir sur des substrats lipopeptidiques variables dont l'amino-alcool branché côté C-terminal diffère, et souligne le rôle central de cet enzyme dans l'élaboration du motif glycannique.

Dans un dernier temps, nous avons étudié l'activité des acétyltransférases potentiellement impliquées dans la synthèse des GPLs. S'agissant des acétyltransférases Atf1 et Atf2, la duplication du gène *atf* de *Msmeg* chez *Mabs* est à l'origine d'une régulation plus complexe de la biosynthèse des GPLs de ce dernier. En effet, si la délétion du gène *atf* entraine bien un défaut d'acétylation des GPLs chez *Msmeg*[343,348] nos résultats démontrent que la délétion simple ou simultanée des gènes *atf1* et *atf2* induisait une faible accumulation de GPLs mono-*O*-acétylés mais n'empêchaient pas la synthèse des GPL di-*O*-acétylés chez *Mabs*. Cette observation nous a conduit à identifier deux nouvelles acétyltransférases codées par les gènes *MAB_1725c* et *MAB_3448* qui se retrouvent hors du locus de biosynthèse des GPLS. L'intervention des produits de ces deux gènes et leur activité métabolique dans l'acétylation des GPLs est déduite de la diminution puis la disparition de GPLs di-*O*-acétylés chez les mutants multiples *Δatf1, atf2, MAB_1725c et Δatf1, atf2, MAB_3448* et *Δatf1, atf2, MAB_1725c, MAB_3448*. Un effet collatéral et inattendu constaté dès la délétion unique du gène *atf2* est la diminution du niveau d'expression des GPL-2a et l'augmentation de la quantité de GPL-3. Pour expliquer ce phénomène, l'une des hypothèses est qu'Atf2 aurait une affinité supérieure pour le GPL-2a par rapport à Gtf3 et que la rapide acétylation du GPL-2a provoquerait une moindre activité de Gtf3, régulant ainsi la proportion de GPL-2a et de GPL-3 dans l'enveloppe de *Mabs* S. Chez le mutant triple *Δatf1, atf2, MAB_1725c* et le mutant quadruple *Δatf1, atf2, MAB_1725c, MAB_3448*, cette déstabilisation de la quantité relative de GPL-2a par rapport à celle de GPL-3 conduit à l'apparition de nouveaux intermédiaires métaboliques mono-*O*-acétylé sur le 6 déoxy-talose ou mono- et tri-*O*-méthylés sur le rhamnose terminal. Néanmoins, aucune différence de phagocytose entre la souche sauvage et les mutants n'a été observée. Comme le GPL-3 est celui qui contribue le plus à la phagocytose du bacille, il est probable que les groupements méthyles supplémentaires empêchent la

liaison avec les récepteurs à mannose et CR-3 qui participaient à la phagocytose des billes recouvertes de GPL-3 ou à la souche surexprimant le gène *gtf3*.

L'aspect des colonies a davantage été impacté par le défaut d'expression des gènes codant les quatre acétyltransférases. Si les délétions simples ont eu peu d'influence, tous les mutants multiples délétés à minima des gènes *atf1* et *atf2* présentaient une morphologie qui tendait graduellement vers celui du morphotype rugueux. Ces résultats sont en accord d'une part, avec les constatations de Recht et collaborateurs sur le mutant d'insertion *Δatf* de *Msmeg* qui présentait un aspect plus rugueux accompagné d'une moindre motilité et d'une moindre capacité à former des biofilms[343], et d'autre part, avec l'aspect rugueux de nos mutants de glycosyltransférases excepté *Δgtf3*. Chez tous les autres mutants, l'absence du 6 déoxy-talose au sein des GPLs identifiés implique l'absence des groupements acétyles et l'observation d'un morphotype rugueux comme c'est le cas lorsque les GPLs sont totalement déacétylés dans le mutant *Δatf1, atf2, MAB_1725c, MAB_3448*.

Pour conclure, la description de la structure du GPL4690 confirme l'existence d'un nouveau GPL propre à *Mabs* R (Figure 81). En effet, la mise en évidence du gène *MAB_4690* dans la virulence de *Mabs* R chez la souris a amené les équipes du Pr Jean-Louis HERRMANN, du Dr Laurent KREMER et du Dr Yann GUERARDEL à tenter d'identifier le produit de ce gène. Les développements chromatographiques ont permis d'isoler ce composé et de séquencer le noyau octapeptidique Thr-Ala-Phe-Ile-(OMe)Tyr-Leu-Thr-Thr par analyse RMN. La présence du résidu para-*O*-méthyl tyrosine a été confirmée par le décalage de masse de 177 a.m.u. entre les fragments a_4 et a_5 détectés sur le spectre de fragmentation MALDI-TOF. Elle constitue une première originalité structurale qui n'avait jamais été décrite dans la littérature chez les mycobactéries, même si on retrouve un résidu de β-hydroxy para-*O*-méthyl tyrosine dans un peptide cyclique de chez *Streptomyces* sp., nommé skyllamycine[401]. Chez cette bactérie, l'activité *O*-méthyl transférase dépendante du donneur S-adénosyl méthionine est assurée par le produit du gène *Sky37*[402].

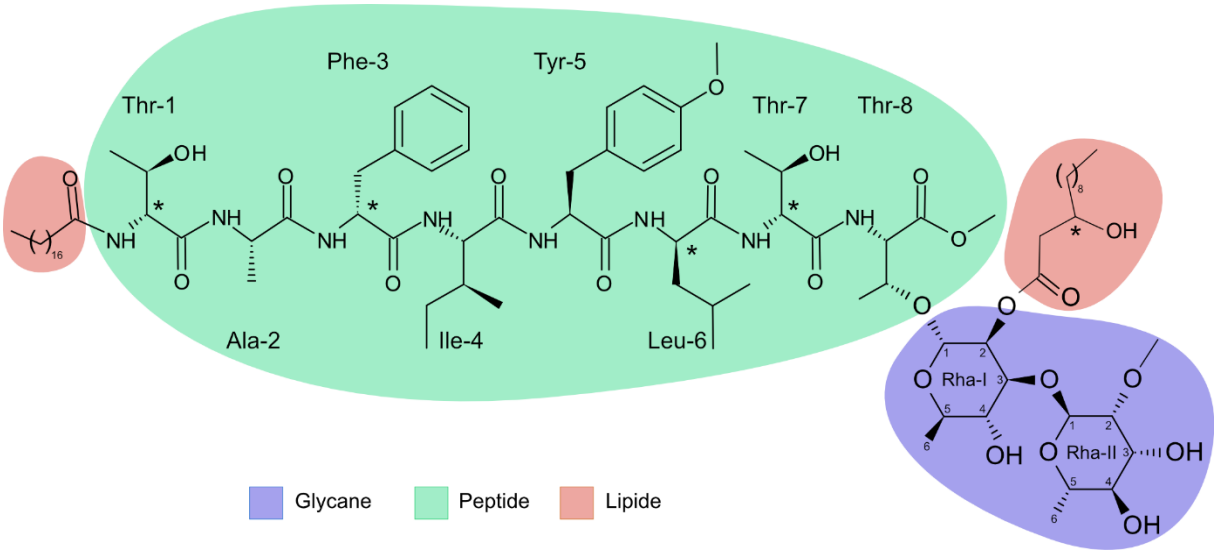


Figure 81 : Structure proposée du GPL4690. Les astérisques bleus indiquent les éléments de structure restant à vérifier : la configuration D/L des acides aminés a été prédite à partir des épimérases codées par les gènes *MAB_4690c* et *MAB_4691c*, la configuration L présentée pour les rhamnoses respecte la littérature sur les mycobactéries, la position du groupement hydroxyle sur l'acide laurique n'a pas été déterminée.

De même, l'analyse RMN du motif glycannique a révélé la présence de deux résidus de rhamnoses, confirmés ensuite par GC-MS. Le premier est lié au résidu de thréonine côté C-terminal, et porte le second via une liaison glycosidique α(1-3). En outre, la position 2 du premier rhamnose est acylée, tandis qu'elle est méthylée pour le second rhamnose. Cet arrangement du glycopeptide s'avère cohérent avec les prédictions de la nature et du nombre d'acides aminés utilisés par les produits des gènes *MAB_4690c* et *MAB_4691c*, ainsi qu'avec la proximité des glycosyltransférases codés par *MAB_4694c* et *MAB_4695c* qui possèdent une forte similarité avec *gtf1*, *gtf2* et *gtf3*. En outre, la présence de quatre épimérases parmi les modules suggère que certains acides aminés basculent de la configuration L à la configuration D. Ce mécanisme est largement connu chez les peptides synthases non ribosomales bactériennes car il confère l'avantage d'échapper à l'hydrolyse de la chaine peptidique[403]. En tenant compte de la position des épimérases, nous proposons à ce stade que T_1, F_3, L_6 et T_7 soient en configuration D, mais cette prédiction nécessite une confirmation par l'analyse stéréochimique de chaque acide aminé. De plus, dans le cas de la thréonine, le carbone asymétrique en position β est en configuration *allo* dans les GPL-2a et GPL-3 de *Mabs*, cette caractéristique restant à vérifier pour le GPL4690.

Concernant la partie lipidique, la présence majoritaire d'acide stéarique résulte de l'interprétation des fragments produits par l'ion m/z 1750 qui augmentent de 28 a.m.u. dans le cas de l'ion m/z 1778 alors que les fragments a restent inchangées. De plus, cet acide gras est présent dans

l'analyse de lipides faite par GC-MS. La nature du lipide acylé sur le premier rhamnose a été déterminée de façon indirecte par la différence de masse des deux GPL observés à *m/z* 1750 et *m/z* 1391. Sur la base de l'analyse de fragmentation en MALDI-QIT-TOF et de la mesure de masse précise en MALDI FT-ICR, ce dernier possède le même noyau octapeptide, mais il perd six atomes d'oxygène et deux insaturations, correspondant au 2-*O*-Me rhamnose terminal et à l'acide hydroxy laurique. Néanmoins, la position de la fonction hydroxyle reste à définir.

La présence conjointe des deux GPLs mono- et di-rhamnosylées suggère que l'activité de l'une des deux rhamnosyltransférases probablement codées par *MAB_4694c* et *MAB_4695c* n'est que partielle, comme constaté dans le cas de Gtf3. L'absence du groupement hydroxy-lauryl dans l'intermédiaire mono-rhamnosylé questionne le lien avec la présence du second rhamnose. La modification de ce dernier par l'une des deux méthyl transférases codées par les gènes *MAB_4696c* et *MAB_4700c* est à confirmer. L'un de ces deux gènes pourrait aussi contribuer à la modification de la tyrosine. De même, le transport du GPL4690 par les produits des gènes *MAB_4703c* à *MAB_4705c* devra être vérifié. Finalement, nous postulons que la présence de la thréonine T_8 après l'action de *MAB_4690c* est nécessaire à la glycosylation. D'ailleurs, la délétion de ce gène a bien entrainé la disparition du GPL4690. A ce stade nous n'avons pas retrouvé d'intermédiaire dans le mutant *ΔMAB_4690c* résultant de l'expression seule de *MAB_4691c* De plus sa délétion entraine également la disparition du GPL4690. Le rôle probable des différents gènes dans la biosynthèse du GPL4690 est résumé à la figure 82. Chacun de ses gènes pourra être sélectivement interrompu afin de déterminer de manière certaine leurs rôles dans les étapes clés de la biosynthèse du GPL4690 comme cela a été réalisé pour les GPLs.

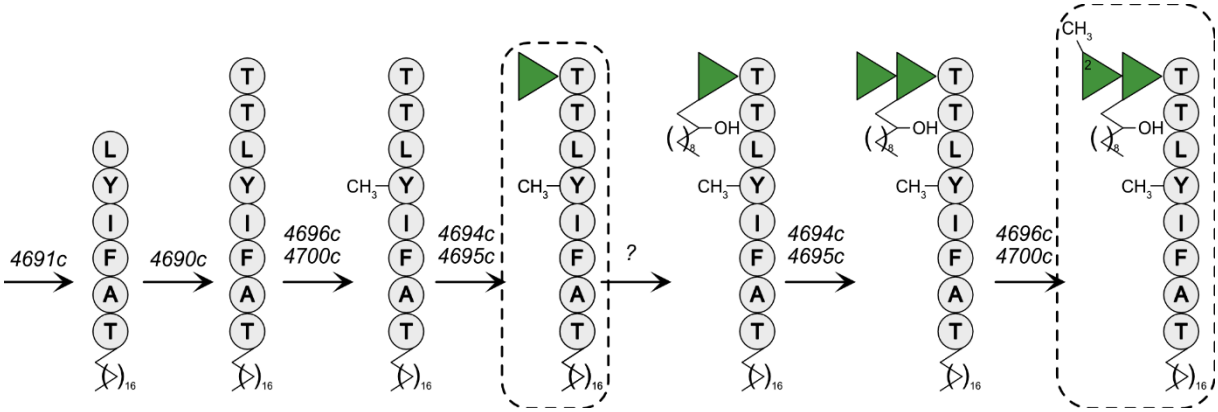


Figure 82 : Schéma hypothétique de biosynthèse du GPL4690. D'après les structures identifiées entourées et l'analyse bioinformatique des gènes au voisinage de *MAB_4690c*, les gènes responsables des éventuelles étapes de biosynthèse sont indiqués.

Seules trois études récentes citent le gène *MAB_4690c* et ses voisins. La première démontre, que son expression est fortement diminuée dans la souche 19977-IP-R, comparativement aux souches 19977-AT-R et CF-R, comme pour les gènes *mps1* et *mps2* du locus de biosynthèse des GPLs. C'est pourquoi, les auteurs suggèrent un mécanisme commun de régulation transcriptionnelle[174]. La seconde paru en 2021, concerne le suivi sur quatre ans de l'évolution génétique des souches cliniques isolées de patients atteints de mucoviscidose [404]. Les auteurs montrent que *Mabs* subit très fréquemment des mutations non silencieuses dans les gènes codant un régulateur transcriptionnel de la famille *TetR*, et impliqué dans la résistance à la clarithromycine, ainsi que dans le gène *embC* contribuant à la biosynthèse du LAM. Plus ponctuellement, ils constatent aussi ces mutations dans le locus de biosynthèse des GPLs ainsi que dans le gène *MAB_4690c*. Enfin, une dernière étude sur la pression de sélection génétique et les liens d'épistasie existant dans le génome de *Mabs*[405], confirme que la mycobactine contribue à une meilleure survie dans les macrophages. Dans cette étude, les auteurs établissent un lien entre le système de sécrétion codé par le gène *MAB_0471c* et la peptide synthase non ribosomal codée par le gène *MAB_3317c* qui participent à la virulence chez la Drosophile. En outre, la peptide synthase codée par le gène *MAB_4691c* a subi une pression de sélection quantitativement proche de *MAB_3317c*, qui n'a pas été validée par une étude de phénotype. De plus, le signal pour le gène *MAB_4690c* est moindre que celui détectée pour *MAB_4691c*. A ce stade, le lipopeptide correspondant à la seule expression de ce gène n'ayant pas été identifié chez le mutant *MAB_4690c* de même que la fonction biologique et les mécanismes de virulence chez la souris des deux glycoformes mises en évidence pour le GPL4690.

Outre le lien avec la pathogénicité de *Mabs* qui reste à élucider, la présence d'un GPL avec des spécificités structurales uniques laisse entrevoir des possibilités de détection et de diagnostic plus précis des infections non tuberculeuses. A ce jour, seul un test commercial de dosage des IgA a été mis en œuvre pour détecter les GPL non spécifiques communs à *Mav* et à *Mabs* chez les patients[406]. Néanmoins, une meilleure spécificité de détection propre à *Mabs* s'avérerait utile pour mieux adapter le traitement antibiotique. Dernièrement, la purification d'une fraction lipoprotéique activatrice de TLR2 à partir de *Mabs* R a permis de déterminer le titre en IgA de patients atteints de mucoviscidose[136], mais le grand nombre de protéines identifiées dans cette fraction entretient le doute sur la cible exactement reconnue[407]. L'amélioration de la spécificité du test a été obtenue en croisant les résultats avec le titrage des IgA dirigés contre la phospholipase C[136]. Si le GPL4690 se révélait un antigène également spécifique, il pourrait s'avérer plus simple à produire en grande quantité et en pureté suffisante pour développer de telles stratégies diagnostic. La recherche de la séquence peptidique

TAFIYLTT sur le serveur UniProt (https://www.uniprot.org/peptide-search/PM20221220195686216ed9458cab69f0be1ce7d9f4/overview) ne renvoie que trois protéines présumées appartenant au genre *Dyadobacter*, alors que les requêtes faites sur les bases de données « Peptide Atlas » ou « Antimicrobial Peptides Database » sont infructueuses.

V MATERIELS ET METHODES

Les protocoles d'extraction et d'analyse CCM se sont principalement inspirés du travail de Besra[389]

V.A EXTRACTION DES GLYCOLIPIDES

A partir d'un culot lyophilisé de 50 mg de mycobactéries, on ajoute 2 mL de méthanol/0,3 % NaCl (10:1) puis 1 mL d'éther de pétrole. On agite 15 min à température ambiante à l'aide d'un barreau aimanté, sur une plaque tournante à 800 rpm. On centrifuge 5 min à 3000 rpm, et on récupère la phase supérieure éther de pétrole. Le culot cellulaire est dispersé dans la fraction méthanolique, et on répète deux fois l'extraction par 1mL d'éther de pétrole. Les trois fractions obtenues d'éther de pétrole sont séchées sous azote, et les lipides apolaires sont solubilisés dans 300 µL de dichlorométhane. Ensuite, la fraction méthanolique est chauffée 5-10 min à 100°C sous reflux puis refroidie à température ambiante. On y ajoute 2,3 mL de chloroforme/méthanol/0,3 % NaCl (9:10:3) et on agite 1h à 800 rpm. Le mélange est centrifugé 5 min à 3000 rpm, et le surnageant est prélevé. On rajoute enfin 750 µL de chloroforme/méthanol/0,3 % NaCl (5:10:4), et on agite 30 min à 800 rpm. Après une 2ème centrifugation 5 min à 3000 rpm, le surnageant est regroupé avec le précédent dans une ampoule à décanter. On y rajoute 1,3 mL de chloroforme puis 1,3 mL de 0,3 % Nacl. La démixtion s'opère et on mélange vigoureusement l'ampoule à décanter durant 5 min, pour finalement récupérer la phase inférieure chloroformique après décantation. Les lipides polaires sont séchés sous azote puis repris dans 300 µL de chloroforme/méthanol (2:1). La phase aqueuse supérieure n'a pas été étudiée. Les volumes sont adaptés à la quantité de bactéries lyophilisées.

V.B Chromatographie couche mince (CCM)

V.B.1 Analytique

Sur une plaque de gel de silice 5-17µm ALUGRAM Xtra SIL (MACHEREY NAGEL GmbH, Duren, Allemagne), on dépose 5, 10 ou 20 µL d'échantillon sur une bande pouvant aller de 5 à 10 mm. Après migration à 4°C, on sèche la plaque sous air chaud et on réalise une migration supplémentaire si nécessaire au-dessus du front de solvant précédent pour améliorer la séparation. Les systèmes de solvant utilisés sont indiqués dans le tableau XX. La cuve de migration est saturée durant 4 heures à 4°C avant toute migration.

V.B.2 Préparative

Dans le cas d'une CCM préparative, 150-300 µL d'échantillon sont déposés en une ligne de 140-160 mm sur une plaque de verre 20cm x 20cm recouverte de gel de silice 60µm déposée dans une cuve contenant le solvant de migration. De même, une 2ème migration peut être effectuée. Un cache recouvert d'aluminium est placé sur la silice afin de ne révéler que les bords de la plaque sur 2 cm.

V.B.3 Révélation

La révélation de la plaque s'effectue en pulvérisant de l'orcinol dans l'acide sulfurique 20% ou de l'acide phosphomolybdique (PMA) 10% dans l'éthanol. La plaque est alors chauffée entre 150°C et 200°C jusqu'à la coloration des bandes de migration. Dans le cas de la chromatographie préparative, la plaque de verre est placée dans une enceinte contenant des granules de diiode, afin de repérer les bandes sur l'ensemble de leur longueur. Après décoloration la silice des bandes repérées est grattée et les glycolipides sont réextraits pendant 1h aux ultrasons dans le mélange chloroforme/méthanol (2:1).

V.C Purification des glycolipides par chromatographie-flash

Les extraits apolaires et polaires ou les fractions purifiées sont mélangés avec de la silice vierge ou greffée C_{18} selon la séparation choisie. Pendant une nuit complète, le mélange est séché à température ambiante sous hotte ventilée. La silice est ensuite collectée et introduite dans une colonne dédiée à l'injection du dépôt sec. Après équilibration de la colonne de séparation, La silice imprégnée est placée en amont de celle-ci.

V.C.1 Séparation par adsorption

Une colonne de 4g de gel de silice de granulométrie 15µm (INTERCHIM, Montluçon, France) est équilibrée avec 100% de chloroforme. Un gradient linéaire démarre sur le système de chromatographie flash Puriflash (INTERCHIM, Montluçon, France) pour atteindre jusqu'à 50 % de méthanol (tableau 26). Des fractions de 5 à 7 mL sont collectées. Après séchage sous vide, leur contenu est vérifié par coloration directe à l'orcinol ou au PMA, puis par CCM, voire MALDI-TOF. Les fractions d'intérêt sont regroupées, séchées sous azote et solubilisées dans le chloroforme/méthanol (2:1).

Tableau 26 : Programme général de gradient pour la séparation des glycolipides par chromatographie d'adsorption sur gel de silice vierge 15µm, le débit est de 5mL/min sur une colonne de 4g. Les pourcentages de méthanol sont indiqués pur chaque glycolipide caractérisé.

Temps (min)	0	5	20	30	45	55	70	85
$CHCl_3$ (%)	100	100	90	90	70	70	50	50
CH_3OH (%)	0	0	10	10	30	30	50	50
Glycolipides apolaires	Ac_2TMM (5%)							
	TDM (10%)							
Glycolipides polaires	GPL-2a (5-7%) ; GPL-3 (7-10%)							
	GPL-4690 (10%)							
	3,4 DAT (15%)							
	2,3 DAT (20%)							
	$AcPIM_2$ (25-30%)							
	PI,PIM_1,PIM_2 (40%)							
	$AcPIM_6$ (50%)							

V.C.2 Séparation phase inverse

Dans le cas du GPL4690, une colonne de 6g de silice greffée C18 et de granulométrie 50µm est équilibrée avec H_2O/CH_3OH 0,1% TFA (1:1) pendant 5 min à 5mL/min. Après injection, le gradient démarre pour atteindre 100% de méthanol 0,1% TFA après 60 min. Cette proportion est maintenue pendant 10 min. Des fractions de 5 à 7 mL sont récupérées pendant le gradient. Après séchage sous azote à 50-70°C, leur contenu est vérifié par coloration directe à l'orcinol-sulfurique, puis par CCM, voire par MALDI-TOF et RMN. Les fractions d'intérêt sont regroupées, séchées sous azote et reprises dans le chloroforme/méthanol (2:1).

V.D Dérivation chimique

V.D.1 Saponification des glycolipides (β-élimination)

Après séchage sous azote de l'extrait ou de la fraction purifiée, on ajoute de la soude 0,1M dans le méthanol/chloroforme (2:1). On place 2h à l'étuve à 37°C. On ajoute 1 mL de chloroforme puis 1 mL d'eau. On agite 30 sec à 1 min au vortex puis on centrifuge 30 sec. On récupère la fraction inférieure chloroformique et on la sèche sous azote pour la reprendre dans le chloroforme/méthanol (2:1).

V.D.2 Méthanolyse (transestérification)

On place au dessicateur pour la nuit un aliquot de l'échantillon. On y ajoute 200 µL de méthanol/HCl 0,5-3N , et on le place 4-24h à l'étuve à 80°C. On ajoute 1 mL de chloroforme (ou heptane) puis 1 mL d'eau. On agite 30 sec à 1 min au vortex puis on centrifuge 30 sec. On élimine l'eau et on répète le lavage avec 1mL d'eau ultrapure. On récupère la fraction inférieure chloroformique et on la sèche sous azote en deçà de 50°C pour reprise dans chloroforme/méthanol (2:1) ou le chloroforme pur.

V.D.3 Hydrolyse alcaline suivie d'une méthyl-estérification

L'échantillon est hydrolysé par 15% de tétrabutyl-ammonium hydroxude pendant 12 à 24h à 100°C, puis méthyl-estérifiée 1 h aux ultrasons après ajout de 150µL d'iodométhane (dérivé très toxique à manipuler avec précaution : hotte + EPI), 2mL de dichlorométhane et 1mL d'eau ultrapure. La phase organique est prélevée, séchée sous azote et lavée plusieurs fois par un mélange chloroforme/eau (1:1) avant analyse CCM ou injection en GC.

V.D.4 Obtention des dérivés itols-acétates des monosaccharides

Avant hydrolyse de l'échantillon par 1mL d'ATFA 4M à 100°C, on ajoute à l'échantillon 1µg de mésoinositol comme standard interne. On sèche alors l'échantillon sous azote à 80°C puis on réserve au dessicateur pour la nuit. Le lendemain, on ajoute 500 µL de $NaBH_4$ 10 mg/mL dans NH_4OH 2M, et on incube 4h à température ambiante. On arrête la réaction par ajout d'acide acétique glacial concentré jusqu'à l'arrêt du dégagement gazeux. On vérifie que le pH est bien acide, et on sèche à 55°C sous flux d'azote en le co-distillant avec du méthanol à 5 reprises. A nouveau, on laisse au dessicateur toute la nuit. Finalement, on ajoute 500 µL d'anhydride acétique pur 4h à 80°C. Pour les lavages, on ajoute 1,5 mL de chloroforme et 1 mL d'eau ultrapure. On agite au vortex 1min, on centrifuge puis on élimine la phase supérieure aqueuse. On lave 4 fois la phase chloroformique avec 1 mL d'eau ultrapure. On ajoute du sulfate de sodium dans la phase chloroformique pour éliminer les traces d'eau, on filtre sur laine de verre et on sèche sous azote en deçà de 50°C. On reprend dans le chloroforme avant injection en GC.

V.D.5 Obtention des dérivés perméthylés du tréhalose

Un aliquot de glycolipides est solubilisé dans le DMSO anhydre auquel est ajouté le même volume d'iodométhane (dérivé très toxique à manipuler avec précaution : hotte + EPI) ainsi qu'une spatule de soude en poudre. Le mélange est placé au bain à ultrasons pendant 2h, et la réaction est

arrêtée par l'ajout d'eau ultrapure à 4°C. Les composés perméthylés sont extraits par l'ajout de chloroforme lavé plusieurs fois par de l'eau ultrapure en proportion (1:1).

V.E CHROMATOGRAPHIE GAZ

V.E.1 Dérivés itols-acétates

1 µL d'échantillon est injecté en mode splitless sur une colonne capillaire Solgel 1MS 30mx0,25mmx0,25µm. Le gradient de température démarre de 120°C jusqu'à 230°C à 3°C/min, puis à 270°C à 10°C/min. Les itols acétates sont détectés après ionisation électronique à 70 eV sur un simple quadripôle 5976 (AGILENT TECHNOLOGIES, Santa Clara, Etats-unis) en mode « full-scan » sur une gamme de masse allant de 50 Da à 500 Da. Des standards de rhamnose, 6-déoxy talose, mannose, glucose et de rhamnoses O-méthylés isolés de spores de *Bacillus cereus* ont été dilués dans les échantillons pour confirmer les temps de rétention et les fragmentations.

V.E.2 FAMEs et dérivés triméthylsilylés

Les esters méthyliques d'acides gras sont séparés sur un système TRACE 1300 (THERMO FISHER SCIENTIFIC, San Jose, Etats-Unis). 1 µL est injecté en mode splitless sur une colonne capillaire SLB-5MS 30mx0,25mmx0,25µm. 5 minutes après l'injection, le gradient de température démarre de 50°C jusqu'à 140°C à 10°C/min, puis jusqu'à 340°C à 15°C/min. Les dérivés méthyliques d'acides gras sont détectés spécifiquement en mode « selected ion monitoring » des ions 74 et 87 Da après ionisation électronique à 70 eV, en même temps que le mode « full scan » de 50 Da à 700 Da sur un simple quadripôle ISQ (THERMO FISHER SCIENTIFIC, San Jose, Etats-Unis).

Les dérivés triméthylsilylés ont été produits à partir des dérivés méthyliques d'acides gras par ajout de 100 µL de BSTFA et 100 µL de pyridine. La réaction s'est déroulée à température ambiante durant 2h avant d'ajouter 300 µL d'heptane. Les dérivés triméthylsilylés ont été séparés avec le même programme de température sur la même colonne et détectés en mode full scan. Un standard méthyl 3-hydroxy octadecanoate (MATREYA LLC, PA, Etats-unis) a été soumis à la réaction afin d'obtenir le spectre de fragmentation EI-MS.

V.F SPECTROMETRIE DE MASSE

V.F.1 MALDI QIT-TOF

Dans un micro-insert de 250 µL, on mélange 5 µL d'acide dihydroxy benzoïque (DHB) 10mg/mL dans le chloroforme/méthanol (1:2) avec 5 µL d'échantillon dans le chloroforme/méthanol (2:1). On

dépose 1 µL sur la cible MALDI. L'acquisition MS ou MS^n se fait sur un système Axima Résonance de la société (SHIMADZU, Kyoto, Japan). L'énergie du LASER est optimisée pour chaque échantillon entre 83 eV et 123 eV, le nombre d'acquisition par spectre est de 300, l'analyseur temps de vol est en mode réflectron, et la gamme de masse sélectionnée va de 50 Da à 3000 Da. L'étalonnage externe est fait sur un mélange de glucidex (polymère de glucoses) couvrant cette gamme de masse. Pour l'analyse MS^2, la sélection de l'ion parent s'est faite à une résolution de 250 et la collision avec une tension entre 300 V et 600 V. Dans le cas de la MS^3, la sélection de l'ion s'est faite avec une résolution standard de 70 et avec une tension variant de 300 V à 600 V.

V.F.2 MALDI et ESI FT-ICR

L'analyse MALDI-FT-ICR s'est faite en mélangeant la matrice DHB 10mg/mL, l'échantillon et l'étalon interne $PEG_{22\text{-}52}$-2000 mM ou le mélange de peptides (angiotensine I, angiotensine II, ATCH) dans une proportion (2/1/1) avant le dépôt sur la cible. L'énergie du LASER est réglée à 25% de la puissance et la gamme de masse de 150 à 3000 Da. L'acquisition de 6 spectres MALDI en mode positif et la mesure de masse précise génère une masse moyenne avec une précision estimée à 0,113 ppm. L'ionisation électrospray s'est faite par dilution au 1/100 de l'échantillon dans le méthanol 1% HCOOH.

V.F.3 ESI-Ion Trap

1µL d'échantillon est injecté sur une colonne 150 x 2,1 mm 2,7 µm Halo C18 après 22 min d'équilibration avec 60% de solvant A eau/ACN (60:40) 0,1% HCOOH et 40% de solvant B isopropanol/ACN (90:10) 0,1% HCOOH à un débit de 0,15 mL/min et 50°C. Le gradient atteint 100% de solvant B à 49 min avec un plateau de 8 min. L'acquisition du courant ionique total s'est faite en mode full-scan à 52000 u.m.a./sec et fragmentation MS^2 des 3 ions les plus intenses tout au long du chromatogramme. Une seconde acquisition réservée à la fragmentation de l'ion *m/z* 1750 a été utilisée pour comparer relativement son intensité dans les extraits polaires totaux.

V.G Résonance Magnétique Nucléaire

Les échantillons ont été repris puis séchés trois fois par un mélange $CDCl_3$ pur (dans le cas des MAMES) $CDCl_3/CD_3OD$ (2:1 ; dans le cas des glycolipides) $CDCl_3/CD_3OH$ (2:1) dans le cas du GPL4690) pour atteindre un volume de 270 µL. L'échantillon a été introduit dans un tube en verre (SHIGEMI, Allison Park, PA, Etats-unis) de diamètre 3 mm. Après introduction dans une sonde TBI permettant d'observer les noyaux d'hydrogène et de carbone et de phosphore, l'acquisition s'est faite à 293K sur une console AVANCE Neo (BRUKER BIOSPIN GmbH, Rheinstetten, Allemagne) et pour lequel la

fréquence de résonance des noyaux d'hydrogène est de 400MHz. Les séquences impulsionnelles utilisées pour chaque expérience homonucléaire ou hétéronucléaire étaient les séquences standards fournies par le constructeur. Après acquisition, les spectres ^{1}H ont subi une correction de phase ainsi qu'une calibration sur le signal du méthanol ($\delta\ ^{1}H/\delta\ ^{13}C$ = 3,31/49,29 ppm) ou sur celui du chloroforme ($\delta\ ^{1}H$ = 7,26 ppm).

La liste des impulsions constructeurs utilisées est la suivante : zg (^{1}H) ; selcssfdizs.2 (^{1}H) ; cosygpqf (COSY) ; dipsi2etgpsi (TOCSY) ; noesygpph (NOESY) ; hsqcedetgpsisp2.2 (HSQC) ; hsqcdietgpsisp2 (HSQC/TOCSY) ; hmbcetgpl3nd (HMBC).

V.H Couplage à des billes fluorescentes

Le dosage des déoxyhexoses contenus dans les GPLs purifiés a été réalisé par la méthode des itols acétates au moyen de standard externe de rhamnose et de 6 déoxy-talose, et d'une correction par l'étalon interne *myo*-inositol. Ensuite, l'équivalent de 40 µg de 3,4 di-O-méthyl rhamnose est aliquoté dans un tube stérile et séché sous une hotte ventilée à 50°C puis solubilisé dans 50µL de butanol. Le protocole de couplage aux billes a été adapté de celui de la littérature[408]. Brièvement, 15 µL de suspension contenant $5{,}4.10^{8}$ billes fluorescentes de 1µm (Em / Ex ; 505 nm / 515 nm ; THERMOFISHER SCIENTIFIC, San Jose, USA) ont été lavées 2 fois avec du carbonate/bicarbonate 0,05M pH 9,6 et remises en suspension dans 1 mL de PBS. 950µL de ce mélange a été ajouté au 50µL de butanol contenant les GPLs et le couplage a été réalisé 2 h à 37° C sous agitation douce. Les billes ont été lavées deux fois avec 1 mL de PBS, puis bloquées pendant 1 h avec 1 mL de BSA à 5 % dans du PBS à 37 °C sous agitation douce. Finalement, les billes ont été lavées deux fois avec 0,5% de BSA dans du PBS puis transférées pour stockage à -20°C.

VI BIBLIOGRAPHIE

1. Global Tuberculosis Report 2022. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022

2. Ahmed K. Global tuberculosis cases increase for the first time in 20 years. *The Guardian*. https://www.theguardian.com/global-development/2022/oct/28/global-tuberculosis-tb-cases-increase. Published October 28, 2022. Accessed December 22, 2022.

3. Floto RA, Olivier KN, Saiman L, et al. US Cystic Fibrosis Foundation and European Cystic Fibrosis Society consensus recommendations for the management of non-tuberculous mycobacteria in individuals with cystic fibrosis. *Thorax*. 2016;71 Suppl 1(Suppl 1):i1-22. doi:10.1136/thoraxjnl-2015-207360

4. Tran T, Bonham AJ, Chan ED, Honda JR. A paucity of knowledge regarding nontuberculous mycobacterial lipids compared to the tubercle bacillus. *Tuberculosis*. 2019;115:96-107. doi:10.1016/j.tube.2019.02.008

5. Trevisan. I Generi e le Specie delle Batteriacee.1889.

6. Lehmann. Atlas und Grundriss der Bakteriologie und Lehrbuch der speziellen bakteriologischen Diagnostik. *Lehmann's Medizin, Handatlanter X*. 1899.

7. Chester FD. A Manual of Determinative Bacteriology. *The MacMillan Co., New York,*. 1901.

8. Marchoux E, Sorel F. Recherches sur la lèpre. *Annales de l'Institut Pasteur, 26 : 675-700*. 1912.

9. Bergey HD. Bergey's Manual of Determinative Bacteriology, 1st ed.1923.

10. Aronson JD. Spontaneous Tuberculosis in Salt Water Fish. *J Infect Dis*. 1926;39(4):315-320. doi:10.1093/infdis/39.4.315

11. Fenner F. The Significance of the Incubation Period in Infectious Diseases. *Med J Aust*. 1950;2(23):813-818. doi:10.5694/j.1326-5377.1950.tb106945.x

12. Moore M, Frerichs JB. An Unusual Acid-Fast Infection of the Knee with Subcutaneous, Abscess-Like Lesions of the Gluteal Region: Report of a Case with a Study of the Organism, Mycobacterium abscessus, n. sp.1. *J Invest Dermatol*. 1953;20(2):133-169. doi:10.1038/jid.1953.18

13. Runyon EH. Anonymous Mycobacteria in Pulmonary Disease. *Med Clin North Am*. 1959;43(1):273-290. doi:10.1016/S0025-7125(16)34193-1

14. Runyon EH. Typical Mycobacteria: Their classification. *Am Rev Respir Dis*. 1965;91(2):288-289. doi:10.1164/arrd.1965.91.2.288

15. Bojalil LF, Cerbòn J, Trujillo A 1962. Adansonian classification of mycobacteria. *Microbiology*. 28(2):333-346. doi:10.1099/00221287-28-2-333

16. Tsukamura M. Identification of mycobacteria. *Tubercle*. 1967;48(4):311-338. doi:10.1016/S0041-3879(67)80040-0

17. Turenne CY. Nontuberculous mycobacteria: Insights on taxonomy and evolution. *Infect Genet Evol*. 2019;72:159-168. doi:10.1016/j.meegid.2019.01.017

18. Wayne LG, Gross WM. Base Composition of Deoxyribonucleic Acid Isolated from Mycobacteria. *J Bacteriol*. 1968;96(6):1915-1919.

19. Lechevalier MP, Horan AC, Lechevalier H. Lipid Composition in the Classification of Nocardiae and Mycobacteria. *J Bacteriol*. 1971;105(1):313-318.

20. Minnikin DE, Minnikin SM, Parlett JH, Goodfellow M, Magnusson M. Mycolic acid patterns of some species of Mycobacterium. *Arch Microbiol*. 1984;139(2):225-231. doi:10.1007/BF00402005

21. Levy-Frebault V, Portaels F. Proposed Minimal Standards for the Genus Mycobacterium and for Description of New Slowly Growing Mycobacterium Species. INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, Apr. 1992, p. 315-323, Vol. 42, n°2

22. Butler WR, Guthertz LS. Mycolic Acid Analysis by High-Performance Liquid Chromatography for Identification of Mycobacterium Species. *Clin Microbiol Rev*. 2001;14(4):704-726. doi:10.1128/CMR.14.4.704-726.2001

23. Hall RM, Ratledge C 1984. Mycobactins as Chemotaxonomic Characters for Some Rapidly Growing Mycobacteria. *Microbiology*. 130(8):1883-1892. doi:10.1099/00221287-130-8-1883

24. Wayne LG, Brenner DJ, Colwell RR, et al. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Int J Syst Evol Microbiol*. 37(4):463-464. doi:10.1099/00207713-37-4-463

25. Baess I. Deoxyribonucleic Acid Relatedness Among Species of Slowly-Growing Mycobacteria. *Acta Pathol Microbiol Scand [B]*. 1979;87B(1-6):221-226. doi:10.1111/j.1699-0463.1979.tb02430.x

26. Imaeda T 1985. Deoxyribonucleic Acid Relatedness Among Selected Strains of Mycobacterium tuberculosis, Mycobacterium bovis, Mycobacterium bovis BCG, Mycobacterium microti, and Mycobacterium africanum. *Int J Syst Evol Microbiol*. 35(2):147-150. doi:10.1099/00207713-35-2-147

27. Pitulle C, Dorsch M, Kazda J, Wolters J, Stackebrandt E. Phylogeny of Rapidly Growing Members of the Genus Mycobacterium. *Int J Syst Bacteriol*. 1992;42:337-343. doi:10.1099/00207713-42-3-337

28. McNabb A, Eisler D, Adie K, et al. Assessment of Partial Sequencing of the 65-Kilodalton Heat Shock Protein Gene (hsp65) for Routine Identification of Mycobacterium Species Isolated from Clinical Sources. *J Clin Microbiol*. 2004;42(7):3000-3011. doi:10.1128/JCM.42.7.3000-3011.2004

29. rpoB-Based Identification of Nonpigmented and Late-Pigmenting Rapidly Growing Mycobacteria | Journal of Clinical Microbiology. Accessed September 5, 2022. https://journals.asm.org/doi/10.1128/JCM.41.12.5699-5708.2003

30. Tortoli E. Microbiological Features and Clinical Relevance of New Species of the Genus Mycobacterium. *Clin Microbiol Rev*. 2014;27(4):727-752. doi:10.1128/CMR.00035-14

31. Fraser C, Fleischmann R. Strategies for whole microbial genome sequencing and analysis. Electrophoresis 1997, IS, 1207-1216.

32. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JMY 2007. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*. 57(1):81-91. doi:10.1099/ijs.0.64483-0

33. Kim M, Oh HS, Park SC, Chun J 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol*. 64(Pt_2):346-351. doi:10.1099/ijs.0.059774-0

34. Tortoli E, Fedrizzi T, Meehan CJ, et al. The new phylogeny of the genus Mycobacterium: The old and the news. *Infect Genet Evol*. 2017;56:19-25. doi:10.1016/j.meegid.2017.10.013

35. Gupta RS, Lo B, Son J. Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium into an Emended Genus Mycobacterium and Four Novel Genera. *Front Microbiol*. 2018;9. doi:10.3389/fmicb.2018.00067

36. Meehan CJ, Barco RA, Loh YHE, Cogneau S, Rigouts L. Reconstituting the genus Mycobacterium. *Int J Syst Evol Microbiol*. 2021;71(9):004922. doi:10.1099/ijsem.0.004922

37. Nouioui I, Carro L, García-López M, et al. Genome-Based Taxonomic Classification of the Phylum Actinobacteria. *Front Microbiol*. 2018;9. Accessed September 1, 2022. https://www.frontiersin.org/articles/10.3389/fmicb.2018.02007

38. Fedrizzi T, Meehan CJ, Grottola A, et al. Genomic characterization of Nontuberculous Mycobacteria. *Sci Rep*. 2017;7(1):45258. doi:10.1038/srep45258

39. Stinear TP, Mve-Obiang A, Small PLC, et al. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of Mycobacterium ulcerans. *Proc Natl Acad Sci U S A*. 2004;101(5):1345-1349. doi:10.1073/pnas.0305877101

40. Crawford JT, Cave MD, Bates JH. Characterization of plasmids from strains of Mycobacterium avium-intracellulare. *Rev Infect Dis*. 1981;3(5):949-952. doi:10.1093/clinids/3.5.949

41. Zainuddin ZF, Dale JW. Does Mycobacterium tuberculosis have plasmids? *Tubercle*. 1990;71(1):43-49. doi:10.1016/0041-3879(90)90060-L

42. Mohamad S, Ismail NN, Parumasivam T, Ibrahim P, Osman H, A. Wahab H. Antituberculosis activity, phytochemical identification of Costus speciosus (J. Koenig) Sm., Cymbopogon citratus (DC. Ex Nees) Stapf., and Tabernaemontana coronaria (L.) Willd. and their effects on the growth kinetics and cellular integrity of Mycobacterium tuberculosis H37Rv. *BMC Complement Altern Med*. 2018;18(1):5. doi:10.1186/s12906-017-2077-5

43. Goren MB, Brennan PJ. Mycobacterial lipids: chemistry and biologic activities. In Tuberculosis (ed. Youmans GP), pp. 63–193. W.B. Saunders, Philadelphia (1979).

44. Anand P, Akhter Y. A review on enzyme complexes of electron transport chain from Mycobacterium tuberculosis as promising drug targets. *Int J Biol Macromol*. 2022;212:474-494. doi:10.1016/j.ijbiomac.2022.05.124

45. Saito H. [Laboratory media for the cultivation of tubercle bacillus]. *Kekkaku*. 1998;73(5):329-337.

46. Agapova A, Serafini A, Petridis M, et al. Flexible nitrogen utilisation by the metabolic generalist pathogen Mycobacterium tuberculosis. *eLife*. 8:e41129. doi:10.7554/eLife.41129

47. Tu HZ, Chang SH, Huaug TS, Huaug WK, Liu YC, Lee SSJ. Microscopic Morphology in Smears Prepared from MGIT Broth Medium for Rapid Presumptive Identification of Mycobacterium tuberculosis complex, Mycobacterium avium complex and Mycobacterium kansasii. *Ann Clin Lab Sci*. 2003;33(2):179-183.

48. Julián E, Roldán M, Sánchez-Chardi A, Astola O, Agustí G, Luquin M. Microscopic Cords, a Virulence-Related Characteristic of Mycobacterium tuberculosis, Are Also Present in Nonpathogenic Mycobacteria. *J Bacteriol*. 2010;192(7):1751-1760. doi:10.1128/JB.01485-09

49. Palomino JC, Martin A, Von Groll A, Portaels F. Rapid culture-based methods for drug-resistance detection in Mycobacterium tuberculosis. *J Microbiol Methods*. 2008;75(2):161-166. doi:10.1016/j.mimet.2008.06.015

50. Yegian D, Baisden L. Factors Affecting the Beading of the Tubercle Bacillus Stained by the Ziehl-Neelsen Technique. *J Bacteriol*. 1942;44(6):667-672.

51. Kinyoun JJ. A MODIFICATION OF PONDER'S STAIN FOR DIPHTHERIA. *Am J Public Health N Y N 1912*. 1915;5(3):246-247. doi:10.2105/ajph.5.3.246

52. Steingart KR, Henry M, Ng V, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis*. 2006;6(9):570-581. doi:10.1016/S1473-3099(06)70578-3

53. Van Deun A, Hossain MA, Gumusboga M, Rieder HL. Ziehl-Neelsen staining: theory and practice [Technical Note]. *Int J Tuberc Lung Dis*. 2008;12(1):108-110.

54. Crothers JW, Laga AC, Solomon IH. Clinical Performance of Mycobacterial Immunohistochemistry in Anatomic Pathology Specimens: The Beginning of the End for Ziehl-Neelsen? *Am J Clin Pathol*. 2021;155(1):97-105. doi:10.1093/ajcp/aqaa119

55. Kent PT, Kubica GP. *Public Health Mycobacteriology: A Guide for the Level III Laboratory*. Centers for Disease Control, Atlanta, GA.; 1985. Accessed September 7, 2022. https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB86216546.xhtml

56. Springer B, Stockman L, Teschner K, Roberts GD, Böttger EC. Two-laboratory collaborative study on identification of mycobacteria: molecular versus phenotypic methods. *J Clin Microbiol*. 1996;34(2):296-303. doi:10.1128/jcm.34.2.296-303.1996

57. Smid I, Salfinger M. Mycobacterial identification by computer-aided gas-liquid chromatography. *Diagn Microbiol Infect Dis*. 1994;19(2):81-88. doi:10.1016/0732-8893(94)90117-1

58. Kellogg JA, Bankert DA, Withers GS, Sweimler W, Kiehn TE, Pfyffer GE. Application of the Sherlock Mycobacteria Identification System Using High-Performance Liquid Chromatography in a Clinical Laboratory. *J Clin Microbiol*. 2001;39(3):964-970. doi:10.1128/JCM.39.3.964-970.2001

59. Jia Khor M, Broda A, Kostrzewa M, Drobniewski F, Larrouy-Maumus G. An Improved Method for Rapid Detection of Mycobacterium abscessus Complex Based on Species-Specific Lipid Fingerprint by Routine MALDI-TOF. *Front Chem*. 2021;9. Accessed May 1, 2022. https://www.frontiersin.org/article/10.3389/fchem.2021.715890

60. Khéchine AE, Couderc C, Flaudrops C, Raoult D, Drancourt M. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Identification of Mycobacteria in Routine Clinical Practice. *PLOS ONE*. 2011;6(9):e24720. doi:10.1371/journal.pone.0024720

61. Alcaide F, Amlerová J, Bou G, et al. How to: identify non-tuberculous Mycobacterium species using MALDI-TOF mass spectrometry. *Clin Microbiol Infect*. 2018;24(6):599-603. doi:10.1016/j.cmi.2017.11.012

62. Cao Y, Wang L, Ma P, Fan W, Gu B, Ju S. Accuracy of Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Identification of Mycobacteria: a systematic review and meta-analysis. *Sci Rep*. 2018;8(1):4131. doi:10.1038/s41598-018-22642-w

63. Rodriguez-Temporal D, Alcaide F, Mareković I, et al. Multicentre study on the reproducibility of MALDI-TOF MS for nontuberculous mycobacteria identification. *Sci Rep*. 2022;12(1):1237. doi:10.1038/s41598-022-05315-7

64. Identification des mycobactéries tuberculeuses. Accessed September 7, 2022. https://microbiologiemedicale.fr/identification-mycobacteries-tuberculeuses/

65. Han XY, Pham AS, Tarrand JJ, Sood PK, Luthra R. Rapid and Accurate Identification of Mycobacteria by Sequencing Hypervariable Regions of the 16S Ribosomal RNA Gene. *Am J Clin Pathol*. 2002;118(5):796-801. doi:10.1309/HN44-XQYM-JMAQ-2EDL

66. Ei PW, Aung WW, Lee JS, Choi GE, Chang CL. Molecular Strain Typing of Mycobacterium tuberculosis: a Review of Frequently Used Methods. *J Korean Med Sci*. 2016;31(11):1673-1683. doi:10.3346/jkms.2016.31.11.1673

67. Larsen MH, Lacourciere K, Parker TM, et al. The Many Hosts of Mycobacteria 8 (MHM8): A conference report. *Tuberculosis*. 2020;121:101914. doi:10.1016/j.tube.2020.101914

68. Thorel MF, Falkinham JO III, Moreau R g. Environmental mycobacteria from alpine and subalpine habitats. *FEMS Microbiol Ecol*. 2004;49(3):343-347. doi:10.1016/j.femsec.2004.04.016

69. Schröder KH, Naumann L, Kroppenstedt RM, Reischl UY 1997. Mycobacterium hassiacum sp. nov., a New Rapidly Growing Thermophilic Mycobacterium. *Int J Syst Evol Microbiol*. 47(1):86-91. doi:10.1099/00207713-47-1-86

70. Norby B, Fosgate GT, Manning EJB, Collins MT, Roussel AJ. Environmental mycobacteria in soil and water on beef ranches: association between presence of cultivable mycobacteria and soil and water physicochemical characteristics. *Vet Microbiol*. 2007;124(1-2):153-159. doi:10.1016/j.vetmic.2007.04.015

71. Drancourt M. Looking in amoebae as a source of mycobacteria. *Microb Pathog*. 2014;77:119-124. doi:10.1016/j.micpath.2014.07.001

72. Whipps CM, Lieggi C, Wagner R. Mycobacteriosis in Zebrafish Colonies. *ILAR J Natl Res Counc Inst Lab Anim Resour*. 2012;53(2):95-105. doi:10.1093/ilar.53.2.95

73. Hilborn ED, Covert TC, Yakrus MA, et al. Persistence of Nontuberculous Mycobacteria in a Drinking Water System after Addition of Filtration Treatment. *Appl Environ Microbiol*. 2006;72(9):5864-5869. doi:10.1128/AEM.00759-06

74. September SM, Brözel VS, Venter SN. Diversity of nontuberculoid Mycobacterium species in biofilms of urban and semiurban drinking water distribution systems. *Appl Environ Microbiol*. 2004;70(12):7571-7573. doi:10.1128/AEM.70.12.7571-7573.2004

75. Tsao HF, Scheikl U, Herbold C, Indra A, Walochnik J, Horn M. The cooling tower water microbiota: Seasonal dynamics and co-occurrence of bacterial and protist phylotypes. *Water Res*. 2019;159:464-479. doi:10.1016/j.watres.2019.04.028

76. Mtetwa HN, Amoah ID, Kumari S, Bux F, Reddy P. The source and fate of Mycobacterium tuberculosis complex in wastewater and possible routes of transmission. *BMC Public Health*. 2022;22(1):145. doi:10.1186/s12889-022-12527-z

77. Le Dantec C, Duguet JP, Montiel A, Dumoutier N, Dubrou S, Vincent V. Chlorine Disinfection of Atypical Mycobacteria Isolated from a Water Distribution System. *Appl Environ Microbiol*. 2002;68(3):1025-1032. doi:10.1128/AEM.68.3.1025-1032.2002

78. Jankute M, Nataraj V, Lee OYC, et al. The role of hydrophobicity in tuberculosis evolution and pathogenicity. *Sci Rep*. 2017;7(1):1315. doi:10.1038/s41598-017-01501-0

79. Harland CW, Rabuka D, Bertozzi CR, Parthasarathy R. The Mycobacterium tuberculosis Virulence Factor Trehalose Dimycolate Imparts Desiccation Resistance to Model Mycobacterial Membranes. *Biophys J*. 2008;94(12):4718-4724. doi:10.1529/biophysj.107.125542

80. Wendt SL, George KL, Parker BC, Gruft H, Falkinham JO. Epidemiology of Infection by Nontuberculous Mycobacteria. *Am Rev Respir Dis*. 1980;122(2):259-263. doi:10.1164/arrd.1980.122.2.259

81. Parker BC, Ford MA, Gruft H, Falkinham JO. Epidemiology of infection by nontuberculous mycobacteria. IV. Preferential aerosolization of Mycobacterium intracellulare from natural waters. *Am Rev Respir Dis*. 1983;128(4):652-656. doi:10.1164/arrd.1983.128.4.652

82. Balamayooran G, Pena M, Sharma R, Truman RW. The armadillo as an animal model and reservoir host for Mycobacterium leprae. *Clin Dermatol*. 2015;33(1):108-115. doi:10.1016/j.clindermatol.2014.07.001

83. Bernut A, Herrmann JL, Ordway D, Kremer L. The Diverse Cellular and Animal Models to Decipher the Physiopathological Traits of Mycobacterium abscessus Infection. *Front Cell Infect Microbiol*. 2017;7. Accessed September 6, 2022. https://www.frontiersin.org/articles/10.3389/fcimb.2017.00100

84. Mudedla S, Avendano EE, Raman G. Non-tuberculous mycobacterium skin infections after tattooing in healthy individuals: A systematic review of case reports. *Dermatol Online J*. 2015;21(6):13030/qt8mr3r4f0.

85. Zimmermann P, Tebruegge M, Curtis N, Ritz N. The management of non-tuberculous cervicofacial lymphadenitis in children: A systematic review and meta-analysis. *J Infect*. 2015;71(1):9-18. doi:10.1016/j.jinf.2015.02.010

86. Wu EL, Al-Heeti O, Hoff BM, et al. Role of Therapeutic Drug Monitoring in the Treatment of Persistent Mycobacterium abscessus Central Nervous System Infection: A Case Report and Review of the Literature. *Open Forum Infect Dis*. 2022;9(8):ofac392. doi:10.1093/ofid/ofac392

87. Shu CC, Wu MF, Pan SW, Wu TS, Lai HC, Lin MC. Host immune response against environmental nontuberculous mycobacteria and the risk populations of nontuberculous mycobacterial lung disease. *J Formos Med Assoc*. 2020;119:S13-S22. doi:10.1016/j.jfma.2020.05.001

88. Marras TK, Mendelson D, Marchand-Austin A, May K, Jamieson FB. Pulmonary Nontuberculous Mycobacterial Disease, Ontario, Canada, 1998–2010 - Volume 19, Number 11—November 2013 - Emerging Infectious Diseases journal - CDC. doi:10.3201/eid1911.130737

89. Lai CC, Tan CK, Chou CH, et al. Increasing Incidence of Nontuberculous Mycobacteria, Taiwan, 2000–2008 - Volume 16, Number 2—February 2010 - Emerging Infectious Diseases journal - CDC. doi:10.3201/eid1602.090675

90. Hoefsloot W, van Ingen J, Andrejak C, et al. The geographic diversity of nontuberculous mycobacteria isolated from pulmonary samples: an NTM-NET collaborative study. *Eur Respir J*. 2013;42(6):1604-1613. doi:10.1183/09031936.00149212

91. van der Werf MJ, Ködmön C, Katalinić-Janković V, et al. Inventory study of non-tuberculous mycobacteria in the European Union. *BMC Infect Dis*. 2014;14:62. doi:10.1186/1471-2334-14-62

92. Vinnard C, Longworth S, Mezochow A, Patrawalla A, Kreiswirth BN, Hamilton K. Deaths Related to Nontuberculous Mycobacterial Infections in the United States, 1999–2014. *Ann Am Thorac Soc*. 2016;13(11):1951-1955. doi:10.1513/AnnalsATS.201606-474BC

93. López-Roa P, Aznar E, Cacho J, et al. Epidemiology of Non-Tuberculous Mycobacteria isolated from clinical specimens in Madrid, Spain, from 2013 to 2017. *Eur J Clin Microbiol Infect Dis*. 2020;39(6):1089-1094. doi:10.1007/s10096-020-03826-7

94. Brode SK, Daley CL, Marras TK. The epidemiologic relationship between tuberculosis and non-tuberculous mycobacterial disease: a systematic review. *Int J Tuberc Lung Dis*. 2014;18(11):1370-1377. doi:10.5588/ijtld.14.0120

95. Adjemian J, Daniel-Wayman S, Ricotta E, Prevots DR. Epidemiology of Nontuberculous Mycobacteriosis. *Semin Respir Crit Care Med*. 2018;39(3):325-335. doi:10.1055/s-0038-1651491

96. Strollo SE, Adjemian J, Adjemian MK, Prevots DR. The Burden of Pulmonary Nontuberculous Mycobacterial Disease in the United States. *Ann Am Thorac Soc*. 2015;12(10):1458-1464. doi:10.1513/AnnalsATS.201503-173OC

97. Diel R, Jacob J, Lampenius N, et al. Burden of non-tuberculous mycobacterial pulmonary disease in Germany. *Eur Respir J*. 2017;49(4):1602109. doi:10.1183/13993003.02109-2016

98. Bemer P, Peuchant O, Guet-Revillet H, et al. Management of patients with pulmonary mycobacteriosis in France: a multicenter retrospective cohort study. *BMC Pulm Med*. 2021;21(1):333. doi:10.1186/s12890-021-01701-5

99. T JAS, J R, Rajan A, Shankar V. Features of the biochemistry of Mycobacterium smegmatis, as a possible model for Mycobacterium tuberculosis. *J Infect Public Health*. 2020;13(9):1255-1264. doi:10.1016/j.jiph.2020.06.023

100. Snapper SB, Melton RE, Mustafa S, Kieser T, Jr WRJ. Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. *Mol Microbiol*. 1990;4(11):1911-1919. doi:10.1111/j.1365-2958.1990.tb02040.x

101. Norgard MV, Imaeda T. Physiological factors involved in the transformation of Mycobacterium smegmatis. *J Bacteriol*. 1978;133(3):1254-1262. doi:10.1128/jb.133.3.1254-1262.1978

102. Hermans J, de Bont JA. Techniques for genetic engineering in mycobacteria. Alternative host strains, DNA-transfer systems and vectors. *Antonie Van Leeuwenhoek*. 1996;69(3):243-256. doi:10.1007/BF00399613

103. Bashiri G, Baker EN. Production of recombinant proteins in Mycobacterium smegmatis for structural and functional studies. *Protein Sci*. 2015;24(1):1-10. doi:10.1002/pro.2584

104. Kubica GP, Baess I, Gordon RE, et al. A Co-operative Numerical Analysis of Rapidly Growing Mycobacteria. *Microbiology*. 73(1):55-70. doi:10.1099/00221287-73-1-55

105. Kusunoki S, Ezaki T 1992. Proposal of Mycobacterium peregrinum sp. nov., nom. rev., and Elevation of Mycobacterium chelonae subsp. abscessus (Kubica et al.) to Species Status: Mycobacterium abscessus comb. nov. *Int J Syst Evol Microbiol*. 42(2):240-245. doi:10.1099/00207713-42-2-240

106. Wilson RW, Steingrube VA, Böttger EC, et al. Mycobacterium immunogenum sp. nov., a novel species related to Mycobacterium abscessus and associated with clinical disease, pseudo-outbreaks and contaminated metalworking fluids: an international cooperative study on mycobacterial taxonomy. *Int J Syst Evol Microbiol*. 2001;51(5):1751-1764. doi:10.1099/00207713-51-5-1751

107. Nash KA, Brown-Elliott BA, Wallace RJ. A novel gene, erm(41), confers inducible macrolide resistance to clinical isolates of Mycobacterium abscessus but is absent from Mycobacterium chelonae. *Antimicrob Agents Chemother*. 2009;53(4):1367-1376. doi:10.1128/AAC.01275-08

108. Leao SC, Tortoli E, Viana-Niero C, et al. Characterization of mycobacteria from a major Brazilian outbreak suggests that revision of the taxonomic status of members of the Mycobacterium

chelonae-M. abscessus group is needed. *J Clin Microbiol*. 2009;47(9):2691-2698. doi:10.1128/JCM.00808-09

109. Adékambi T, Reynaud-Gaubert M, Greub G, et al. Amoebal Coculture of "Mycobacterium massiliense" sp. nov. from the Sputum of a Patient with Hemoptoic Pneumonia. *J Clin Microbiol*. 2004;42(12):5493-5501. doi:10.1128/JCM.42.12.5493-5501.2004

110. rpoB gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of Mycobacterium bolletii sp. nov., Mycobacterium phocaicum sp. nov. and Mycobacterium aubagnense sp. nov. | Microbiology Society. Accessed September 12, 2022. https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.63969-0

111. Choi GE, Cho YJ, Koh WJ, Chun J, Cho SN, Shin SJ. Draft Genome Sequence of Mycobacterium abscessus subsp. bolletii BDT. *J Bacteriol*. 2012;194(10):2756-2757. doi:10.1128/JB.00354-12

112. Leao SC, Tortoli E, Euzéby JP, Garcia MJ 2011. Proposal that Mycobacterium massiliense and Mycobacterium bolletii be united and reclassified as Mycobacterium abscessus subsp. bolletii comb. nov., designation of Mycobacterium abscessus subsp. abscessus subsp. nov. and emended description of Mycobacterium abscessus. *Int J Syst Evol Microbiol*. 61(9):2311-2313. doi:10.1099/ijs.0.023770-0

113. Cho YJ, Yi H, Chun J, et al. The Genome Sequence of 'Mycobacterium massiliense' Strain CIP 108297 Suggests the Independent Taxonomic Status of the Mycobacterium abscessus Complex at the Subspecies Level. *PLOS ONE*. 2013;8(11):e81560. doi:10.1371/journal.pone.0081560

114. Sassi M, Drancourt M. Genome analysis reveals three genomospecies in Mycobacterium abscessus. *BMC Genomics*. 2014;15(1):359. doi:10.1186/1471-2164-15-359

115. Tortoli E, Kohl TA, Brown-Elliott BA, et al. Emended description of Mycobacterium abscessus, Mycobacterium abscessus subsp. abscessus and Mycobacteriumabscessus subsp. bolletii and designation of Mycobacteriumabscessus subsp. massiliense comb. nov. *Int J Syst Evol Microbiol*. 66(11):4471-4479. doi:10.1099/ijsem.0.001376

116. Kim HY, Kim BJ, Kook Y, et al. Mycobacterium massiliense is differentiated from Mycobacterium abscessus and Mycobacterium bolletii by erythromycin ribosome methyltransferase gene (erm) and clarithromycin susceptibility patterns. *Microbiol Immunol*. 2010;54(6):347-353. doi:10.1111/j.1348-0421.2010.00221.x

117. Cardenas DD, Yasmin T, Ahmed S. A Rare Insidious Case of Skin and Soft Tissue Infection Due to Mycobacterium abscessus: A Case Report. *Cureus*. 14(6):e25725. doi:10.7759/cureus.25725

118. Deng L, Luo YZ, Liu F, Yu XH. Subcutaneous infection caused by Mycobacterium abscessus following cosmetic injections of botulinum toxin: A case report. *World J Clin Cases*. 2022;10(18):6141-6147. doi:10.12998/wjcc.v10.i18.6141

119. Clabots D, Serrat A. Mycobacterium abscessus peritonitis and ventriculitis associated with ventriculoperitoneal shunt. *IDCases*. 2022;27:e01445. doi:10.1016/j.idcr.2022.e01445

120. Sermet-Gaudelus I, Le Bourgeois M, Pierre-Audigier C, et al. Mycobacterium abscessus and Children with Cystic Fibrosis. *Emerg Infect Dis*. 2003;9(12):1587-1591. doi:10.3201/eid0912.020774

121. Roux AL, Catherinot E, Ripoll F, et al. Multicenter Study of Prevalence of Nontuberculous Mycobacteria in Patients with Cystic Fibrosis in France. *J Clin Microbiol*. 2009;47(12):4124-4128. doi:10.1128/JCM.01257-09

122. Jönsson BE, Gilljam M, Lindblad A, Ridell M, Wold AE, Welinder-Olsson C. Molecular epidemiology of Mycobacterium abscessus, with focus on cystic fibrosis. *J Clin Microbiol*. 2007;45(5):1497-1504. doi:10.1128/JCM.02592-06

123. Bryant JM, Grogono DM, Greaves D, et al. Whole-genome sequencing to identify transmission of Mycobacterium abscessus between patients with cystic fibrosis: a retrospective cohort study. *The Lancet*. 2013;381(9877):1551-1560. doi:10.1016/S0140-6736(13)60632-7

124. Bryant JM, Grogono DM, Rodriguez-Rincon D, et al. Population-level genomics identifies the emergence and global spread of a human transmissible multidrug-resistant nontuberculous mycobacterium. *Science*. 2016;354(6313):751-757. doi:10.1126/science.aaf8156

125. Tortoli E, Kohl TA, Trovato A, et al. Mycobacterium abscessus in patients with cystic fibrosis: low impact of inter-human transmission in Italy. *Eur Respir J*. 2017;50(1). doi:10.1183/13993003.02525-2016

126. Wetzstein N, Diricks M, Kohl TA, et al. Molecular Epidemiology of Mycobacterium abscessus Isolates Recovered from German Cystic Fibrosis Patients. *Microbiol Spectr*. 10(4):e01714-22. doi:10.1128/spectrum.01714-22

127. Yan J, Kevat A, Martinez E, et al. Investigating transmission of Mycobacterium abscessus amongst children in an Australian cystic fibrosis centre. *J Cyst Fibros*. 2020;19(2):219-224. doi:10.1016/j.jcf.2019.02.011

128. Lim AYH, Chotirmall SH, Fok ETK, et al. Profiling non-tuberculous mycobacteria in an Asian setting: characteristics and clinical outcomes of hospitalized patients in Singapore. *BMC Pulm Med*. 2018;18(1):85. doi:10.1186/s12890-018-0637-1

129. Nagano H, Kinjo T, Nei Y, Yamashiro S, Fujita J, Kishaba T. Causative species of nontuberculous mycobacterial lung disease and comparative investigation on clinical features of Mycobacterium abscessus complex disease: A retrospective analysis for two major hospitals in a subtropical region of Japan. *PLOS ONE*. 2017;12(10):e0186826. doi:10.1371/journal.pone.0186826

130. Huang HL, Cheng MH, Lu PL, et al. Epidemiology and Predictors of NTM Pulmonary Infection in Taiwan - a Retrospective, Five-Year Multicenter Study. *Sci Rep*. 2017;7(1):16300. doi:10.1038/s41598-017-16559-z

131. Yoshida M, Chien JY, Morimoto K, et al. Molecular Epidemiological Characteristics of Mycobacterium abscessus Complex Derived from Non-Cystic Fibrosis Patients in Japan and Taiwan. Doucet-Populaire FC, ed. *Microbiol Spectr*. 2022;10(3):e00571-22. doi:10.1128/spectrum.00571-22

132. Abdelaal HFM, Chan ED, Young L, Baldwin SL, Coler RN. Mycobacterium abscessus: It's Complex. *Microorganisms*. 2022;10(7):1454. doi:10.3390/microorganisms10071454

133. Chest X-Ray Finding of Pulmonary Tuberculosis and Nontuberculous Mycobacterial Lung Diseases in Patients with Acid-Fast Bacilli Smear-Positive Sputum. Accessed September 14, 2022. https://austinpublishinggroup.com/tuberculosis/fulltext/atrt-v2-id1008.php

134. Chin KL, Sarmiento ME, Alvarez-Cabrera N, Norazmi MN, Acosta A. Pulmonary non-tuberculous mycobacterial infections: current state and future management. *Eur J Clin Microbiol Infect Dis*. 2020;39(5):799-826. doi:10.1007/s10096-019-03771-0

135. Genc GE, Demir M, Yaman G, Kayar B, Koksal F, Satana D. Evaluation of MALDI-TOF MS for identification of nontuberculous mycobacteria isolated from clinical specimens in mycobacteria growth indicator tube medium. *New Microbiol*. 2018;41(3):214-219.

136. Le Moigne V, Roux AL, Mahoudo H, et al. IgA Serological Response for the Diagnosis of Mycobacterium abscessus Infections in Patients with Cystic Fibrosis. *Microbiol Spectr*. 10(3):e00192-22. doi:10.1128/spectrum.00192-22

137. Bordin A, Pandey S, Coulter C, et al. Rapid macrolide and amikacin resistance testing for Mycobacterium abscessus in people with cystic fibrosis. *J Med Microbiol*. 70(4):001349. doi:10.1099/jmm.0.001349

138. Harrison J, Weaver JA, Desai M, Cox JAG. In vitro efficacy of relebactam versus avibactam against Mycobacterium abscessus complex. *Cell Surf*. 2021;7:100064. doi:10.1016/j.tcsw.2021.100064

139. Bastian S, Veziris N, Roux AL, et al. Assessment of Clarithromycin Susceptibility in Strains Belonging to the Mycobacterium abscessus Group by erm(41) and rrl Sequencing. *Antimicrob Agents Chemother*. 2011;55(2):775-781. doi:10.1128/AAC.00861-10

140. Ryskova L, Bolehovska R, Kukla R, et al. Mycobacterioses Induced by Mycobacterium abscessus: Case Studies Indicating the Importance of Molecular Analysis for the Identification of Antibiotic Resistance. *Antibiotics*. 2022;11(7):873. doi:10.3390/antibiotics11070873

141. Rominski A, Selchow P, Becker K, Brülle JK, Dal Molin M, Sander P. Elucidation of Mycobacterium abscessus aminoglycoside and capreomycin resistance by targeted deletion of three putative resistance genes. *J Antimicrob Chemother*. 2017;72(8):2191-2200. doi:10.1093/jac/dkx125

142. Naor N, Zarbib E, Barkan D. Mycobacterium abscessus Tetracycline-Modifying Monooxygenase MAB_1496c Appears Not to Be Sufficient to Cause Resistance to Tetracycline When Expressed in Mycobacterium smegmatis. *Microbiol Spectr*. 10(4):e02346-22. doi:10.1128/spectrum.02346-22

143. Rominski A, Roditscheff A, Selchow P, Böttger EC, Sander P. Intrinsic rifamycin resistance of Mycobacterium abscessus is mediated by ADP-ribosyltransferase MAB_0591. *J Antimicrob Chemother*. 2017;72(2):376-384. doi:10.1093/jac/dkw466

144. Egorova A, Jackson M, Gavrilyuk V, Makarov V. Pipeline of anti-Mycobacterium abscessus small molecules: Repurposable drugs and promising novel chemical entities. *Med Res Rev*. 2021;41(4):2350-2387. doi:10.1002/med.21798

145. Cattoir V. Identification moléculaire des mycobactéries et détection de la résistance aux antibiotiques. *Ann Biol Clin (Paris)*. 2004;62(4):405-413.

146. Gutiérrez AV, Richard M, Roquet-Banères F, Viljoen A, Kremer L. The TetR Family Transcription Factor MAB_2299c Regulates the Expression of Two Distinct MmpS-MmpL Efflux Pumps Involved in Cross-Resistance to Clofazimine and Bedaquiline in Mycobacterium abscessus. *Antimicrob Agents Chemother*. 2019;63(10):e01000-19. doi:10.1128/AAC.01000-19

147. Halloum I, Viljoen A, Khanna V, et al. Resistance to Thiacetazone Derivatives Active against Mycobacterium abscessus Involves Mutations in the MmpL5 Transcriptional Repressor MAB_4384. *Antimicrob Agents Chemother*. 2017;61(4):e02509-16. doi:10.1128/AAC.02509-16

148. Bakala N'Goma JC, Le Moigne V, Soismier N, et al. Mycobacterium abscessus phospholipase C expression is induced during coculture within amoebae and enhances M. abscessus virulence in mice. *Infect Immun*. 2015;83(2):780-791. doi:10.1128/IAI.02032-14

149. Buchmeier N, Blanc-Potard A, Ehrt S, Piddington D, Riley L, Groisman EA. A parallel intraphagosomal survival strategy shared by Mycobacterium tuberculosis and Salmonella enterica. *Mol Microbiol*. 2000;35(6):1375-1382. doi:10.1046/j.1365-2958.2000.01797.x

150. St John G, Brot N, Ruan J, et al. Peptide methionine sulfoxide reductase from Escherichia coli and Mycobacterium tuberculosis protects bacteria against oxidative damage from reactive nitrogen intermediates. *Proc Natl Acad Sci U S A*. 2001;98(17):9901-9906. doi:10.1073/pnas.161295398

151. Siegrist MS, Unnikrishnan M, McConnell MJ, et al. Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proc Natl Acad Sci*. 2009;106(44):18792-18797. doi:10.1073/pnas.0900589106

152. Kim YS, Yang CS, Nguyen LT, et al. Mycobacterium abscessus ESX-3 plays an important role in host inflammatory and pathological responses during infection. *Microbes Infect*. 2017;19(1):5-17. doi:10.1016/j.micinf.2016.09.001

153. Lagune M, Le Moigne V, Johansen MD, et al. The ESX-4 substrates, EsxU and EsxT, modulate Mycobacterium abscessus fitness. *PLoS Pathog*. 2022;18(8):e1010771. doi:10.1371/journal.ppat.1010771

154. Ripoll F, Pasek S, Schenowitz C, et al. Non Mycobacterial Virulence Genes in the Genome of the Emerging Pathogen Mycobacterium abscessus. *PLOS ONE*. 2009;4(6):e5660. doi:10.1371/journal.pone.0005660

155. Arruda S, Bomfim G, Knights R, Huima-Byron T, Riley LW. Cloning of an M. tuberculosis DNA Fragment Associated with Entry and Survival Inside Cells. *Science*. 1993;261(5127):1454-1457.

156. Viljoen A, Dubois V, Girard-Misguich F, Blaise M, Herrmann JL, Kremer L. The diverse family of MmpL transporters in mycobacteria: from regulation to antimicrobial developments. *Mol Microbiol*. 2017;104(6):889-904. doi:10.1111/mmi.13675

157. Tahlan K, Wilson R, Kastrinsky DB, et al. SQ109 Targets MmpL3, a Membrane Transporter of Trehalose Monomycolate Involved in Mycolic Acid Donation to the Cell Wall Core of

Mycobacterium tuberculosis. *Antimicrob Agents Chemother*. 2012;56(4):1797-1809. doi:10.1128/AAC.05708-11

158. Dubois V, Viljoen A, Laencina L, et al. MmpL8MAB controls Mycobacterium abscessus virulence and production of a previously unknown glycolipid family. *Proc Natl Acad Sci U S A*. 2018;115(43):E10147-E10156. doi:10.1073/pnas.1812984115

159. Converse SE, Mougous JD, Leavell MD, Leary JA, Bertozzi CR, Cox JS. MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. *Proc Natl Acad Sci U S A*. 2003;100(10):6121-6126. doi:10.1073/pnas.1030024100

160. Gomez M, Doukhan L, Nair G, Smith I. sigA is an essential gene in Mycobacterium smegmatis. *Mol Microbiol*. 1998;29(2):617-628. doi:10.1046/j.1365-2958.1998.00960.x

161. Sun R, Converse PJ, Ko C, Tyagi S, Morrison NE, Bishai WR. Mycobacterium tuberculosis ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set. *Mol Microbiol*. 2004;52(1):25-38. doi:10.1111/j.1365-2958.2003.03958.x

162. Calamita H, Ko C, Tyagi S, Yoshimatsu T, Morrison NE, Bishai WR. The Mycobacterium tuberculosis SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence. *Cell Microbiol*. 2005;7(2):233-244. doi:10.1111/j.1462-5822.2004.00454.x

163. Jensen-Cain DM, Quinn FD. Differential expression of sigE by Mycobacterium tuberculosis during intracellular growth. *Microb Pathog*. 2001;30(5):271-278. doi:10.1006/mpat.2001.0431

164. Raman S, Song T, Puyang X, Bardarov S, Jacobs WR, Husson RN. The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in Mycobacterium tuberculosis. *J Bacteriol*. 2001;183(20):6119-6125. doi:10.1128/JB.183.20.6119-6125.2001

165. Song T, Dove SL, Lee KH, Husson RN. RshA, an anti-sigma factor that regulates the activity of the mycobacterial stress response sigma factor SigH. *Mol Microbiol*. 2003;50(3):949-959. doi:10.1046/j.1365-2958.2003.03739.x

166. Chen JM, German GJ, Alexander DC, Ren H, Tan T, Liu J. Roles of Lsr2 in colony morphology and biofilm formation of Mycobacterium smegmatis. *J Bacteriol*. 2006;188(2):633-641. doi:10.1128/JB.188.2.633-641.2006

167. Colangeli R, Haq A, Arcus VL, et al. The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc Natl Acad Sci U S A*. 2009;106(11):4414-4418. doi:10.1073/pnas.0810126106

168. Le Moigne V, Bernut A, Cortès M, et al. Lsr2 Is an Important Determinant of Intracellular Growth and Virulence in Mycobacterium abscessus. *Front Microbiol*. 2019;10. Accessed May 22, 2022. https://www.frontiersin.org/article/10.3389/fmicb.2019.00905

169. Byrd TF, Lyons CR. Preliminary characterization of a Mycobacterium abscessus mutant in human and murine models of infection. *Infect Immun*. 1999;67(9):4700-4707. doi:10.1128/IAI.67.9.4700-4707.1999

170. Howard ST, Rhoades E, Recht J, et al. Spontaneous reversion of Mycobacterium abscessus from a smooth to a rough morphotype is associated with reduced expression of glycopeptidolipid and reacquisition of an invasive phenotype. *Microbiology*. 2006;152(6):1581-1590. doi:10.1099/mic.0.28625-0

171. Catherinot E, Clarissou J, Etienne G, et al. Hypervirulence of a Rough Variant of the *Mycobacterium abscessus* Type Strain. *Infect Immun*. 2007;75(2):1055-1058. doi:10.1128/IAI.00835-06

172. Sánchez-Chardi A, Olivares F, Byrd TF, Julián E, Brambilla C, Luquin M. Demonstration of Cord Formation by Rough Mycobacterium abscessus Variants: Implications for the Clinical Microbiology Laboratory. *J Clin Microbiol*. 2011;49(6):2293-2295. doi:10.1128/JCM.02322-10

173. Kim BJ, Kim BR, Lee SY, Kook YH, Kim BJ. Rough colony morphology of Mycobacterium massiliense Type II genotype is due to the deletion of glycopeptidolipid locus within its genome. *BMC Genomics*. 2013;14(1):890. doi:10.1186/1471-2164-14-890

174. Pawlik A, Garnier G, Orgeur M, et al. Identification and characterization of the genetic changes responsible for the characteristic smooth-to-rough morphotype alterations of clinically persistent Mycobacterium abscessus. *Mol Microbiol*. 2013;90(3):612-629. doi:10.1111/mmi.12387

175. Park IK, Hsu AP, Tettelin H, et al. Clonal Diversification and Changes in Lipid Traits and Colony Morphology in Mycobacterium abscessus Clinical Isolates. *J Clin Microbiol*. 2015;53(11):3438-3447. doi:10.1128/JCM.02015-15

176. Llorens-Fons M, Pérez-Trujillo M, Julián E, et al. Trehalose Polyphleates, External Cell Wall Lipids in Mycobacterium abscessus, Are Associated with the Formation of Clumps with Cording Morphology, Which Have Been Associated with Virulence. *Front Microbiol*. 2017;8. Accessed May 21, 2022. https://www.frontiersin.org/article/10.3389/fmicb.2017.01402

177. Halloum I, Carrère-Kremer S, Blaise M, et al. Deletion of a dehydratase important for intracellular growth and cording renders rough Mycobacterium abscessus avirulent. *Proc Natl Acad Sci*. 2016;113(29):E4228-E4237. doi:10.1073/pnas.1605477113

178. Tsai SH, Shen GH, Lin CH, Liau JR, Lai HC, Hu ST. Mab_3168c, a Putative Acetyltransferase, Enhances Adherence, Intracellular Survival and Antimicrobial Resistance of Mycobacterium abscessus. *PLOS ONE*. 2013;8(6):e67563. doi:10.1371/journal.pone.0067563

179. Liu TY, Tsai SH, Chen JW, Wang YC, Hu ST, Chen YY. Mab_3083c Is a Homologue of RNase J and Plays a Role in Colony Morphotype, Aggregation, and Sliding Motility of Mycobacterium abscessus. *Microorganisms*. 2021;9(4):676. doi:10.3390/microorganisms9040676

180. Rhoades ER, Archambault AS, Greendyke R, Hsu FF, Streeter C, Byrd TF. Mycobacterium abscessus Glycopeptidolipids Mask Underlying Cell Wall Phosphatidyl-myo-Inositol Mannosides Blocking Induction of Human Macrophage TNF-α by Preventing Interaction with TLR2. *J Immunol*. 2009;183(3):1997-2007. doi:10.4049/jimmunol.0802181

181. Davidson LB, Nessar R, Kempaiah P, Perkins DJ, Byrd TF. Mycobacterium abscessus Glycopeptidolipid Prevents Respiratory Epithelial TLR2 Signaling as Measured by HβD2 Gene Expression and IL-8 Release. *PLOS ONE*. 2011;6(12):e29148. doi:10.1371/journal.pone.0029148

182. Nessar R, Reyrat JM, Davidson LB, Byrd TFY 2011. Deletion of the mmpL4b gene in the Mycobacterium abscessus glycopeptidolipid biosynthetic pathway results in loss of surface colonization capability, but enhanced ability to replicate in human macrophages and stimulate their innate immune response. *Microbiology*. 2011;157(4):1187-1195. doi:10.1099/mic.0.046557-0

183. Malcolm KC, Nichols EM, Caceres SM, et al. Mycobacterium abscessus Induces a Limited Pattern of Neutrophil Activation That Promotes Pathogen Survival. *PLOS ONE*. 2013;8(2):e57402. doi:10.1371/journal.pone.0057402

184. Qvist T, Eickhardt S, Kragh KN, et al. Chronic pulmonary disease with Mycobacterium abscessus complex is a biofilm infection. *Eur Respir J*. 2015;46(6):1823-1826. doi:10.1183/13993003.01102-2015

185. Bernut A, Herrmann JL, Kissa K, et al. Mycobacterium abscessus cording prevents phagocytosis and promotes abscess formation. *Proc Natl Acad Sci*. 2014;111(10):E943-E952. doi:10.1073/pnas.1321390111

186. Bernut A, Nguyen-Chi M, Halloum I, Herrmann JL, Lutfalla G, Kremer L. Mycobacterium abscessus-Induced Granuloma Formation Is Strictly Dependent on TNF Signaling and Neutrophil Trafficking. *PLOS Pathog*. 2016;12(11):e1005986. doi:10.1371/journal.ppat.1005986

187. Roux AL, Viljoen A, Bah A, et al. The distinct fate of smooth and rough Mycobacterium abscessus variants inside macrophages. *Open Biol*. 2016;6(11):160185. doi:10.1098/rsob.160185

188. Kam JY, Hortle E, Krogman E, et al. Rough and smooth variants of Mycobacterium abscessus are differentially controlled by host immunity during chronic infection of adult zebrafish. *Nat Commun*. 2022;13:952. doi:10.1038/s41467-022-28638-5

189. Daley CL, Iaccarino JM, Lange C, et al. Treatment of Nontuberculous Mycobacterial Pulmonary Disease: An Official ATS/ERS/ESCMID/IDSA Clinical Practice Guideline. *Clin Infect Dis*. 2020;71(4):e1-e36. doi:10.1093/cid/ciaa241

190. Brown-Elliott BA, Nash KA, Wallace RJ. Antimicrobial Susceptibility Testing, Drug Resistance Mechanisms, and Therapy of Infections with Nontuberculous Mycobacteria. *Clin Microbiol Rev*. 2012;25(3):545-582. doi:10.1128/CMR.05030-11

191. Bich Hanh BT, Quang NT, Park Y, et al. Omadacycline Potentiates Clarithromycin Activity Against Mycobacterium abscessus. *Front Pharmacol*. 2021;12. https://www.frontiersin.org/article/10.3389/fphar.2021.790767

192. Wang G, Tang J, Feng J, et al. Activity of Oritavancin and Its Synergy with Other Antibiotics against Mycobacterium abscessus Infection In Vitro and In Vivo. *Int J Mol Sci*. 2021;22(12):6346. doi:10.3390/ijms22126346

193. Dedrick RM, Guerrero-Bustamante CA, Garlena RA, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. *Nat Med*. 2019;25(5):730-733. doi:10.1038/s41591-019-0437-z

194. Dedrick RM, Freeman KG, Nguyen JA, et al. Nebulized Bacteriophage in a Patient With Refractory Mycobacterium abscessus Lung Disease. *Open Forum Infect Dis*. 2022;9(7):ofac194. doi:10.1093/ofid/ofac194

195. Daffé M, Crick DC, Jackson M. Genetics of Capsular Polysaccharides and Cell Envelope (Glyco)lipids. *Microbiol Spectr*. 2014;2(4):MGM2-0021-2013. doi:10.1128/microbiolspec.MGM2-0021-2013

196. Jankute M, Cox JAG, Harrison J, Besra GS. Assembly of the Mycobacterial Cell Wall. *Annu Rev Microbiol*. 2015;69(1):405-423. doi:10.1146/annurev-micro-091014-104121

197. Kaur D, Guerin ME, Skovierová H, Brennan PJ, Jackson M. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of Mycobacterium tuberculosis. *Adv Appl Microbiol*. 2009;69:23-78. doi:10.1016/S0065-2164(09)69002-X

198. Brennan PJ, Nikaido H. THE ENVELOPE OF MYCOBACTERIA. Annu. Rev. Biochem. 1995. 64:29-6.1.

199. Daffé M. The cell envelope of tubercle bacilli. *Tuberculosis*. 2015;95:S155-S158. doi:10.1016/j.tube.2015.02.024

200. Seeliger J, Moody DB. Monstrous Mycobacterial Lipids. *Cell Chem Biol*. 2016;23(2):207-209. doi:10.1016/j.chembiol.2016.02.004

201. Chatterjee D, Khoo KH. The surface glycopeptidolipids of mycobacteria: structures and biological properties: *Cell Mol Life Sci*. 2001;58(14):2018-2042. doi:10.1007/PL00000834

202. Hessel M. van Dijk J, van der Marel GA, Codée JDC. Developments in the Synthesis of Mycobacterial Phenolic Glycolipids. *Chem Rec*. 2021;21(11):3295-3312. doi:10.1002/tcr.202100200

203. Ortalo-Magné A, Dupont MA, Lemassu A, Andersen AB, Gounon P, Daffé M. Molecular composition of the outermost capsular material of the tubercle bacillus. *Microbiol Read Engl*. 1995;141 (Pt 7):1609-1620. doi:10.1099/13500872-141-7-1609

204. Lemassu A, Ortalo-Magné A, Bardou F, Silve G, Lanéelle MA, Daffé M. Extracellular and surface-exposed polysaccharides of non-tuberculous mycobacteria. *Microbiol Read Engl*. 1996;142 (Pt 6):1513-1520. doi:10.1099/13500872-142-6-1513

205. Patterson JH, Waller RF, Jeevarajah D, Billman-Jacobe H, McCONVILLE MJ. Mannose metabolism is required for mycobacterial growth. *Biochem J*. 2003;372(1):77-86. doi:10.1042/bj20021700

206. Haites RE, Morita YS, McConville MJ, Billman-Jacobe H. Function of Phosphatidylinositol in Mycobacteria*. *J Biol Chem*. 2005;280(12):10981-10987. doi:10.1074/jbc.M413443200

207. Ballou CE, Vilkas E, Lederer E. Structural Studies on the Myo-inositol Phospholipids of Mycobacterium tuberculosis (var. bovis, strain BCG). *J Biol Chem*. 1963;238(1):69-76. doi:10.1016/S0021-9258(19)83963-7

208. Lee YC, Ballou CE. Complete Structures of the Glycophospholipids of Mycobacteria*. *Biochemistry*. 1965;4(7):1395-1404. doi:10.1021/bi00883a026

209. Korduláková J, Gilleron M, Mikušová K, et al. Definition of the First Mannosylation Step in Phosphatidylinositol Mannoside Synthesis: PimA IS ESSENTIAL FOR GROWTH OF MYCOBACTERIA*. *J Biol Chem*. 2002;277(35):31335-31344. doi:10.1074/jbc.M204060200

210. Boldrin F, Ventura M, Degiacomi G, et al. The Phosphatidyl-myo-Inositol Mannosyltransferase PimA Is Essential for Mycobacterium tuberculosis Growth In Vitro and In Vivo. *J Bacteriol*. 2014;196(19):3441-3451. doi:10.1128/JB.01346-13

211. Guerin ME, Kaur D, Somashekar BS, et al. New Insights into the Early Steps of Phosphatidylinositol Mannoside Biosynthesis in Mycobacteria. *J Biol Chem*. 2009;284(38):25687-25696. doi:10.1074/jbc.M109.030593

212. Palčeková Z, Gilleron M, Angala SK, et al. Polysaccharide Succinylation Enhances the Intracellular Survival of Mycobacterium abscessus. *ACS Infect Dis*. 2020;6(8):2235-2248. doi:10.1021/acsinfecdis.0c00361

213. KREMER L, GURCHA SS, BIFANI P, et al. Characterization of a putative α-mannosyltransferase involved in phosphatidylinositol trimannoside biosynthesis in Mycobacterium tuberculosis. *Biochem J*. 2002;363(3):437-447. doi:10.1042/bj3630437

214. Morita YS, Sena CBC, Waller RF, et al. PimE Is a Polyprenol-phosphate-mannose-dependent Mannosyltransferase That Transfers the Fifth Mannose of Phosphatidylinositol Mannoside in Mycobacteria*. *J Biol Chem*. 2006;281(35):25143-25155. doi:10.1074/jbc.M604214200

215. Larrouy-Maumus G, Puzo G. Mycobacterial envelope lipids fingerprint from direct MALDI-TOF MS analysis of intact bacilli. Tuberculosis; 95 (2015) 75-85

216. Alexander DC, Jones JRW, Tan T, Chen JM, Liu J. PimF, a Mannosyltransferase of Mycobacteria, Is Involved in the Biosynthesis of Phosphatidylinositol Mannosides and Lipoarabinomannan*. *J Biol Chem*. 2004;279(18):18824-18833. doi:10.1074/jbc.M400791200

217. Burguière A, Hitchen PG, Dover LG, et al. LosA, a Key Glycosyltransferase Involved in the Biosynthesis of a Novel Family of Glycosylated Acyltrehalose Lipooligosaccharides from Mycobacterium marinum*. *J Biol Chem*. 2005;280(51):42124-42133. doi:10.1074/jbc.M507500200

218. Korduláková J, Gilleron M, Puzo G, et al. Identification of the Required Acyltransferase Step in the Biosynthesis of the Phosphatidylinositol Mannosides of Mycobacterium Species*. *J Biol Chem*. 2003;278(38):36285-36295. doi:10.1074/jbc.M303639200

219. Khoo KH, Dell A, Morris HR, Breman PJ, Chatterjee D. Structural definition of acylated phosphatidylinositol mannosides from Mycobacterium tuberculosis: definition of a common

anchor for lipomannan and lipoarabinomannan. *Glycobiology*. 1995;5(1):117-127. doi:10.1093/glycob/5.1.117

220. Gilleron M et al. Structural Study of the LipoMannans from Mycobacterium bovis BCG: Characterisation of Multiacylated Forms of the Phosphatidyl-myo-inositol Anchor. :14.

221. Hsu FF, Turk J, Owens RM, Rhoades ER, Russell DG. Structural characterization of phosphatidyl-myo-inositol mannosides from Mycobacterium bovis Bacillus Calmette Guérin by multiple-stage quadrupole ion-trap mass spectrometry with electrospray ionization. I. PIMs and lyso-PIMs. *J Am Soc Mass Spectrom*. 2007;18(3):466-478. doi:10.1016/j.jasms.2006.10.012

222. Guerin ME, Korduláková J, Alzari PM, Brennan PJ, Jackson M. Molecular Basis of Phosphatidyl-myo-inositol Mannoside Biosynthesis and Regulation in Mycobacteria*. *J Biol Chem*. 2010;285(44):33577-33583. doi:10.1074/jbc.R110.168328

223. Ortalo-Magné A, Lemassu A, Lanéelle MA, et al. Identification of the surface-exposed lipids on the cell envelopes of Mycobacterium tuberculosis and other mycobacterial species. *J Bacteriol*. 1996;178(2):456-461. doi:10.1128/jb.178.2.456-461.1996

224. Torrelles JB, Azad AK, Schlesinger LS. Fine discrimination in the recognition of individual species of phosphatidyl-myo-inositol mannosides from Mycobacterium tuberculosis by C-type lectin pattern recognition receptors. *J Immunol Baltim Md 1950*. 2006;177(3):1805-1816. doi:10.4049/jimmunol.177.3.1805

225. Guerardel Y, Maes E, Elass E, et al. Structural study of lipomannan and lipoarabinomannan from Mycobacterium chelonae. Presence of unusual components with alpha 1,3-mannopyranose side chains. *J Biol Chem*. 2002;277(34):30635-30648. doi:10.1074/jbc.M204398200

226. Palčeková Z, Angala SK, Belardinelli JM, et al. Disruption of the SucT acyltransferase in Mycobacterium smegmatis abrogates succinylation of cell envelope polysaccharides. *J Biol Chem*. 2019;294(26):10325-10335. doi:10.1074/jbc.RA119.008585

227. Shi L, Berg S, Lee A, et al. The Carboxy Terminus of EmbC from Mycobacterium smegmatis Mediates Chain Length Extension of the Arabinan in Lipoarabinomannan*. *J Biol Chem*. 2006;281(28):19512-19526. doi:10.1074/jbc.M513846200

228. Birch HL, Alderwick LJ, Appelmelk BJ, et al. A truncated lipoglycan from mycobacteria with altered immunological properties. *Proc Natl Acad Sci*. 2010;107(6):2634-2639. doi:10.1073/pnas.0915082107

229. Seidel M, Alderwick LJ, Birch HL, Sahm H, Eggeling L, Besra GS. Identification of a Novel Arabinofuranosyltransferase AftB Involved in a Terminal Step of Cell Wall Arabinan Biosynthesis in Corynebacterianeae, such as Corynebacterium glutamicum and Mycobacterium tuberculosis*. *J Biol Chem*. 2007;282(20):14729-14740. doi:10.1074/jbc.M700271200

230. Dinadayala P, Kaur D, Berg S, et al. Genetic basis for the synthesis of the immunomodulatory mannose caps of lipoarabinomannan in Mycobacterium tuberculosis. *J Biol Chem*. 2006;281(29):20027-20035. doi:10.1074/jbc.M603395200

231. Kaur D, Obregón-Henao A, Pham H, Chatterjee D, Brennan PJ, Jackson M. Lipoarabinomannan of Mycobacterium: Mannose capping by a multifunctional terminal mannosyltransferase. *Proc Natl Acad Sci*. 2008;105(46):17973-17977. doi:10.1073/pnas.0807761105

232. Gilleron M, Himoudi N, Adam O, et al. Mycobacterium smegmatis phosphoinositols-glyceroarabinomannans. Structure and localization of alkali-labile and alkali-stable phosphoinositides. *J Biol Chem*. 1997;272(1):117-124. doi:10.1074/jbc.272.1.117

233. Treumann A, Xidong F, McDonnell L, et al. 5-methylthiopentose: a new substituent on lipoarabinomannan in Mycobacterium tuberculosis11Edited by J. Karn. *J Mol Biol*. 2002;316(1):89-100. doi:10.1006/jmbi.2001.5317

234. Guérardel Y, Maes E, Briken V, et al. Lipomannan and Lipoarabinomannan from a Clinical Isolate of Mycobacterium kansasii: NOVEL STRUCTURAL FEATURES AND APOPTOSIS-INDUCING PROPERTIES*. *J Biol Chem*. 2003;278(38):36637-36651. doi:10.1074/jbc.M305427200

235. Shukla S, Richardson ET, Athman JJ, et al. Mycobacterium tuberculosis Lipoprotein LprG Binds Lipoarabinomannan and Determines Its Cell Envelope Localization to Control Phagolysosomal Fusion. *PLOS Pathog*. 2014;10(10):e1004471. doi:10.1371/journal.ppat.1004471

236. Gaur RL, Ren K, Blumenthal A, et al. LprG-Mediated Surface Expression of Lipoarabinomannan Is Essential for Virulence of Mycobacterium tuberculosis. *PLOS Pathog*. 2014;10(9):e1004376. doi:10.1371/journal.ppat.1004376

237. Fukuda T, Matsumura T, Ato M, et al. Critical roles for lipomannan and lipoarabinomannan in cell wall integrity of mycobacteria and pathogenesis of tuberculosis. *mBio*. 2013;4(1):e00472-00412. doi:10.1128/mBio.00472-12

238. Torrelles JB, DesJardin LE, MacNeil J, et al. Inactivation of Mycobacterium tuberculosis mannosyltransferase pimB reduces the cell wall lipoarabinomannan and lipomannan content and increases the rate of bacterial-induced human macrophage cell death. *Glycobiology*. 2009;19(7):743-755. doi:10.1093/glycob/cwp042

239. Turnbull WB, Stalford SA. Methylthioxylose – a jewel in the mycobacterial crown? *Org Biomol Chem*. 2012;10(30):5698-5706. doi:10.1039/C2OB25630D

240. Raymond JB, Mahapatra S, Crick DC, Pavelka MS. Identification of the namH gene, encoding the hydroxylase responsible for the N-glycolylation of the mycobacterial peptidoglycan. *J Biol Chem*. 2005;280(1):326-333. doi:10.1074/jbc.M411006200

241. Lavollay M, Arthur M, Fourgeaud M, et al. The peptidoglycan of stationary-phase Mycobacterium tuberculosis predominantly contains cross-links generated by L,D-transpeptidation. *J Bacteriol*. 2008;190(12):4360-4366. doi:10.1128/JB.00239-08

242. McNeil M, Daffe M, Brennan PJ. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J Biol Chem*. 1990;265(30):18200-18206.

243. Daffe M, Brennan PJ, McNeil M. Predominant structural features of the cell wall arabinogalactan of Mycobacterium tuberculosis as revealed through characterization of oligoglycosyl alditol

fragments by gas chromatography/mass spectrometry and by 1H and 13C NMR analyses. *J Biol Chem*. 1990;265(12):6734-6743.

244. Mikusová K, Belánová M, Korduláková J, et al. Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. *J Bacteriol*. 2006;188(18):6592-6598. doi:10.1128/JB.00489-06

245. Kremer L, Dover LG, Morehouse C, et al. Galactan biosynthesis in Mycobacterium tuberculosis. Identification of a bifunctional UDP-galactofuranosyltransferase. *J Biol Chem*. 2001;276(28):26430-26440. doi:10.1074/jbc.M102022200

246. Alderwick LJ, Seidel M, Sahm H, Besra GS, Eggeling L. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in Mycobacterium tuberculosis. *J Biol Chem*. 2006;281(23):15653-15661. doi:10.1074/jbc.M600045200

247. Skovierová H, Larrouy-Maumus G, Pham H, et al. Biosynthetic origin of the galactosamine substituent of Arabinogalactan in Mycobacterium tuberculosis. *J Biol Chem*. 2010;285(53):41348-41355. doi:10.1074/jbc.M110.188110

248. Dautin N, de Sousa-d'Auria C, Constantinesco-Becker F, et al. Mycoloyltransferases: A large and major family of enzymes shaping the cell envelope of Corynebacteriales. *Biochim Biophys Acta BBA - Gen Subj*. 2017;1861(1, Part B):3581-3592. doi:10.1016/j.bbagen.2016.06.020

249. Asselineau J, Lederer E. Structure of the mycolic acids of Mycobacteria. *Nature*. 1950;166(4227):782-783. doi:10.1038/166782a0

250. Pawełczyk J, Kremer L. The Molecular Genetics of Mycolic Acid Biosynthesis. *Microbiol Spectr*. 2014;2(4):2.4.24. doi:10.1128/microbiolspec.MGM2-0003-2013

251. Marrakchi H, Lanéelle MA, Daffé M. Mycolic Acids: Structures, Biosynthesis, and Beyond. *Chem Biol*. 2014;21(1):67-85. doi:10.1016/j.chembiol.2013.11.011

252. Ausina V, Luquin M, García Barceló M, et al. Mycobacterium alvei sp. nov. *Int J Syst Bacteriol*. 1992;42(4):529-535. doi:10.1099/00207713-42-4-529

253. Villeneuve M, Kawai M, Kanashima H, Watanabe M, Minnikin DE, Nakahara H. Temperature dependence of the Langmuir monolayer packing of mycolic acids from Mycobacterium tuberculosis. *Biochim Biophys Acta BBA - Biomembr*. 2005;1715(2):71-80. doi:10.1016/j.bbamem.2005.07.005

254. Marrakchi H, Lanéelle MA, Daffé M. Mycolic Acids: Structures, Biosynthesis, and Beyond. *Chem Biol*. 2014;21(1):67-85. doi:10.1016/j.chembiol.2013.11.011

255. Bailo R, Radhakrishnan A, Singh A, Nakaya M, Fujiwara N, Bhatt A. The mycobacterial desaturase DesA2 is associated with mycolic acid biosynthesis. *Sci Rep*. 2022;12(1):6943. doi:10.1038/s41598-022-10589-y

256. Glickman MS, Cox JS, Jacobs WR. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of Mycobacterium tuberculosis. *Mol Cell*. 2000;5(4):717-727. doi:10.1016/s1097-2765(00)80250-6

257. Glickman MS. The mmaA2 gene of Mycobacterium tuberculosis encodes the distal cyclopropane synthase of the alpha-mycolic acid. *J Biol Chem*. 2003;278(10):7844-7849. doi:10.1074/jbc.M212458200

258. Glickman MS, Cahill SM, Jacobs WR. The Mycobacterium tuberculosis cmaA2 gene encodes a mycolic acid trans-cyclopropane synthetase. *J Biol Chem*. 2001;276(3):2228-2233. doi:10.1074/jbc.C000652200

259. Behr MA, Schroeder BG, Brinkman JN, Slayden RA, Barry CE. A point mutation in the mma3 gene is responsible for impaired methoxymycolic acid production in Mycobacterium bovis BCG strains obtained after 1927. *J Bacteriol*. 2000;182(12):3394-3399. doi:10.1128/JB.182.12.3394-3399.2000

260. Dubnau E, Chan J, Raynaud C, et al. Oxygenated mycolic acids are necessary for virulence of Mycobacterium tuberculosis in mice. *Mol Microbiol*. 2000;36(3):630-637. doi:10.1046/j.1365-2958.2000.01882.x

261. Barkan D, Rao V, Sukenick GD, Glickman MS. Redundant Function of cmaA2 and mmaA2 in Mycobacterium tuberculosis cis Cyclopropanation of Oxygenated Mycolates. *J Bacteriol*. 2010;192(14):3661-3668. doi:10.1128/JB.00312-10

262. Lefebvre C, Boulon R, Ducoux M, et al. HadD, a novel fatty acid synthase type II protein, is essential for alpha- and epoxy-mycolic acid biosynthesis and mycobacterial fitness. *Sci Rep*. 2018;8(1):6034. doi:10.1038/s41598-018-24380-5

263. Di Capua CB, Belardinelli JM, Carignano HA, Buchieri MV, Suarez CA, Morbidoni HR. Unveiling the Biosynthetic Pathway for Short Mycolic Acids in Nontuberculous Mycobacteria: Mycobacterium smegmatis MSMEG_4301 and Its Ortholog Mycobacterium abscessus MAB_1915 Are Essential for the Synthesis of α′-Mycolic Acids. *Microbiol Spectr*. 10(4):e01288-22. doi:10.1128/spectrum.01288-22

264. Di Capua CB, Belardinelli JM, Buchieri MV, Bortolotti A, Franceschelli JJ, Morbidoni HRY 2018. Deletion of MSMEG_1350 in Mycobacterium smegmatis causes loss of epoxy-mycolic acids, fitness alteration at low temperature and resistance to a set of mycobacteriophages. *Microbiology*. 164(12):1567-1582. doi:10.1099/mic.0.000734

265. Backus KM, Dolan MA, Barry CS, et al. The Three Mycobacterium tuberculosis Antigen 85 Isoforms Have Unique Substrates and Activities Determined by Non-active Site Regions*. *J Biol Chem*. 2014;289(36):25041-25053. doi:10.1074/jbc.M114.581579

266. Kalscheuer R, Weinrick B, Veeraraghavan U, Besra GS, Jacobs WR. Trehalose-recycling ABC transporter LpqY-SugA-SugB-SugC is essential for virulence of Mycobacterium tuberculosis. *Proc Natl Acad Sci*. 2010;107(50):21761-21766. doi:10.1073/pnas.1014642108

267. Rombouts Y, Brust B, Ojha AK, et al. Exposure of Mycobacteria to Cell Wall-inhibitory Drugs Decreases Production of Arabinoglycerolipid Related to Mycolyl-arabinogalactan-peptidoglycan Metabolism. *J Biol Chem*. 2012;287(14):11060-11069. doi:10.1074/jbc.M111.327387

268. Tillequin F. [Trehala, a meeting point between zoology, botany, chemistry, and biochemistry]. *Rev Hist Pharm*. 2009;57(362):163-172.

269. Kalscheuer R, Koliwer-Brandl H. Genetics of Mycobacterial Trehalose Metabolism. *Microbiol Spectr*. 2014;2(3):2.3.04. doi:10.1128/microbiolspec.MGM2-0002-2013

270. Daffé M, Lacave C, Lanéelle MA, Gillois M, Lanéelle G. Polyphthienoyl trehalose, glycolipids specific for virulent strains of the tubercle bacillus. *Eur J Biochem*. 1988;172(3):579-584. doi:10.1111/j.1432-1033.1988.tb13928.x

271. Lemassu A, Lanéelle MA, Daffé M, Revised structure of a trehalose-containing immunoreactive glycolipid of Mycobacterium tuberculosis. FEMS Microbiology Letters 78 (1991) 171-176.

272. Besra GS, Bolton RC, McNeil MR, et al. Structural elucidation of a novel family of acyltrehaloses from Mycobacterium tuberculosis. *Biochemistry*. 1992;31(40):9832-9837. doi:10.1021/bi00155a040

273. Sempere MA, Valero-Guillén PL, De Godos A, Martín-Luengo F 1993. A Triacyltrehalose Containing 2-Methyl-Branched Unsaturated Fatty Acyl Groups Isolated from Mycobacterium Fortuitum. *Microbiology*. 1993;139(3):585-590. doi:10.1099/00221287-139-3-585

274. Ariza MA, Martín-Luengo F, Valero-Guillén PL. A family of diacyltrehaloses isolated from Mycobacterium fortuitum. *Microbiol Read Engl*. 1994;140 (Pt 8):1989-1994. doi:10.1099/13500872-140-8-1989

275. Ariza MA, Valero-Guillén PL. Delineation of molecular species of a family of diacyltrehaloses from Mycobacterium fortuitum by mass spectrometry. *FEMS Microbiol Lett*. 1994;119(3):279-282. doi:10.1111/j.1574-6968.1994.tb06901.x

276. Purdy GE, Hsu FF. Complete Characterization of Polyacyltrehaloses from Mycobacterium tuberculosis H37Rv Biofilm Cultures by Multiple-Stage Linear Ion-Trap Mass Spectrometry Reveals a New Tetraacyltrehalose Family. *Biochemistry*. 2021;60(5):381-397. doi:10.1021/acs.biochem.0c00956

277. Hatzios SK, Schelle MW, Holsclaw CM, et al. PapA3 Is an Acyltransferase Required for Polyacyltrehalose Biosynthesis in Mycobacterium tuberculosis. *J Biol Chem*. 2009;284(19):12745-12751. doi:10.1074/jbc.M809088200

278. Belardinelli JM, Larrouy-Maumus G, Jones V, Sorio de Carvalho LP, McNeil MR, Jackson M. Biosynthesis and Translocation of Unsulfated Acyltrehaloses in Mycobacterium tuberculosis*. *J Biol Chem*. 2014;289(40):27952-27965. doi:10.1074/jbc.M114.581199

279. Saavedra R, Segura E, Leyva R, Esparza LA, López-Marín LM. Mycobacterial Di-O-Acyl-Trehalose Inhibits Mitogen- and Antigen-Induced Proliferation of Murine T Cells In Vitro. *Clin Diagn Lab Immunol*. 2001;8(6):1081-1088. doi:10.1128/CDLI.8.6.1-91-1088.2001

280. Rousseau C, Neyrolles O, Bordat Y, et al. Deficiency in mycolipenate- and mycosanoate-derived acyltrehaloses enhances early interactions of Mycobacterium tuberculosis with host cells. *Cell Microbiol*. 2003;5(6):405-415. doi:10.1046/j.1462-5822.2003.00289.x

281. Goren MB. Sulfolipid I of Mycobacterium tuberculosis, strain H37Rv II. Structural studies. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1970;210(1):127-138. doi:10.1016/0005-2760(70)90068-8

282. Ruhl CR, Pasko BL, Khan HS, et al. Mycobacterium tuberculosis Sulfolipid-1 Activates Nociceptive Neurons and Induces Cough. *Cell*. 2020;181(2):293-305.e11. doi:10.1016/j.cell.2020.02.026

283. Bloch H, Sorkin E, Erlenmeyer H. A Toxic Lipid Component of the Tubercle Bacillus ("Cord Factor")1. *Am Rev Tuberc*. Published online April 23, 2019. Accessed May 21, 2022. http://www.atsjournals.org/doi/pdf/10.1164/art.1953.67.5.629

284. Behling CA, Bennett B, Takayama K, Hunter RL. Development of a trehalose 6,6'-dimycolate model which explains cord formation by Mycobacterium tuberculosis. *Infect Immun*. 1993;61(6):2296-2303. doi:10.1128/iai.61.6.2296-2303.1993

285. Noll H, Bloch H, Asselineau J, Lederer E. The chemical structure of the cord factor of Mycobacterium tuberculosis. *Biochim Biophys Acta*. 1956;20:299-309. doi:10.1016/0006-3002(56)90289-X

286. Mompon B, FEDERICi C, Toubiana R, Lederer E. ISOLATION AND STRUCTURA L DETERMINATION OF A "CORD-FACTOR" (IREFIALOSE 6,6' DIMYCOLATE) FROM MYCOBACTERIUM SM~.GM/~ TIS. :5.

287. Fujita Y, Naka T, McNeil MR, Yano I 2005. Intact molecular characterization of cord factor (trehalose 6,6'-dimycolate) from nine species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology*. 151(10):3403-3416. doi:10.1099/mic.0.28158-0

288. Fujita Y, Naka T, Doi T, Yano I 2005. Direct molecular mass determination of trehalose monomycolate from 11 species of mycobacteria by MALDI-TOF mass spectrometry. *Microbiology*. 151(5):1443-1452. doi:10.1099/mic.0.27791-0

289. Dokic A, Peterson E, Arrieta-Ortiz ML, et al. Mycobacterium abscessus biofilms produce an extracellular matrix and have a distinct mycolic acid profile. *Cell Surf*. 2021;7:100051. doi:10.1016/j.tcsw.2021.100051

290. Belisle JT, Vissa VD, Sievert T, Takayama K, Brennan PJ, Besra GS. Role of the Major Antigen of Mycobacterium tuberculosis in Cell Wall Biogenesis. *Science*. 1997;276(5317):1420-1422. doi:10.1126/science.276.5317.1420

291. Matsunaga I, Naka T, Talekar RS, et al. Mycolyltransferase-mediated Glycolipid Exchange in Mycobacteria∗. *J Biol Chem*. 2008;283(43):28835-28841. doi:10.1074/jbc.M805776200

292. Indrigo J, Hunter RL, Actor JKY 2002. Influence of trehalose 6,6'-dimycolate (TDM) during mycobacterial infection of bone marrow macrophages. *Microbiology*. 148(7):1991-1998. doi:10.1099/00221287-148-7-1991

293. Indrigo J, Hunter RL, Actor JK. Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology*. 2003;149(8):2049-2059. doi:10.1099/mic.0.26226-0

294. Ishikawa E, Ishikawa T, Morita YS, et al. Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J Exp Med*. 2009;206(13):2879-2888. doi:10.1084/jem.20091750

295. Feinberg H, Jégouzo SAF, Rowntree TJW, et al. Mechanism for Recognition of an Unusual Mycobacterial Glycolipid by the Macrophage Receptor Mincle*. *J Biol Chem*. 2013;288(40):28457-28465. doi:10.1074/jbc.M113.497149

296. Hunter RL, Venkataprasad N, Olsen MR. The role of trehalose dimycolate (cord factor) on morphology of virulent M. tuberculosis in vitro. *Tuberculosis*. 2006;86(5):349-356. doi:10.1016/j.tube.2005.08.017

297. Asselineau CP, Montrozier HL. Etude du processus de biosynthese des acides phléiques, acides polyinsaturés synthétisés par Mycobacterium phlei. *Eur J Biochem*. 1976;63(2):509-518. doi:10.1111/j.1432-1033.1976.tb10254.x

298. Burbaud S, Laval F, Lemassu A, Daffé M, Guilhot C, Chalut C. Trehalose Polyphleates Are Produced by a Glycolipid Biosynthetic Pathway Conserved across Phylogenetically Distant Mycobacteria. *Cell Chem Biol*. 2016;23(2):278-289. doi:10.1016/j.chembiol.2015.11.013

299. Thouvenel L, Prevot G, Chiaradia L, et al. The final assembly of trehalose polyphleates takes place within the outer layer of the mycobacterial cell envelope. *J Biol Chem*. 2020;295(32):11184-11194. doi:10.1074/jbc.RA120.013299

300. Hunter SW, Murphy RC, Clay K, Goren MB, Brennan PJ. Trehalose-containing lipooligosaccharides. A new class of species-specific antigens from Mycobacterium. *J Biol Chem*. 1983;258(17):10481-10487. doi:10.1016/S0021-9258(17)44481-4

301. Nataraj V, Pang PC, Haslam S, et al. MKAN27435 Is Required for the Biosynthesis of Higher Subclasses of Lipooligosaccharides in Mycobacterium kansasii. *PLOS ONE*. 2015;10:e0122804. doi:10.1371/journal.pone.0122804

302. Saadat S, Ballou CE. Pyruvylated glycolipids from Mycobacterium smegmatis. Structures of two oligosaccharide components. *J Biol Chem*. 1983;258(3):1813-1818. doi:10.1016/S0021-9258(18)33060-6

303. Besra GS, Khoo KH, Belisle JT, et al. New pyruvylated, glycosylated acyltrehaloses from Mycobacterium smegmatis strains, and their implications for phage resistance in mycobacteria. *Carbohydr Res*. 1994;251:99-114. doi:10.1016/0008-6215(94)84279-5

304. Ren H, Dover LG, Islam ST, et al. Identification of the lipooligosaccharide biosynthetic gene cluster from Mycobacterium marinum. *Mol Microbiol*. 2007;63(5):1345-1359. doi:10.1111/j.1365-2958.2007.05603.x

305. Rombouts Y, Burguière A, Maes E, et al. Mycobacterium marinum Lipooligosaccharides Are Unique Caryophyllose-containing Cell Wall Glycolipids That Inhibit Tumor Necrosis Factor-α Secretion in Macrophages. *J Biol Chem*. 2009;284(31):20975-20988. doi:10.1074/jbc.M109.011429

306. Rombouts Y, Elass E, Biot C, et al. Structural Analysis of an Unusual Bioactive N-Acylated Lipo-Oligosaccharide LOS-IV in Mycobacterium marinum. *J Am Chem Soc*. 2010;132(45):16073-16084. doi:10.1021/ja105807s

307. Daffe M, McNeil M, Brennan PJ. Novel type-specific lipooligosaccharides from Mycobacterium tuberculosis. *Biochemistry*. 1991;30(2):378-388. doi:10.1021/bi00216a011

308. van der Woude AD, Sarkar D, Bhatt A, et al. Unexpected Link between Lipooligosaccharide Biosynthesis and Surface Protein Release in Mycobacterium marinum. *J Biol Chem*. 2012;287(24):20417-20429. doi:10.1074/jbc.M111.336461

309. Alibaud L, Pawelczyk J, Gannoun-Zaki L, et al. Increased Phagocytosis of Mycobacterium marinum Mutants Defective in Lipooligosaccharide Production: A STRUCTURE-ACTIVITY RELATIONSHIP STUDY*. *J Biol Chem*. 2014;289(1):215-228. doi:10.1074/jbc.M113.525550

310. Papa F, Cruaud P, Luquin M, Thorel MF, Goh KS, David HL. Isolation and characterization of serologically reactive lipooligosaccharides from Mycobacterium tuberculosis. *Res Microbiol*. 1993;144(2):91-99. doi:10.1016/0923-2508(93)90021-S

311. Lemassu A, Lévy-Frébault VV, Lanéelle MA, Daffé M. Lack of correlation between colony morphology and lipooligosaccharide content in the Mycobacterium tuberculosis complex. *J Gen Microbiol*. 1992;138(7):1535-1541. doi:10.1099/00221287-138-7-1535

312. Laneelle G, Asselineau J. Structure d'un glycoside de peptidolipide isolé d'une mycobactérie. *Eur J Biochem*. 1968;5(4):487-491. doi:10.1111/j.1432-1033.1968.tb00396.x

313. Rivière M, Puzo G. A new type of serine-containing glycopeptidolipid from Mycobacterium xenopi. *J Biol Chem*. 1991;266(14):9057-9063. doi:10.1016/S0021-9258(18)31552-7

314. Riviere M, Puzo G. Use of proton NMR ROESY for structural determination of O-glycosylated amino acids from a serine-containing glycopeptidolipid antigen. *Biochemistry*. 1992;31(14):3575-3580. doi:10.1021/bi00129a004

315. Besra GS, McNeil MR, Rivoire B, et al. Further structural definition of a new family of glycopeptidolipids from Mycobacterium xenopi. *Biochemistry*. 1993;32(1):347-355. doi:10.1021/bi00052a043

316. Draper PY 1974. The Mycoside Capsule of Mycobacterium avium 357. *Microbiology*. 83(2):431-433. doi:10.1099/00221287-83-2-431

317. Ichimura N, Kasama T. Identification of Valine- or Leucine-Containing Glycopeptidolipids from Mycobacterium avium–intracellulare Complex. *Curr Microbiol*. 2012;64(6):561-568. doi:10.1007/s00284-012-0107-6

318. Eckstein TM, Silbaq FS, Chatterjee D, Kelly NJ, Brennan PJ, Belisle JT. Identification and Recombinant Expression of a Mycobacterium avium Rhamnosyltransferase Gene (rtfA) Involved in Glycopeptidolipid Biosynthesis. *J Bacteriol*. 1998;180(21):5567-5573.

319. Miyamoto Y, Mukai T, Maeda Y, et al. Characterization of the Fucosylation Pathway in the Biosynthesis of Glycopeptidolipids from Mycobacterium avium Complex. *J Bacteriol*. 2007;189(15):5515-5522. doi:10.1128/JB.00344-07

320. Sweet L, Zhang W, Torres-Fewell H, Serianni A, Boggess W, Schorey J. Mycobacterium avium Glycopeptidolipids Require Specific Acetylation and Methylation Patterns for Signaling through Toll-like Receptor 2*. *J Biol Chem*. 2008;283(48):33221-33231. doi:10.1074/jbc.M805539200

321. Miyamoto Y, Mukai T, Naka T, et al. Novel Rhamnosyltransferase Involved in Biosynthesis of Serovar 4-Specific Glycopeptidolipid from Mycobacterium avium Complex. *J Bacteriol*. 2010;192(21):5700-5708. doi:10.1128/JB.00554-10

322. Barrow W, Davis T, Wright E, Labrousse V, Bachelet M, Rastogi N, "Immunomodulatory spectrum of lipids associated with Mycobacterium avium serovar 8" Infection and Immunity (1995, Vol. 63, n°1, 126-133).

323. Fujiwara N, Nakata N, Maeda S, et al. Structural Characterization of a Specific Glycopeptidolipid Containing a Novel N-Acyl-Deoxy Sugar from Mycobacterium intracellulare Serotype 7 and Genetic Analysis of Its Glycosylation Pathway. *J Bacteriol*. 2007;189(3):1099-1108. doi:10.1128/JB.01471-06

324. Miyamoto Y, Mukai T, Maeda Y, et al. The Mycobacterium avium Complex gtfTB Gene Encodes a Glucosyltransferase Required for the Biosynthesis of Serovar 8-Specific Glycopeptidolipid. *J Bacteriol*. 2008;190(24):7918-7924. doi:10.1128/JB.00911-08

325. Khoo KH, Jarboe E, Barker A, Torrelles J, Kuo CW, Chatterjee D. Altered Expression Profile of the Surface Glycopeptidolipids in Drug-resistant Clinical Isolates of Mycobacterium aviumComplex*. *J Biol Chem*. 1999;274(14):9778-9785. doi:10.1074/jbc.274.14.9778

326. Marin LML, Laneelle MA, Prome D, Daffe M, Laneelle G, Prome JC. Glycopeptidolipids from Mycobacterium fortuitum: a variant in the structure of C-mycoside. ACS Publications. doi:10.1021/bi00107a024

327. Lopez Marin LM, Laneelle MA, Prome D, Laneelle G, Prome JC, Daffe M. Structure of a novel sulfate-containing mycobacterial glycolipid. *Biochemistry*. 1992;31(45):11106-11111. doi:10.1021/bi00160a021

328. Lopez-Marin LM, Quesada D, Lakhdar-Ghazal F, Tocanne JF, Laneelle G. Interactions of Mycobacterial Glycopeptidolipids with Membranes: Influence of Carbohydrate on Induced Alterations. *Biochemistry*. 1994;33(23):7056-7061. doi:10.1021/bi00189a006

329. Hamid ME, Ridell M, Minnikin DE, Goodfellow M. Serotaxonomic Analysis of Glycolipids from Mycobacterium chelonae-M. fortuitum Complex and Bovine Farcy Strains. *Zentralblatt Für Bakteriol*. 1998;288(1):23-34. doi:10.1016/S0934-8840(98)80094-7

330. Khoo KH, Chatterjee D, Dell A, Morris HR, Brennan PJ, Draper P. Novel O-Methylated Terminal Glucuronic Acid Characterizes the Polar Glycopeptidolipids of Mycobacterium habana Strain TMC 5135. *J Biol Chem*. 1996;271(21):12333-12342. doi:10.1074/jbc.271.21.12333

331. Mederos LM, Valdivia JA, Sempere MA, Valero-Guillén PL 1998. Analysis of lipids reveals differences between 'Mycobacterium habana' and Mycobacterium simiae. *Microbiology*. 144(5):1181-1188. doi:10.1099/00221287-144-5-1181

332. Mederos L, Valdivia J a., Valero-Guillén P l. New variants of polar glycopeptidolipids detected in Mycobacterium simiae, including 'habana' strains, as evidenced by electrospray ionization-ion trap-mass spectrometry. *J Appl Microbiol*. 2008;105(2):602-614. doi:10.1111/j.1365-2672.2008.03796.x

333. Daffe M, Laneelle MA, Puzo G. Structural elucidation by field desorption and electron-impact mass spectrometry of the c-mycosides isolated from Mycoba cterium smegma tis. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1983;751(3):439-443. doi:10.1016/0005-2760(83)90304-1

334. Billman-Jacobe H, McConville MJ, Haites RE, Kovacevic S, Coppel RL. Identification of a peptide synthetase involved in the biosynthesis of glycopeptidolipids of Mycobacterium smegmatis. *Mol Microbiol*. 1999;33(6):1244-1253. doi:10.1046/j.1365-2958.1999.01572.x

335. Miyamoto Y, Mukai T, Nakata N, et al. Identification and Characterization of the Genes Involved in Glycosylation Pathways of Mycobacterial Glycopeptidolipid Biosynthesis. *J Bacteriol*. 2006;188(1):86-95. doi:10.1128/JB.188.1.86-95.2006

336. Ojha AKr, Varma S, Chatterji DY 2002. Synthesis of an unusual polar glycopeptidolipid in glucose-limited culture of Mycobacterium smegmatis. *Microbiology*. 148(10):3039-3048. doi:10.1099/00221287-148-10-3039

337. Mukherjee R, Gomez M, Jayaraman N, Smith I, Chatterji D 2005. Hyperglycosylation of glycopeptidolipid of Mycobacterium smegmatis under nutrient starvation: structural studies. *Microbiology*. 151(7):2385-2392. doi:10.1099/mic.0.27908-0

338. Bakli M, Karim L, Mokhtari-Soulimane N, Merzouk H, Vincent F. Biochemical characterization of a glycosyltransferase Gtf3 from Mycobacterium smegmatis: a case study of improved protein solubilization. *3 Biotech*. 2020;10(10):436. doi:10.1007/s13205-020-02431-x

339. Patterson JH, McConville MJ, Haites RE, Coppel RL, Billman-Jacobe H. Identification of a Methyltransferase from Mycobacterium smegmatis Involved in Glycopeptidolipid Synthesis*. *J Biol Chem*. 2000;275(32):24900-24906. doi:10.1074/jbc.M000147200

340. Jeevarajah D, Patterson JH, Taig E, Sargeant T, McConville MJ, Billman-Jacobe H. Methylation of GPLs in Mycobacterium smegmatis and Mycobacterium avium. *J Bacteriol*. 2004;186(20):6792-6799. doi:10.1128/JB.186.20.6792-6799.2004

341. Jeevarajah D, Patterson JH, McConville MJ, Billman-Jacobe H 2002. Modification of glycopeptidolipids by an O-methyltransferase of Mycobacterium smegmatisaaThe GenBank accession number for the sequence determined in this work is AY138899. *Microbiology*. 148(10):3079-3087. doi:10.1099/00221287-148-10-3079

342. Vats A, Singh AK, Mukherjee R, et al. Retrobiosynthetic Approach Delineates the Biosynthetic Pathway and the Structure of the Acyl Chain of Mycobacterial Glycopeptidolipids*. *J Biol Chem*. 2012;287(36):30677-30687. doi:10.1074/jbc.M112.384966

343. Recht J, Kolter R. Glycopeptidolipid Acetylation Affects Sliding Motility and Biofilm Formation in Mycobacterium smegmatis. *J Bacteriol*. 2001;183(19):5718-5724. doi:10.1128/JB.183.19.5718-5724.2001

344. Gao J, Sampson NS. A GMC Oxidoreductase Homologue Is Required for Acetylation of Glycopeptidolipid in Mycobacterium smegmatis. *Biochemistry*. 2014;53(4):611-613. doi:10.1021/bi4015083

345. Villeneuve C, Etienne G, Abadie V, et al. Surface-exposed Glycopeptidolipids of Mycobacterium smegmatis Specifically Inhibit the Phagocytosis of Mycobacteria by Human Macrophages: IDENTIFICATION OF A NOVEL FAMILY OF GLYCOPEPTIDOLIPIDS*. *J Biol Chem*. 2003;278(51):51291-51300. doi:10.1074/jbc.M306554200

346. Villeneuve C, Gilleron M, Maridonneau-Parini I, Daffé M, Astarie-Dequeker C, Etienne G. Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *J Lipid Res*. 2005;46(3):475-483. doi:10.1194/jlr.M400308-JLR200

347. Lopez-Marin LM, Gautier N, Laneelle MA, Silve G, Daffe M. Structures of the glycopeptidolipid antigens of Mycobacterium abscessus and Mycobacterium chelonae and possible chemical basis of the serological cross-reactions in the Mycobacterium fortuitum complex. *Microbiology*. 1994;140(5):1109-1118. doi:10.1099/13500872-140-5-1109

348. Ripoll F, Deshayes C, Pasek S, et al. Genomics of glycopeptidolipid biosynthesis in Mycobacterium abscessus and M. chelonae. *BMC Genomics*. 2007;8(1):114. doi:10.1186/1471-2164-8-114

349. Wiersma CJ, Belardinelli JM, Avanzi C, et al. Cell Surface Remodeling of Mycobacterium abscessus under Cystic Fibrosis Airway Growth Conditions. *ACS Infect Dis*. 2020;6(8):2143-2154. doi:10.1021/acsinfecdis.0c00214

350. Bernut A, Viljoen A, Dupont C, et al. Insights into the smooth-to-rough transitioning in Mycobacterium bolletii unravels a functional Tyr residue conserved in all mycobacterial MmpL family members. *Mol Microbiol*. 2016;99(5):866-883. doi:10.1111/mmi.13283

351. Viljoen A, Gutiérrez AV, Dupont C, Ghigo E, Kremer L. A Simple and Rapid Gene Disruption Strategy in Mycobacterium abscessus: On the Design and Application of Glycopeptidolipid Mutants. *Front Cell Infect Microbiol*. 2018;8. Accessed April 18, 2022. https://www.frontiersin.org/article/10.3389/fcimb.2018.00069

352. Deshayes C, Bach H, Euphrasie D, et al. MmpS4 promotes glycopeptidolipids biosynthesis and export in Mycobacterium smegmatis. *Mol Microbiol*. 2010;78(4):989-1003. doi:10.1111/j.1365-2958.2010.07385.x

353. Sondén B, Kocíncová D, Deshayes C, et al. Gap, a mycobacterial specific integral membrane protein, is required for glycolipid transport to the cell surface. *Mol Microbiol*. 2005;58(2):426-440. doi:10.1111/j.1365-2958.2005.04847.x

354. Whang J, Back YW, Lee KI, et al. Mycobacterium abscessus glycopeptidolipids inhibit macrophage apoptosis and bacterial spreading by targeting mitochondrial cyclophilin D. *Cell Death Dis*. 2017;8(8):e3012-e3012. doi:10.1038/cddis.2017.420

355. Clary G, Sasindran SJ, Nesbitt N, et al. Mycobacterium abscessus Smooth and Rough Morphotypes Form Antimicrobial-Tolerant Biofilm Phenotypes but Are Killed by Acetic Acid. *Antimicrob Agents Chemother*. 2018;62(3):e01782-17. doi:10.1128/AAC.01782-17

356. Vilkas E, Miquel AM, Lederer E. [On the isolation and structure of fortuitine, peptidolipid from Mycobacterium fortuitum]. *Biochim Biophys Acta*. 1963;70:217-218. doi:10.1016/0006-3002(63)90745-5

357. Barber M, Jollěs P, Vilkas E, Lederer E. Determination of amino acid sequences in oligopeptides by mass spectrometry. I. The structure of fortuitine, an acyl-nonapeptide methyl ester. *Biochem Biophys Res Commun*. 1965;18(4):469-473. doi:10.1016/0006-291X(65)90775-8

358. Ptak M, Vilkas E, Brevard C. A 400 MHz 1H NMR study of fortuitin, a natural linear lipopeptide. *Biochem Biophys Res Commun*. 1983;113(1):121-126. doi:10.1016/0006-291X(83)90440-0

359. Riviere M, Puzo G, Wright E, Barrow W, "A unique phenylalanine-containing lipopeptide isolated from a rough-colony variant of Mycobacterium avium." *Eur J Biochem*. 1996;241(2):682-690. doi:10.1111/j.1432-1033.1996.00682.x

360. Laneelle G. Mise en evidence d'une conformation stable d'un peptidolipide. *FEBS Lett*. 1969;4(3):210-212. doi:10.1016/0014-5793(69)80237-1

361. Eckstein TM, Chandrasekaran S, Mahapatra S, et al. A Major Cell Wall Lipopeptide of Mycobacterium avium subspecies paratuberculosis*. *J Biol Chem*. 2006;281(8):5209-5215. doi:10.1074/jbc.M512465200

362. Biet F, Bay S, Thibault VC, et al. Lipopentapeptide induces a strong host humoral response and distinguishes Mycobacterium avium subsp. paratuberculosis from M. avium subsp. avium. *Vaccine*. 2008;26(2):257-268. doi:10.1016/j.vaccine.2007.10.059

363. Bannantine J, Etienne G, Laval F, Stabel J, Lemassu A, Daffé M, Bayles D, Ganneau C, Bonhomme F, Branger M, Cochard T, Bay S, Biet F "Cell wall peptidolipids of Mycobacterium avium: from genetic prediction to exact structure of a nonribosomal peptide" Molecular Microbiology (2017, 105 (4), 525-539) doi:10.1111/mmi.13717.

364. Bay S, Begg D, Ganneau C, et al. Engineering Synthetic Lipopeptide Antigen for Specific Detection of Mycobacterium avium subsp. paratuberculosis Infection. *Front Vet Sci*. 2021;8. https://www.frontiersin.org/article/10.3389/fvets.2021.637841

365. Fregnan GB, Smith DW, Randall HM. BIOLOGICAL AND CHEMICAL STUDIES ON MYCOBACTERIA. *J Bacteriol*. 1961;82(4):517-527.

366. Demarteau-Ginsburg H, Lederer E. Sur la structure chimique du mycoside B. *Biochim Biophys Acta BBA - Spec Sect Lipids Relat Subj*. 1963;70:442-451. doi:10.1016/0926-6542(63)90063-5

367. Navalkar RG, Wiegeshaus E, Kondo E, Kim HK, Smith DW. Mycoside G, a Specific Glycolipid in Mycobacterium marinum (Balnei). *J Bacteriol*. 1965;90(1):262-265.

368. MacLennan AP, Randall HM, Smith DW. The occurrence of methyl ethers of rhamnose and fucose in specific glycolipids of certain mycobacteria. *Biochem J*. 1961;80(2):309-318.

369. Gastambide-Odier M, Sarda P. Contribution à l'étude de la structure et de la biosynthèse de glycolipides spécifiques isolés de mycobactéries: les mycosides A et B. *Pneumologie*. 1970;142(2):241-255. doi:10.1007/BF02095222

370. Fournié JJ, Rivière M, Puzo G. Structural elucidation of the major phenolic glycolipid from Mycobacterium kansasii. I. Evidence for tetrasaccharide structure of the oligosaccharide moiety. *J Biol Chem*. 1987;262(7):3174-3179. doi:10.1016/S0021-9258(18)61486-3

371. Fournié JJ, Rivière M, Papa F, Puzo G. Structural elucidation of the major phenolic glycolipid from Mycobacterium kansasii. II. Presence of a novel dideoxyhexose. *J Biol Chem*. 1987;262(7):3180-3184. doi:10.1016/S0021-9258(18)61487-5

372. Gilleron M, Venisse A, Rivière M, Servin P, Puzo G. Carbohydrate epitope structural elucidation by 1H-NMR spectroscopy of a new Mycobacterium kansasii phenolic glycolipid antigen. *Eur J Biochem*. 1990;193(2):449-457. doi:10.1111/j.1432-1033.1990.tb19359.x

373. Watanabe M, Aoyagi Y, Ohta A, Minnikin DE. Structures of phenolic glycolipids from Mycobacterium kansasii. *Eur J Biochem*. 1997;248(1):93-98. doi:10.1111/j.1432-1033.1997.00093.x

374. Daffé M, Lanéelle MA, Lacave C, Lanéelle G. Monoglycosyldiacylphenol-phthiocerol of Mycobacterium tuberculosis and Mycobacterium bovis. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1988;958(3):443-449. doi:10.1016/0005-2760(88)90230-5

375. Chatterjee D, Bozic CM, Knisley C, Cho SN, Brennan PJ. Phenolic glycolipids of Mycobacterium bovis: new structures and synthesis of a corresponding seroreactive neoglycoprotein. *Infect Immun*. 1989;57(2):322-330.

376. Vercellone A, Puzo G. New-found phenolic glycolipids in Mycobacterium bovis BCG: Presence of a diglycosylated glycolipid. *J Biol Chem*. 1989;264(13):7447-7454. doi:10.1016/S0021-9258(18)83254-9

377. Daffé M, Lacave C, Lanéelle MA, Lanéelle G. Structure of the major triglycosyl phenol-phthiocerol of Mycobacterium tuberculosis (strain Canetti). *Eur J Biochem*. 1987;167(1):155-160. doi:10.1111/j.1432-1033.1987.tb13317.x

378. Daffé M, Servin P. Scalar, dipolar-correlated and J-resolved 2D-NMR spectroscopy of the specific phenolic mycoside of Mycobacterium tuberculosis. *Eur J Biochem*. 1989;185(1):157-162. doi:10.1111/j.1432-1033.1989.tb15097.x

379. Daffé M. Further specific triglycosyl phenol phthiocerol diester from Mycobacterium tuberculosis. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1989;1002(2):257-260. doi:10.1016/0005-2760(89)90295-6

380. Watanabe M, Yamada Y, Iguchi K, Minnikin DE. Structural elucidation of new phenolic glycolipids from Mycobaclerium tuberculosis. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1994;1210(2):174-180. doi:10.1016/0005-2760(94)90118-X

381. Pérez E, Constant P, Lemassu A, Laval F, Daffé M, Guilhot C. Characterization of Three Glycosyltransferases Involved in the Biosynthesis of the Phenolic Glycolipid Antigens from the Mycobacterium tuberculosis Complex*. *J Biol Chem*. 2004;279(41):42574-42583. doi:10.1074/jbc.M406246200

382. Sarda P, Gastambide-Odier M. Structure chimique de l'aglycone du mycoside g de mycobacterium marinum. *Chem Phys Lipids*. 1967;1(5):434-444. doi:10.1016/0009-3084(67)90022-9

383. Dobson G, Minnikin DE, Besra GS, Mallet AI, Magnusson M. Characterisation of phenolic glycolipids from Mycobacterium marinum. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1990;1042(2):176-181. doi:10.1016/0005-2760(90)90004-H

384. Daffé M, Varnerot A, Lévy-Frébault VV. The phenolic mycoside of Mycobacterium ulcerans: structure and taxonomic implications. *J Gen Microbiol*. 1992;138(1):131-137. doi:10.1099/00221287-138-1-131

385. Hunter SW, Brennan PJ. A novel phenolic glycolipid from Mycobacterium leprae possibly involved in immunogenicity and pathogenicity. *J Bacteriol*. 1981;147(3):728-735.

386. Hunter SW, Fujiwara T, Brennan PJ. Structure and antigenicity of the major specific glycolipid antigen of Mycobacterium leprae. *J Biol Chem*. 1982;257(24):15072-15078. doi:10.1016/S0021-9258(18)33395-7

387. Hunter SW, Brennan PJ. Further specific extracellular phenolic glycolipid antigens and a related diacylphthiocerol from Mycobacterium leprae. *J Biol Chem*. 1983;258(12):7556-7562. doi:10.1016/S0021-9258(18)32214-2

388. Daffé M, Lanéelle MA. Diglycosyl phenol phthiocerol diester of Mycobacterium leprae. *Biochim Biophys Acta BBA - Lipids Lipid Metab*. 1989;1002(3):333-337. doi:10.1016/0005-2760(89)90347-0

389. Besra GS. Preparation of cell-wall fractions from mycobacteria. *Methods Mol Biol Clifton NJ*. 1998;101:91-107. doi:10.1385/0-89603-471-2:91

390. Daher W, Leclercq LD, Viljoen A, et al. O-Methylation of the Glycopeptidolipid Acyl Chain Defines Surface Hydrophobicity of Mycobacterium abscessus and Macrophage Invasion. *ACS Infect Dis*. 2020;6(10):2756-2770. doi:10.1021/acsinfecdis.0c00490

391. Daher W, Leclercq LD, Johansen MD, et al. Glycopeptidolipid glycosylation controls surface properties and pathogenicity in Mycobacterium abscessus. *Cell Chem Biol*. Published online March 28, 2022:S2451-9456(22)00125-8. doi:10.1016/j.chembiol.2022.03.008

392. Eckstein TM, Belisle JT, Inamine JM. Proposed pathway for the biosynthesis of serovar-specific glycopeptidolipids in Mycobacterium avium serovar 2. *Microbiology*. 2003;149(10):2797-2807. doi:10.1099/mic.0.26528-0

393. Rao V, Fujiwara N, Porcelli SA, Glickman MS. Mycobacterium tuberculosis controls host innate immune activation through cyclopropane modification of a glycolipid effector molecule. *J Exp Med*. 2005;201(4):535-543. doi:10.1084/jem.20041668

394. Rao V, Gao F, Chen B, Jacobs WR, Glickman MS. Trans -cyclopropanation of mycolic acids on trehalose dimycolate suppresses Mycobacterium tuberculosis –induced inflammation and virulence. *J Clin Invest*. 2006;116(6):1660-1667. doi:10.1172/JCI27335

395. Rainwater DL, Kolattukudy PE. Fatty acid biosynthesis in Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin. Purification and characterization of a novel fatty acid synthase, mycocerosic acid synthase, which elongates n-fatty acyl-CoA with methylmalonyl-CoA. *J Biol Chem*. 1985;260(1):616-623. doi:10.1016/S0021-9258(18)89777-0

396. Fernandes ND, Kolattukudy PE. A Newly Identified Methyl-branched Chain Fatty Acid Synthesizing Enzyme from Mycobacterium tuberculosis var. bovis BCG*. *J Biol Chem*. 1998;273(5):2823-2828. doi:10.1074/jbc.273.5.2823

397. Rens C, Chao J, Sexton D, Tocheva E, Av-Gay Y "Roles for phthiocerol dimycocerosate lipids in Mycobacterium tuberculosis pathogenesis" Microbiology (2021; 167: 001042) DOI 10.1099/mic.0.001042

398. Okamoto Y, Fujita Y, Naka T, Hirai M, Tomiyasu I, Yano I. Mycobacterial sulfolipid shows a virulence by inhibiting cord factor induced granuloma formation and TNF-α release. *Microb Pathog*. 2006;40(6):245-253. doi:10.1016/j.micpath.2006.02.002

399. Miura M, Nonomura Y. Evaluation of Hydrophobicity for Fatty Acids Using Reversed-phase Thin Layer Chromatography. *J Oleo Sci*. 2019;68(7):665-670. doi:10.5650/jos.ess19016

400. Kocíncová D, Singh AK, Beretti JL, et al. Spontaneous transposition of IS1096 or ISMsm3 leads to glycopeptidolipid overproduction and affects surface properties in Mycobacterium smegmatis. *Tuberculosis*. 2008;88(5):390-398. doi:10.1016/j.tube.2008.02.005

401. Toki S, Agatsuma T, Ochiai K, et al. RP-1776, a Novel Cyclic Peptide Produced by Streptomyces sp., Inhibits the Binding of PDGF to the Extracellular Domain of Its Receptor. *J Antibiot (Tokyo)*. 2001;54(5):405-414. doi:10.7164/antibiotics.54.405

402. Pohle S, Appelt C, Roux M, Fiedler HP, Süssmuth RD. Biosynthetic Gene Cluster of the Non-ribosomally Synthesized Cyclodepsipeptide Skyllamycin: Deciphering Unprecedented Ways of Unusual Hydroxylation Reactions. *J Am Chem Soc*. 2011;133(16):6194-6205. doi:10.1021/ja108971p

403. Süssmuth RD, Mainz A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew Chem Int Ed*. 2017;56(14):3770-3821. doi:10.1002/anie.201609079

404. Lewin A, Kamal E, Semmler T, et al. Genetic diversification of persistent Mycobacterium abscessus within cystic fibrosis patients. *Virulence*. 2021;12(1):2415-2429. doi:10.1080/21505594.2021.1959808

405. Boeck L, Burbaud S, Skwark M, et al. Mycobacterium abscessus pathogenesis identified by phenogenomic analyses. *Nat Microbiol*. 2022;7(9):1431-1441. doi:10.1038/s41564-022-01204-x

406. Tam WO, Wong CF, Wong SSY, Kwan CLY. Serological diagnosis of Mycobacterium avium complex lung diseases by enzyme immunoassay of IgA antibodies against MAC-specific glycopeptidolipid core antigen. *J Clin Tuberc Mycobact Dis*. 2020;21:100189. doi:10.1016/j.jctube.2020.100189

407. Le Moigne V, Roux AL, Jobart-Malfait A, et al. A TLR2-Activating Fraction From Mycobacterium abscessus Rough Variant Demonstrates Vaccine and Diagnostic Potential. *Front Cell Infect Microbiol*. 2020;10:432. doi:10.3389/fcimb.2020.00432

408. Schlesinger LS, Hull SR, Kaufman TM. Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages. *J Immunol Baltim Md 1950*. 1994;152(8):4070-4079.

RESUME

Mycobacterium abscessus subs abscessus (Mabs) est une mycobactérie opportuniste qui appartient au complexe MABSC avec *subs massiliense* (*Mmas*) et *subs bolletti (Mbol)*. Au sein de ce complexe, les études épidémiologiques démontrent que trois clones dominants émergent dans le monde, majoritairement parmi les patients atteints de mucoviscidose, ainsi que chez les individus sains. De fait, *Mabs* provoque principalement des pathologies pulmonaires, et plus rarement des infections des tissus cutanés, mous ou nerveux. Cette espèce qui est la plus résistante aux antibiotiques est décrite sous deux formes, lisse (S) et rugueuse (R), qui diffèrent par leur virulence et par la composition glycolipidique de leur enveloppe. En effet, la production de grande quantité de glycopeptidolipides (GPLs) chez *Mabs* S est associée à une faible virulence. A l'inverse, ces mêmes GPLs ne se retrouvent pas dans l'enveloppe de la forme rugueuse qui se montre très virulente, plus tardivement lors du cycle infectieux.

Afin de mieux comprendre la pathogénicité de *Mabs*, notre étude a consisté en la description exhaustive des structures chimiques des glycolipides pariétaux dont on connait leur effet immunomodulateur. L'extraction et la purification par chromatographie suivies d'analyses par spectrométrie de masse et par résonance magnétique nucléaire, ont mis en évidence deux familles de composés communes à *Mabs* R et S. Tout d'abord, la description de di-*O*-acyl-tréhaloses originaux révèle la présence possible d'acides mycoliques α et α' identiques à ceux caractérisés dans le tréhalose dimycolate (TDM) et d'acides gras à chaine courte et polyméthylbranchée, de même qu'une variation dans le positionnement des lipides sur le tréhalose en 2,3 ou en 3,4. D'autre part, la résolution de la structure des phosphatidyl *myo*-inositol dimannosides (PIM_2), à l'origine de la synthèse des lipoglycannes, indique la présence de 2-*O*-méthyl mannose.

De plus, nos travaux ont permis de préciser les voies privilégiées de production des GPLs grâce à la comparaison des structures retrouvées chez *Mabs* S sauvage et les mutants ciblés au locus de leur biosynthèse. Le 1er gène *fmt* est impliqué dans la modification du lipide, l'hydrophobicité de l'enveloppe, et l'invasion des macrophages. Parmi les trois gènes codant les glycosyltransférases, *gtf2* et *gtf3* ont respectivement un rôle d'initiation de la glycosylation et de synthèse du GPL-3 qui est la forme la plus phagocytée *in-vitro*. En outre, la poursuite de nos travaux sur une 3ème série de gènes codant les acétyltransférases dont *atf1* et *atf2,* démontrent non seulement l'implication de *MAB_1725c* et *MAB_3448c* hors du locus de biosynthèse des GPLs, dans l'acétylation de ces derniers, mais aussi le rôle prépondérant d'*atf2* dans la régulation du ratio GPL-2a/GPL-3.

Enfin, nos travaux se sont focalisés sur l'étude d'un mutant rugueux *ΔMAB_4690* faiblement virulent chez la souris et pour lequel un glycolipide pariétal disparait. L'analyse structurale approfondie de ce composé, couplée à la description bio-informatique du locus auquel appartient le gène *MAB_4690*, confirme la présence d'un nouveau GPL, nommé GPL4690, chez *Mabs* R. Comme les GPLs de *Mabs* S, cette entité moléculaire se décompose en trois parties : un noyau octapeptidique dont la séquence est TAFIYLTT ; un lipide acylé du côté N-terminal ; un motif 2-*O*-Me-α-rhamnopyranosyl(1-3)2-*O*-hydroxylauryl-α-rhamnopyranose glycosylé sur le résidu de thréonine côté C-terminal. Même si ce dernier semble produit en petite quantité, il laisse entrevoir des perspectives concernant l'explication de la virulence de la souche rugueuse, ainsi que des possibilités d'identification plus spécifique de *Mabs*.

ABSTRACT

Mycobacterium abscessus subs abscessus (Mabs) is an opportunist mycobacterium wich belong to the MABSC complex with *subs bolletti* (*Mbol*) and *subs massiliense* (*Mmas*). Inside this complex, epidemiological studies show that three dominant circulating clones appear worldwide among cystic fibrosis population as well as in heathy people. Indeed, *Mabs* is mainly responsible for pulmonary disease but it can also infect skin and soft tissues as well as nervous system. In addition, *Mabs* is the most drug-resistant mycobacterium and can be described in two morphotypes, smooth (S) and rough (R), which are separated according to their virulence and the glycolipids content of their cell-wall. The highly producing GPLs *Mabs* S has lower virulence compare to *Mabs* R which doesn't synthetise GPLs in its cell wall and appear later during the course of infection.

In order to better understand the pathogenicity of *Mabs*, our aim was to describe all free glycolipids inside the cell wall their interaction with innate immune system. Extraction and chromatography purification followed by mass spectrometry and nuclear magnetic resonance analysis highlights three glycolipid's family in both morphotypes. First, di-*O*-acyltrehalose structural studies show possible presence of α and α' mycolic acids similar to those from trehalose dimycolate and short polymethyl branched chain lipid. In addition, trehalose's acylation is observed either at position 2 and 3 or at position 3 and 4. On the other hand, study of the phosphatidyl myo-inositol dimannosides (PIM_2) that give rise to lypoglycan's biosynthesis, reveal specific 2-*O*-methyl mannose residue.

Furthermore, we describe in our study preferred pathways dedicated to GPL's biosynthesis thanks to structural comparison beetween wild-type *Mabs* S GPL-2a or GPL-3 as well as new accumulating GPLs in several mutants. the first gene *fmt* is involved in lipid β methoxylation, cell-wall hydrophobicity and macrophage invasion. Among the three glycosyltransferases coding genes, *gtf2* is required for glycosylation initiation while *gtf3* leads to the production of triglycosylated GPL-3 that is highly phagocytosized *in-vitro*. Further work on a third set of genes encoding acetyltransferases, including *atf1* and *atf2*, demonstrates not only the impact of *MAB_1725c* and *MAB_3448c*, located outside the GPL biosynthesis locus, on their acetylation, but also the prominent role of *atf2* in the regulation GPL-2a/GPL-3 ratio.

Finally, our work focused on a rough *ΔMAB_4690* mutant weakly virulent in mice. An unidentified compound disappears from its total polar lipid extract. Extensive structural analysis of this purified compound, coupled with bioinformatic description of the locus to which the *MAB_4690* gene belongs, confirms the presence of a novel GPL, named GPL4690. Like the GPLs of *Mabs* S, this molecular entity is split into three parts: an octapeptide core which sequence is TAFIYLTT ; an acylated lipid at the N-terminal side ; and a 2-O-Me-α-rhamnopyranosyl(1-3)2-O-hydroxylauryl-α-rhamnopyranose motif glycosylated on the threonine residue at the C-terminal side. Although GPL4690 appears to be produced in few amount, it offers insights into the explanation of the virulence of wild type *Mabs* R, as well as possibilities for more specific identification of this subspecies.