UNIVERSITÉ DE LILLE – FACULTÉ DES SCIENCES ET TECHNOLOGIES

École Doctorale Biologie – Santé de Lille (ED 446)

Thèse de Doctorat présentée par

Lin SHEN

pour l'obtention du grade de Docteur de l'Université de Lille Discipline : Biochimie et Biologie moléculaire

Caractérisation d'une galactosidase impliquée dans le

catabolisme de la paroi des mycobactéries

Réalisée au sein de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF, UMR CNRS 8576)

Soutenue publiquement le 13 Décembre 2019 devant le jury composé de :

Rapporteurs :	Dr. Anne LEMASSU Maîtresse de Conférences, IPBS, UMR5089, Toulouse
	Pr. Françoise ROUTIER Professeur des Universités, Faculté de médecine de Hanovre, Allemagne
Examinateurs :	Dr. Alain BAULARD Directeur de recherche INSERM, Institut Pasteur, U1019-UMR 8204, Lille
	Dr. Christian CHALUT Chargé de recherche CNRS, IPBS, UMR5089, Toulouse
Directeur de thèse :	Dr. Yann GUÉRARDEL Directeur de recherche CNRS, UGSF UMR8576, Villeneuve d'Ascq
Co-encadrant :	Dr. Christophe MARILLER Maître de Conférences, UGSF UMR8576, Villeneuve d'Ascq

Ces travaux ont été réalisés au sein du laboratoire : Unité de Glycobiologie Structurale et Fonctionnelle (UGSF, UMR CNRS 8576) Bâtiment C9, Avenue Mendeleïev Cité Scientifique, 59655 Villeneuve d'Ascq France

Remerciement

Avant tout, je tiens à remercier le **Pr. Christophe D'Hulst**, directeur de l'unité, pour m'avoir accueillie au sein de l'UGSF.

Je remercie ensuite l'ensemble des membres du jury d'avoir accepté d'évaluer ces travaux de thèse. Je tiens à remercier le **Dr. Anne Lemassu** et le **Pr. Françoise Routier** de me faire l'honneur d'évaluer ces travaux en tant que rapporteurs. Je remercie également le **Dr. Alain Baulard** et le **Dr. Christian Chalut** d'avoir accepté d'être examinateurs.

J'adresse aussi mes remerciements au **Dr. William Helbert** et au **Dr. Romain Veyron-Churlet** pour avoir participé à mes différents comités de suivi individuel et pour leurs précieux conseils.

Je tiens à remercier très chaleureusement mon directeur de thèse, le **Dr. Yann Guérardel**, de m'avoir encadrée et de m'avoir transmis ses connaissances. Je te remercie également de m'avoir donné énormément d'opportunités pour participer à de nombreux événements et d'avoir fait de mes trois ans de thèse une expérience autant enrichissante. Je voudrais te remercier non seulement en tant que directeur, mais aussi en tant que personne qui m'a encouragée et m'a toujours soutenue dans les moments difficiles.

Je tiens à exprimer ma plus profonde reconnaissance au **Dr. Christophe Mariller**, mon encadrant de thèse. C'est grâce à toi que j'ai commencé à m'intéresser à la recherche. Je me souviens encore de la première fois où je t'avais demandé si tu prenais des stagiaires en troisième année de licence. Depuis, je n'ai plus quitté ce laboratoire. Pendant ces six ans, j'ai pu développer mes compétences grâce à tes enseignements, tes conseils pertinents, tes explications détaillées mais aussi grâce à ta patience, pour avoir répété je ne sais combien de fois quand je ne comprenais pas à cause de la langue française. Tu m'as toujours fait confiance même quand je ne me fais moi-même pas confiance. Tu es toujours là pour m'aider, me soutenir sans dire un mot de plus. C'est grâce à toi que je n'ai plus peur nombreuse de choses et que je commence à avoir confiance en moi. Un énorme merci d'avoir autant contribué à ma réussite et merci de ton soutien!

J'adresse aussi mes remerciements au **Pr. Kaoru Takegawa** pour m'avoir m'accueillie avec bienveillance pendant deux mois dans son laboratoire au Japon. Cette expérience précieuse m'a permis d'ouvrir les yeux et de découvrir une autre culture de recherche très riche.

Je souhaite également remercier toutes les personnes qui ont participé à ce projet :

- Le Dr. Albertus Viljoen, le Dr. Iman Halloum et le Dr. Laurent Kremer pour leur collaboration, leur expertise dans le domaine des mycobactéries dans ce projet et leur conseils pertinents.
- Le **Dr. Sydney Villaume**, **Loïc Chêne**, le **Pr. Stéphane P. Vincent** pour nous avoir fourni de nombreux produits chimiques précieux et leurs conseils pertinents.
- Le **Dr. Maju Joe** et le **Pr. Todd L. Lowary** pour les produits synthétiques d'oligosaccharides et leurs précieux conseils.
- Les membres de la plateforme PAGés et plus particulièrement le Dr. Emmanuel Maes et le Dr. Bernadette Codeville pour les analyses en GC/MS. Le Dr. Frédéric Krzewinski et le Dr. Nao Yamakawa pour les nombreux coups de main.
- Le Dr. Shin-Yi Yu pour m'avoir formée à l'analyse des sucres avec gentillesse et pour m'avoir encouragée pendant ces trois ans. Mais aussi pour tous les repas chinois (ou autre) ensemble, ce qui m'a permis de me ressourcer pendant les moments difficiles.
- Le Dr. Olivier Vidal de m'avoir ouvert les portes du laboratoire de microbiologie, de m'avoir aidée pour la partie biologie moléculaire et microbiologie. Merci pour tes conseils, ta disponibilité et ta bonne humeur.
- Le **Dr. Geoffroy de Bettignies** de m'avoir appris la partie mutagenèse dirigée.
- Le Dr. Isabelle Huvent de m'avoir appris pour la mutagenèse dirigée, le clonage et tous les conseils précieux sur la biologie moléculaire.
- Les membres de l'équipe « Glycobiologie Structurale des Interactions Hôtes-Pathogènes » avec qui j'ai passé six ans merveilleux. Merci à : Christophe, Cédric, Elisabeth, Faustine, Florence, Frédéric, Charlotte et Sabrine.

J'adresse également des remerciements aux anciens doctorants de l'équipe : **Clément, Pierre-André, Alexandre, James** et **Thomas**, merci pour la bonne humeur et la bonne ambiance.

Je tiens à remercier plus particulièrement ma collègue/copine **Clémence**. Merci pour d'être toujours là pour moi, d'écouter mes plaintes de tout et de rien et d'avoir toujours corrigé mes fautes d'orthographe pendant ces cinq dernières années. Je n'y serais pas arrivée sans ton soutien. Et merci à tes parents aussi pour leur accueil chaleureux.

Merci à tous les collègues et compagnons qui sont ou ont été présents dans l'UGSF : Violaine, Camille, Angélina, Pascal, Vincent, Louis-David, Audrey, Anne-Sophie, Sumeyye, Maxance, Elodie, Alexandre, Charles, Ninon, Moyira et tous les autres.

Merci à toute l'équipe administrative et au personnel de l'unité pour leur aide et leur travail au quotidien.

Je remercie **mes amis** « non-scientifique » pour les bons moments passés ensemble, de me soutenir et de me changer les idées.

Enfin, je souhaiterais exprimer un énorme merci à **mes parents** qui ont toujours cru en moi, de m'avoir fait confiance, de m'avoir soutenu physiquement et moralement. Surtout ma maman, de répondre à mes appels à n'importe quelle heure, d'essayer de me consoler, de m'avoir dit « Tu vas sûrement y arriver, mais si jamais tu n'y arrives pas, ce n'est pas grave ! ». 最后,我想感谢我的父母,多谢你们一路的体谅与支持,在每次我想要放弃的时候,都不厌其烦地对我进行开导和鼓励。多亏你们的这份理解,才能有这份论文的终稿!

Table de matières

Remerciement	5
Table de matières	9
Liste des figures	14
Abréviations	19
Avant-propos	25
Partie I : Généralités	27
I. Épidémiologie de la tuberculose et du complexe Mycobacterium t	uberculosis 29
1. La tuberculose : état des lieux	29
2. Agent infectieux	
2.1 Actinobactéries	
2.1.1 Taxonomie des actinobactéries	
2.1.2 Classification	
2.1.3 Genre Mycobacterium	
2.1.3.1 Le complexe Mycobacterium tuberculosis	
2.1.3.2 Les mycobactéries non tuberculeuses	
3. Cycle infectieux de la tuberculose	41
4. Réponses immunitaires	44
5. Diagnostic et traitement de la tuberculose	46
5.1 Diagnostic de l'infection tuberculeuse latente	46
5.1.1 Test TCT	46
5.1.2 Le dosage IGRA	47
5.2 Diagnostic de la tuberculose active	47
5.2.1 La radiographie du thorax	47
5.2.2 La microscopie des frottis d'expectoration	47
5.2.3 Les tests moléculaires	
5.2.4 Examen de culture	
5.2.5 Coloration de Ziehl-Nelsen	
5.3 Diagnostic de la résistance aux médicaments antituberculeux	
5.4 Traitements antibiotiques	
5.4.1 Isoniazide	50

5.4.2 Rifampicine	50
5.4.3 Pyrazinamide	51
5.4.4 Ethambutol	51
5.4.5 Streptomycine	52
6. La tuberculose multirésistante	53
7. Conclusion du chapitre	54
II. Paroi cellulaire de Mycobacterium tuberculosis	56
1. Structure de la paroi cellulaire de Mycobacterium tuberculosis	61
1.1 Le peptidoglycane	61
1.2 L'arabinogalactane	62
1.3 Les acides mycoliques	65
2. Biosynthèses du complexe mAGP	67
2.1 Le peptidoglycane	67
2.1.1 Étapes cytoplasmiques	67
2.1.2 Étapes associées à la membrane	68
2.1.3 Assemblage du peptidoglycane dans le périplasme	70
2.2 L'arabinogalactane	70
2.2.1 L'unité <i>Linker</i>	70
2.2.2 La chaîne galactane	71
2.2.3 La chaîne arabinane	72
2.2.4 Modifications de l'arabinogalactane	75
2.3 L'acide mycolique	76
III. Les glycosides Hydrolases	
1. Vue globale des glycosidases	80
2. Mécanisme d'action des glycosidases	80
2.1 Mécanisme d'inversion	81
2.2 Mécanisme de rétention	81
2.3 Nécessité d'un ion métallique	83
3. Élucidations du mécanisme de la glycosidase	84
3.1 Stéréochimie des glycosidases	84
3.2 Identification des résidus catalytiques	85
3.2.1 Identification des résidus nucléophiles	86
3.2.2 Identification des résidus acides / basiques	88
4. Classifications des glycosidases	
5. Structure des glycosidases	91
5.1 Tonneau ($eta/lpha)_8$	91
5.2 β-Jelly roll	92

5.3 β -propeller	93
5.4 Tonneau ($lpha/lpha)_6$	93
5.5 Hélice-β	93
5.6 α+β	94
6. Topologie des sites actifs	94
6.1 Poche	94
6.2 Crevasse ouverte ou sillon	94
6.3 Niche	94
6.4 Tunnel	95
7. Les glycosidases de Mycobacterium tuberculosis	95
7.1 Les α-glucanes	97
7.2 Les β-glycanases	98
7.3 Enzymes impliquées dans le remodelage du peptidoglycane	99
7.4 Les α-mannosidases	99
8. Conclusion du chapitre	
IV. Remodelage de la paroi	101
1. Dégradation, remodelage et recyclage du peptidoglycane	102
2. Les Acides mycoliques	
3. Catabolisme de l'arabinogalactane	
4. Conclusion du chapitre	
Partie II : Obiectifs	
Partie III : Résultats et discussion	
I. Découverte de la protéine GlfH1	
1. Mise en évidence d'une nouvelle galactosidase chez <i>M. tuberculosis</i>	
1.1 Identification de GlfH1	
1.2 Production de GlfH1	
1.2.1 Expression de GlfH1 chez <i>E. coli</i>	
1.2.2 Extraction des protéines	
1.3 Isolement de GlfH1	
1.4 Empreinte peptidique massique (PMF)	
1.4.1 GIfH1	
1.4.2 Protéine associée	131
II. Caractérisation de GlfH1	
1. Paramètres enzymatiques de GlfH1	
1.1 Validation du substrat synthétique : le pNP-Galf	

1.2 Calcul de la vitesse enzymatique	135
1.3 Détermination des conditions de vitesse initiale	135
1.4 Détermination du pH optimal	136
1.5 Stabilité	137
1.6 Détermination des constantes michaeliennes : K_M et k_{cat}	138
1.7 Rôle du calcium	139
1.8 Spécificité de GlfH1 vis-à-vis du monosaccharide libéré	141
1.9 Inhibition par l'éthambutol	142
1.10 Trois formes différentes de GlfH1	144
2. Étude du site actif de GlfH1	146
2.1 Stéréochilmie du mécanisme catalytique : rétention ou inversion ?	146
2.2 Identification des résidus catalytiques	150
2.3 Détermination du site catalytique de GlfH1	151
2.3.1 Identification du site catalytique par alignement multiple de séquences	151
2.3.2 Modélisation par homologie de GlfH1	152
2.3.3 Mutagenèse dirigée	155
III. Activité de GlfH1	159
1. Activité de GlfH1 vis-à-vis de l'arabinogalactane synthétique	159
1.1 Analyse HPAEC	
1.2 Confirmation par MALDI-TOF	164
2. Activité de GlfH1 sur son substrat naturel, l'arabinogalactane	
2.1 Analyse en HPAEC	
2.2 Confirmations des monosaccharide libéré par GC/MS	
2.3 Analyse par RMN	
3. Hydrolyse sélective des liaisons eta 1-5 et eta 1-6 entre les résidus de Gal f	173
IV. Étudier la fonction de GlfH1	177
1. Identification de du gène orthologue de glfH1 chez M. smegmatis	
2. Validation du g èneorthologue de glfH1 chez M. smegmatis : MSMEG_5877	
3. Phénotype chez l'amibe	
4. Activité de MSMEG_5877 vis-à-vis de l'arabinogalactane de M. smegmatis	180
Partie IV : Conclusion	184
Partie V : Matériels et Méthodes	190
I. Considérations générales	192
II. Production, purification et analyse de la protéine recombinante obtenue	192

1. Souches bactériennes, plasmides et milieux de culture	192
2. Purification des plasmides	193
3. Expression de la protéine recombinante	194
4. Extraction des protéines	195
5. Purification de GlfH1 par chromatographie d'affinité sur métal immobilisé (IMAC)	196
6. Dessalage de GlfH1 par chromatographie de gel filtration	196
7. Dosage de protéine par la méthode BCA	197
8. Electrophorèse en gel de polyacrylamide en présence de dodécyl sulfate de sodium (SDS-PAGE)	197
9. Clonage de GlfH1	197
10. Mutagenèse dirigée	199
11. Empreinte peptidique massique (<i>Peptide Mass Fingerprinting</i> , PMF)	202
III. Analyse bioinformatique	203
IV. Caractérisation de l'activité enzymatique de GlfH1	204
V. Modélisation par homologie	204
VI. Préparation de l'arabinogalactane de Mycobacterium bovis BCG	205
1. Culture de <i>M. bovis BCG</i>	205
2. Purification de l'arabinogalactane de Mycobacterium bovis BCG	205
2.1 Préparation du complexe mAGP	205
2.2 Purification de l'AG	206
VII. Conditions hydrolyse par GlfH1	206
1. Mode d'action de GlfH1	206
2. Hydrolyse d'oligosaccharides synthétiques par GlfH1	207
3. Hydrolyse des polysaccharides naturels par GlfH1	207
VIII. Méthodes analytiques	207
1. Analyse HPAEC-PAD	207
2. Analyse par chromatographie en phase gazeuse (CPG)	208
2.1 Dérivation par triméthylsilylation (TMS)	208
2.2 Analyses en GC-FID et GC/MS	208
3. Analyse spectromètre de masse MALDI-TOF	209
4. Préparations des échantillons pour la Résonance magnétique nucléaire (RMN)	209
Annexes	. 210
Bibliographie	. 214

Liste des figures

Figure 1 : Taux d'incidence de la tuberculose	30
Figure 2 : Taxonomie du phylum Actionobacteria	32
Figure 3 : Arbre phylogénétique du phylum Actinobacteria	34
Figure 4 : Représentation de la structure de l'enveloppe du groupe CMNR	35
Figure 5 : Arbre phylogénétique du genre <i>Mycobacterium</i>	36
Figure 6: Représentation schématique de la relation évolutive du complexe	de
M. tuberculosis	38
Figure 7 : Classification des mycobactéries non tuberculeuses	40
Figure 8 : Progression de la tuberculose, de l'infection à la tuberculose active	41
Figure 9 : Cycle infectieux de la tuberculose	42
Figure 10 : Réponses immunitaires protectrices pendant l'infection de M. tuberculosis	44
Figure 11 : Organisation de l'enveloppe cellulaire des mycobactéries	56
Figure 12 : Représentation schématique de l'enveloppe cellulaire de <i>M. tuberculosis</i>	58
Figure 13 : Représentation schématique du complexe mAGP	60
Figure 14 : Structure du mycolyl-arabinogalactane (mAG)	63
Figure 15 : Représentation structurale des acides mycoliques de M. tuberculosis avec ces tro	ois
types principaux	66
Figure 16 : Schéma de la biosynthèse, la dégradation et le recyclage du peptidoglycane d	les
espèces Mycobacterium avec les différentes enzymes impliquées	69
Figure 17 : Biosynthèse de l'arabinogalactane	74
Figure 18: Représentation schématique de biosynthèse des acides mycoliques	de
M. tuberculosis	77
Figure 19: Mécanisme typique des glycosidases agissant par inversion de la configuration	on
anomériques du substrat	81
Figure 20 : Mécanisme typique des glycosidases agissant par rétention de la configuration	on
anomériques du substrat	83
Figure 21 : Mécanisme de réaction proposé pour une α -mannosidase de la famille GH92 a	git
par inversion	84
Figure 22 : Récupération de l'activité catalytique d'une glycosidase par l'azoture	86

Figure 23 : Inactivation et réactivation d'une glycosidase agit en β -rétention par le 2-déoxy-2-
fluoro glycosides
Figure 24 : L'inactivation et la réactivation d'une glucosidase agit en α -rétention par le
fluorure de 5-fluoro- α -D-glucopyranosyl
Figure 25 : Structures des glycosidases
Figure 26 : Topologies des sites actifs des glycosidases
Figure 27 : Voies de biosynthèses et de dégradation du tréhalose chez <i>M. tuberculosis</i> 98
Figure 28 : Le remodelage de la paroi de <i>M. tuberculosis</i>
Figure 29: Représentation schématique de la chaîne arabinane de l'arabinogalactane en
présentant les différents sites de coupure 106
Figure 30 : Le motif 5- <i>O</i> -mycolyl-β-Ara <i>f</i> -(1→2)-5- <i>O</i> -mycolyl-α-Ara <i>f</i> est retrouvé à l'extrémité
terminale non réductrice du mAGP et du DMAG107
Figure 31 : Structure du L-Arabinofuranose et du D-Galactofuranose
Figure 32 : Chromatogramme représentant l'hydrolyse de l'arabinogalactane de M. bovis BCG
par GlfH1 en HPAEC 118
Figure 33 : L'annotation de la séquence GlfH1 avec la présence d'un peptide signal 118
Figure 34: Alignement multiple des séquences de GlfH1(NP_217612.1) avec d'autres
glycosidases mycobactériennes de la famille GH5_13119
Figure 35 : Analyse phylogénétique des glycosidases mycobactériennes de la sous-famille
GH5_13 en utilisant la méthode PHYLIP Neighbor Joining 120
Figure 36: Analyse phylogénique de glycosidases mycobactériennes de la sous-famille
GH5_13 à l'aide de la méthode PhyML Maximum Likehood120
Figure 37 : Représentation schématique à l'aide du logiciel pDRAW32 de la séquence codante
du vecteur pET32a-GlfH1 122
Figure 38 : Analyse par SDS-PAGE des extraits cellulaires protéiques de souches d'E. coli BL21
pET32a-Rv3096 cultivées en milieu LB 123
Figure 39 : Analyse par SDS-PAGE des fractions solubles et insolubles de souches d'E. coli BL21-
pET32a- GlfH1 cultivées en milieu audoinducteur (AIM) à 25 °C, obtenu par sonication, avec
ou sans la présence de détergents125
The second s
Figure 40 : Activite totale (AT) et activite specifique (AS) de GITH1 exprimee dans E. coll en
fonction de la méthode de lyse et du détergent employé

Figure 42 : Analyse par SDS-PAGE des extraits protéiques de souches d'E. coli BL21 pET32a-
Rv3096 cultivées en milieu autoinducteur (AIM) à 25 °C après purification IMAC 130
Figure 43 : Spectre de masse MALDI-TOF des peptides trypsiques de GlfH1 132
Figure 44 : Spectre de masse MALDI-TOF des peptides trypsiques de la protéine inconnue
associée à GlfH1 lors de son isolement 132
Figure 45 : Représentation graphique de l'activité enzymatique en µmole/min en fonction de
différentes concentrations de GlfH1 134
Figure 46: Représentation graphique de l'absorbance à 405 nm en fonction de la
concentration en pNp 135
Figure 47 : Détermination des conditions expérimentales de vitesse initiale
Figure 48 : Représentation graphique de l'activité enzymatique en fonction du pH 137
Figure 49 : Représentation graphique de la stabilité de l'activité enzymatique de GlfH1 en
fonction du pH 138
Figure 50: Représentation graphique de l'activité enzymatique en fonction de la
concentration en <i>p</i> NP-Gal <i>f</i>
Figure 51 : Pourcentage d'activité résiduelle de GlfH1 en présence de différents chélateurs ou
cations divalents
Figure 52 : Activité de GlfH1 vis-à-vis de douze substrats synthétiques
Figure 53 : Inhibition de GlfH1 par l'éthambutol 143
Figure 54 : Modèle d'inhibition en mode mixte et non compétitif 143
Figure 55 : Activité enzymatique résiduellev de GlfH1 en présence de méthanol 146
Figure 56 : Analyse par CPG après triméthylsylilation 147
Figure 57 : Analyse GC / MS du produit de la réaction de GlfH1 149
Figure 58 : Détermination la nature des résidus catalytiques
Figure 59 : Alignement de séquences multiples de GlfH1 avec d'autres glycosidases de la
famille GH5_13 151
Figure 60 : Superposition des quatre modèles 3-D de GlfH1 152
Figure 61 : Le modèle par homologie de GlfH1 s'inscrit parfaitement dans la mesure de volume
obtenue par Dynamic Light Scattering (DLS) 153
Figure 62 : Modèle par homologie de GlfH1 obtenu avec i-Tasser 154
Figure 63: Stratégie du 2-fluoro-pNP-glycoside pour l'identification du résidu catalytique
réalisant 157

Figure 64 : Structure des oligosaccharides synthétiques imitant la structure des chaînes
galactanes de l'arabinogalactane 160
Figure 65 : Chromatogramme d'hydrolyse de l'oligosaccharide <i>Lin</i> -Gal <i>f</i> ₅ par GlfH1 161
Figure 66 : Chromatogramme d'hydrolyse de l'oligosaccharide <i>Lin</i> -Gal <i>f</i> ₆ par GlfH1 162
Figure 67 : Chromatogramme d'hydrolyse de l'oligosaccharide Lin-Galf ₇ par GlfH1 162
Figure 68 : Chromatogramme d'hydrolyse de l'oligosaccharide <i>Lin</i> -Gal <i>f</i> ₁₁ par GlfH1 163
Figure 69 : Chromatogramme d'hydrolyse de l'oligosaccharide Lin-Galf ₁₂ par GlfH1 163
Figure 70 : Spectrométrie de masse MALDI-TOF après hydrolyse de l'oligosaccharide Lin-Gal ₁₂
par GlfH1
Figure 71: Représentation schématique de l'action de GlfH1 sur l'arabinogalactane de
M. tuberculosis
Figure 72 : Chromatogramme représentant l'hydrolyse de l'arabinogalactane, AGP et mAGP
par GlfH1 167
Figure 73 : Analyse GC/MS sur le produit réactionnel de l'arabinogalactane hydrolysé par
GIfH1
Figure 74 : Analyse GC/MS sur le produit réactionnel de l'AGP hydrolysé par GlfH1 169
Figure 75 : Analyse GC/MS sur le produit réactionnel du mAGP hydrolysé par GlfH1 170
Figure 76 : Analyse RMN de l'activité de GlfH1 171
Figure 77 : Chromatogramme représentant l'hydrolyse des substrats synthétiques 2PA-Gal $f_{\beta_{1-}}$
$_5$ Galf et 2PA-Gal $f_{\beta_{1-6}}$ Galf par 0 à 30 µg de GlfH1173
Figure 78: Représentations graphiques du pourcentage d'hydrolyse des substrats
synthétiques par GlfH1 en fonction de la quantité de GlfH1 174
Figure 79 : Chromatogramme représentant l'hydrolyse des substrats synthétiques 2PA-Gal $f_{\beta_{1-}}$
$_5$ Gal f et 2PA-Gal $f_{\beta_{1-6}}$ Gal f par 3 µg de GlfH1 175
Figure 80 : Représentations graphiques de la cinétique d'hydrolyse des substrats synthétiques
par GlfH1 175
Figure 81: Alignement des séquences de GlfH1 avec MSMEG_5877 par l'algorithme
Needleman & Wunsh 177
Figure 82 : La croissance des différents mutants dans l'amibe 179
Figure 83 : Impact de la délétion de GlfH1 sur la structure de l'arabinogalactane 180
Figure 84 : Spectres RMN 1D ¹ H (900 MHz) obtenus sur les arabinogalactanes purifiés des
souches WT, ΔglfH1 et ΔglfH1 complémentation de M. smegmatis

Figure 85 : Quantifications des monosaccharides individuels sur l'arabinogalacta	ne purifiée de
M. smegmatis	
Figure 86 : Représentation schématique du vecteur pET32a (Invitrogen)	195

Abréviations

2F-DGF	2,4 dinitrophényl 2-deoxy-2-fluoro-β-D-galactofuranoside
2PA	2-aminopyridine
ACN	Acétonitrile
ADN	Acide DésoxyriboNucléique
AG	ArabinoGalactane
AGP	ArabinoGalactanePeptidoglycane
AIM	Auto Induction Media
Ara <i>f</i>	D-arabinofuranose
Ara <i>f</i> T	Arabinofuranosyltransférase
AraT	Arabinosyl-transférase
ARN	Acide RiboNucléique
AS	Activité Spécifique
АТ	Activité Totale
ATFA	Acide TriFluoroAcétique
ATFA	Acide TriFluoroAcétique
BAAR	Bacille Acido-Alcoolo-Résistant
BCA	BiCinchoninic Acid
BCG	Bacille de Calmette et Guérin
Bic	Bicarbonate d'ammonium
ВК	Bacille de Koch
BSTFA	Bis Silyle Trifluoroacétamide
BTZ	Benzothiazinone
CAZy	Carbohydrate-Active enZyme
CD	Cellule dendritique
Chaps	3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate
СНСА	α-Cyano-4-hydroxycinnamic acid
СМН	Complexe Majeur d'Histocompatibilité

CMI	Concentration minimale d'inhibition	
CMNR	Corynebacterium, Mycobacterium, Nocardia et Rhodococcus	
СМТ	Compexe Mycobacterium Tuberculosis	
СоА	Coenzyme A	
COSY	COrrelation SpectroscopY	
CPG	Chromatographie en Phase Gazeuse	
cv	Column Volume	
DAG	Diacylglycérol	
DAT	Tréhalose Diacylé	
DDM	n-Dodecyl β-D-maltoside	
DHB	DiHydroxyBenzoïque	
DMAG	5-O-mycolyl-β-Araf-(1->2)-5-O-mycolyl-α-Araf-(1->1)-Gro	
DMSO	Diméthylsulfoxyde	
DP	Degré de polymérisation	
DPA	DecaprenylmonoPhosphoryl-D-Arabinose	
DPPR	DecaprenylPhosphoryl-5-PhosphoRibose	
DPR	Décaprényl phosphoribose	
DTT	Dithiothréitol	
EDTA	EthylèneDiamineTétraAcétique	
EGTA	EthylèneGlycérol-bis(2-aminoethylether)-N,N,N',N'-TétraAcétique	
ELISA	Enzyme-Linked ImmunoSorbent Assay	
ЕМВ	Éthambutol	
EtOH	Éthanol	
FAS	Fatty Acid Synthase	
Gal <i>f</i>	D-galactofuranose	
GalNAc	N-AcetylGalactosamine	
Galp	Galactopyranose	
GalT	Galactosyltransférase	
GC-FID	Chromatographie en phase gazeuse couplée à un détecteur à ionisation de flamme	
GC/MS	Gas Chromatography Mass Spectrometry	

GH	Glycoside Hydrolase	
GlcNAc	N-AcetylGlucosamine	
Glcp	Glucopyranose	
GLP	Glycolipide phénolique	
GPL	Glycopeptidolipide	
HPAEC-PAD	High Performance Anion Exchange Chromatography-Pulsed Amperometric Detection	
hpi	heures après l'infection	
HPLC	Chromatographie en phase liquide à haute pression	
HSQC	Heteronuclear Single Quantum Correlation	
IDR	Interadermo-réaction à la tuberculine	
IFNγ	Interféron-y	
IGRA	In-vitro Interferon Gamma Release Assay	
IMAC	Chromatographie d'Affinité pour ions Métalliques Immobilisés	
INH	Isoniazide	
IPTG	Isopropyl β-D-1-thiogalactopyranoside	
ITL	Infection Tuberculeuse Latente	
IUBMB-CE	International Union of Biochemistry and Molecular Biology Enzyme Commission	
LAM	LipoArabinoMannane	
LB	Luria-Bertani	
LDAO	N,N-Dimethyldodecylamine N-oxide	
Ldt	L,D-transpeptidases	
Lin	linker	
LM	LipoMannane	
<i>m</i> -DAP	meso-DiAminoPimélique	
mAG	mycolyl-ArabinoGalactane	
mAGP	mycolyl-ArabinoGalactane-Peptidoglycane	
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation -Time Of Flight	
Man <i>p</i>	Mannopyranose	
MCS	Multiple Cloning Site	
MDR	Multiple Drug Resistance	

ME	Membrane Externe	
MeOH	Méthanol	
MNT	Mycobactéries Non Tuberculeuses	
MP	Membrane Plasmique	
МТВ	Mycobacterium tuberculosis	
MurNAc	Acide N-acétylmuramique	
MurNGly	Acide N-glycolylmuramique	
NB-DNJ	N-butyl-deoxynojirimycine	
OBG	Octyl-beta-Glucoside	
OMS	Organisation Mondiale de la santé	
ΡΑΤ	Tréhaloe Pentacylé	
PBS	Phosphate Buffered Saline	
PCR	Polymerase Chaine Reaction	
PDIM	Phthiocérol DIMycocérosate	
PEP	Phospho-énol-pyruvate	
PG	PeptidoGlycane	
PIM	Phosphatidyl- <i>myo</i> -Inositol	
PMF	Peptide Mass Fingerprinting	
pNPGal <i>f</i>	p-Nitrophenol-beta-D-Galactofuranose	
pNPGal <i>p</i>	p-Nitrophenol-α-Galactopyranose	
pRpp	5-phosphoribosyl-1-pyrophosphate	
PZA	Pyrazinamide	
qsp	quantité suffisante pour	
Rha	Rhamnose	
RIF	Rifampicine	
RMN	Résonance Magnétique Nucléaire	
Rpf	Resuscitation-promoting factor	
SAB	Sérum Albumine Bovine	
SDS	dodécylsulfate de sodium	
SDS-PAGE	Electrophorèse en gel de polyacrylamide en présence de Dodécyl Sulfate Sodium	

SL	Sulphatide
SM	Streptomycine
TAE	Tris-Acétate-EDTA
TAG	Triacylglycérol
ТВ	tuberculose
тст	Test Cutané à la Tuberculine
TDM	Tréhalose DiMGycolate
TFA	Acide trifluoroacétique
тмм	Tréhalose MonoMycolate
TMS	Triméthylsilyle
TNF-α	Facteur de nécrose tumorale $lpha$
ТРР	Thymidine TriPhosphate
Trx	Thiorédoxine
Tween 20	Polyethylene glycol sorbitan monolaurate
VIH	Virus d'Immunodéficience Humaine
WT	Wild Type
XDR	eXtensively Drug Resistance
Xylp	Xylopyranose

Avant-propos

La tuberculose (TB) est une maladie ancienne - des études montrent qu'elle affecte l'homme depuis plusieurs milliers d'années¹. La cause de la tuberculose est restée inconnue jusqu'en 1882, année où le Dr. Robert Koch a annoncé la découverte du bacille Mycobacterium tuberculosis. À la fin des années 1800, la tuberculose était l'une des principales causes de décès dans certains pays européens. À partir des années 1940, grâce à la découverte des premiers médicaments anti-tuberculeux, les taux de cas nationaux ont diminué jusqu'à 10 % par an et les taux de mortalité ont diminué encore plus rapidement. Dans certains pays européens, la tuberculose en est arrivée à être considérée comme une maladie du passé. Pour de nombreux pays, toutefois, la tuberculose reste toujours un problème majeur de santé publique. Et ceci en dépit du fait que, avec un diagnostic opportun et un traitement médicamenteux correct, la plupart des personnes qui développent la maladie peuvent être guéries. Mais malgré les progrès importants réalisés et le fait que les taux d'infection et de décès ont régulièrement baissé, il y a environ 10 millions de personnes dans le monde souffrant encore de la tuberculose chaque année et elle reste l'une des 10 principales causes de décès. Elle est la cause principale de décès par un agent infectieux unique, devant le VIH / sida.

Dans ce cadre de lutte contre la tuberculose, la paroi de *Mycobacterium tuberculosis* constitue une cible thérapeutique de première importance. Plusieurs molécules thérapeutiques ciblent déjà la biosynthèse des constituants de la paroi. Par exemple, l'isoniazide et l'éthionamide sont des inhibiteurs de la synthèse des acides mycoliques, l'éthambutol de la synthèse de l'arabinogalactane et, enfin, plusieurs molécules de la classe des carbapénèmes de la synthèse du peptidoglycane. Cependant, avec l'apparition de la tuberculose multirésistantes (MDR-TB) et la tuberculose ultrarésistante (XDR-TB), il est impératif de trouver de nouvelles cibles thérapeutiques pour éradiquer la tuberculose avant 2030 d'après les objectifs de l'OMS¹. Pour cela, les recherches fondamentales sur la paroi cellulaire des mycobactéries visant à découvrir de nouvelles cibles thérapeutiques potentielles sont nécessaires. Dans ce contexte, la mise en évidence d'un catabolisme de la paroi et notamment de son polysaccharide principal, l'arabinogalactane serait un premier pas important vers ce but. C'est pourquoi ce mémoire présente l'identification et la caractérisation de la première glycosidase dont la cible est la chaine galactane de l'arabinogalactane mycobactérien.

Ce manuscrit de thèse contient deux parties principales, la première partie est la présentation générale du sujet. Dans cette partie, les informations concernant la tuberculose tels que la mode d'infection, les diagnostics et les traitements seront d'abord décrits. Puis la structure et la biosynthèse des différents composés de la paroi cellulaire de *M. tuberculosis* vont abordés. Ensuite, les caractéristiques principales des glycosidases, y compris des glycosidases de *M. tuberculosis*, seront présentées. Enfin, le concept de catabolisme de la paroi mycobactérienne sera abordé.

La deuxième partie du manuscrit présentent les travaux réalisés pendant ces trois ans de thèse et relatant l'identification et la caractérisation de la galactofuranohydrolase GlfH1 de *M. tuberculosis.*

Partie I : Généralités

I. Épidémiologie de la tuberculose et du complexe *Mycobacterium tuberculosis*

1. La tuberculose : état des lieux

La tuberculose (TB) est une des maladies infectieuses les plus anciennement connues. Elle est l'une des 10 principales causes de décès et se classant en première position des maladies dues à un agent infectieux unique (au-dessus du VIH / sida). En 2017, l'Organisation Mondiale de la Santé (OMS) a recensé 1,3 million de décès dus à cette pathologie parmi les personnes non porteuses du VIH et 300 000 décès supplémentaires parmi les personnes avec VIH positifs¹. Dix millions de nouvelles infections, soit 133 cas pour 100 000 habitants ont été également estimés. La tuberculose est présente dans tous les pays et touche les personnes de tout âge. Les estimations pour 2017 ont montré que 90 % des cas sont des adultes, 64 % des hommes et 9 % des personnes infectées par le VIH (72 % en Afrique). Les deux tiers des patients se trouvent soit en Inde (27 %), en Chine (9 %), en Indonésie (8 %), aux Philippines (6 %), au Pakistan (5 %), au Nigéria (4 %), au Bangladesh (4 %) ou en Afrique du Sud (3 %). Selon les estimations, le nombre de personnes développant la tuberculose chaque année est en baisse, bien que cette diminution soit très lente. Environ 1,7 milliard de personnes, soit 23 % de la population mondiale, seraient infectées par l'infection tuberculeuse latente (ITL) et risqueraient donc de développer une tuberculose active au cours de leur vie. Entre 5 % et 15 % des personnes infectées évolueront (de quelques mois à quelques années) vers une tuberculose active².



Figure 1 : Taux d'incidence de la tuberculose en 2017 adapté de Global Tuberculosis Report 2018¹.

Entre 2000 et 2017, le taux de mortalité par tuberculose a chuté de 42 % grâce à l'utilisation de médicaments antituberculeux, ces derniers s'attaquant principalement à la biosynthèse de composés spécifiques présents dans la paroi bactérienne. Cependant, ces traitements perdent en efficacité en raison de l'apparition de souches résistantes. Selon le nombre de médicaments auxquels résiste le pathogène, il s'agit de tuberculose multirésistante (*Multiple Drug Resistance,* MDR-TB), résistance à au moins l'isoniazide et la rifampicine ou ultrarésistante (*eXtensively Drug Resistant,* XDR-TB) où résistance à pratiquement tous les antituberculeux) ³. On estime à 558 000 le nombre de nouveaux cas de tuberculose résistante à la rifampicine (RR-TB) dans le monde, dont près de la moitié dans trois pays : l'Inde (24 %), la Chine (13 %) et la Russie (10 %). Parmi les cas de patients atteints par une tuberculose résistante à la rifampicine, 82 % ont une tuberculose multirésistante (MDR-TB). Selon l'étude de l'OMS en 2017, 3,5 % des nouveaux cas de tuberculose et 18 % des cas précédemment traités sont atteints de RR/MDR-TB. 8,5 % des cas de MDR-TB sont en réalités des cas de XDR-TB.

C'est pour cela que régulièrement, des recherches fondamentales sont effectuées afin de comprendre le mécanisme d'action des agents anti-TB courants et de trouver de nouvelles cibles pharmacologiques permettant la synthèse de nouveaux médicaments.

2. Agent infectieux

L'agent infectieux de la tuberculose est *Mycobacterium tuberculosis*. Les mycobactéries appartiennent au genre *Mycobacterium*, à la famille des *Mycobacteriaceae*, au sous-ordre des *Corynebacterineae*, à l'ordre des *Actinomycetales* et à la classe des *Actinobacteria*.

2.1 Actinobactéries

Le phylum *Actinobacteria* représente l'un des plus grands groupes d'unité taxonomique parmi les 18 groupes de bactéries reconnus jusqu'au présent⁴. Ce sont des bactéries aérobies à Gram positif avec un taux de G+C très important : de 51 % chez certains *corynebacteria* et jusqu'à plus de 70 % chez les *Streptomyces* et *Frankia*⁵. Il existe toutefois deux exceptions : le *Tropheryma whipplei* et le *Gardnerella vaginalis* qui contiennent moins de 50 % de G+C dans leur gènome^{6,7}.

La majorité des actinobactéries se trouvent dans le sol ou en milieux aquatiques, comme les genres *Streptomyces, Micromonospora, Rhodococcus* et *Salinispora*. Cependant certaines d'entre elles sont des agents infectieux pour les plantes ou les animaux comme les genres *Corynebacterium, Mycobacterium* et *Nocardia*. Elles peuvent aussi être des symbiotes de plantes comme *Frankia*, ou des commensaux de gastro-intestinales comme *Bifidobacterium*⁶. Elles sont majoritairement chimiohétérotrophiques et utilisent une grande variété de substrats carbonés organiques, y compris des polysaccharides complexes^{6,8}.

Les actinobactéries ont un impact important sur la santé humaine en raison de leurs diverses propriétés physiologiques et de métaboliques. Grâce à leur métabolisme secondaire, elle produit deux-tiers des molécules antibiotiques naturelles utilisées en traitement clinique, de composés anticancéreux, d'anthelminthiques et d'antifongiques également^{6,9}.

Au niveau écologique, elles sont indispensables pour la décomposition des matières végétales et en particulier de la lignocellulose ce qui permet de produire des biocarburants⁵. Elles jouent aussi un rôle crucial dans le recyclage des biomatériaux réfractaires à leur décomposition et humification¹⁰.

2.1.1 Taxonomie des actinobactéries

Le phylum des *Actinobacteria* est défini en se basant sur l'arbre phylogénique de l'ARNr 16S et contient 6 classes : *Actinobacteria, Acidimicrobiia, Coriobacteriia, Nitriliruptoria, Rubrobacteria* et *Thermoleophilia*⁴. La classe *Actinobacteria* comprend au totale 15 ordres⁴(Fig 2). Cependant, l'utilisation des séquences nucléotidiques de l'ARNr 16S reste ambigüe car elle ne distingue pas bien les genres ni les espèces. C'est pour cela que des marqueurs génétiques comme *rpoB* et *ssgB* ont été utilisés pour la discrimination des genres proches¹¹. De plus, ses caractéristiques morphologiques, physiologiques et chimiotaxonomiques permettent de réaliser des classifications préliminaires¹².





2.1.2 Classification

Le phylum *Actinobacteria* au niveau du genre et d'espèce peut être basée sur la morphologie microscopique, les caractères chimiotaxonomiques ou les caractéristiques moléculaires comme la séquence de l'ARNr 16S.

La classification morphologique du phylum *Actinobacteria* au niveau du genre et d'espèce se distingue en fonction de la présence ou non d'un mycélium ou plus précisément, d'un

mycélium aérien, de la couleur de ce mycélium, de la structure et de la forme de ses spores, et des pigments mélanoïdes produits par la bactérie⁶.

La classification chimiotaxonomique se base sur la distribution de composants chimiques comme les acides aminés, les lipides, les protéines, les ménaquinones, les acides muramiques, les sucres et la composition de l'ADN.

La classification moléculaire se base sur le séquençage du gène de l'ARNr 16S et l'hybridation ADN-ADN. Grâce à cette classification, l'ordre des *Actinomycetales* a été ségrégé en 14 sousordres : *Actinopolysporineae, Catenulisporineae, Corynebacterineae, Frankineae, Glycomycineae, Jiangellineae, Kineosporineae, Micrococcineae, Micromonosporineae, Propionibacterineae, Pseudonocardineae, Streptomycineae,* et *Streptosporangineae*¹³. De plus, le séquençage des différents génomes du phylum *Actinobacteria* a permis l'identification de 43 familles et de 203 genres^{10,14}(Fig 3). Tous les groupes précédemment classés dans l'ordre ont été considérés comme monophylétiques sur la base de leurs critères moléculaires, mais certains groupes paraphylétiques, c'est-à-dire un groupe d'organismes comprenant un ancêtre mais pas tous ses descendants¹⁵, ont été trouvés dans le sous-ordre.



Figure 3 : Arbre phylogénétique basé sur 97 séquences génomiques du phylum Actinobacteria ⁶.

Les genres *Corynebacterium, Mycobacterium, Nocardia* et *Rhodococcus* forment un taxon monophylétique, c'est-à-dire un groupe d'organismes regroupant les espèces ancestrales uniques ainsi que la totalité de ses descendants, au sein des *Actinobacteria* qui s'appelle le groupe CMNR¹⁴. Ce groupe partage une enveloppe particulière (Fig 4) qui contient des acides mycoliques, le peptidoglycane et divers polysaccharides dont notamment l'arabinogalactane, le lipomannane et le lipoarabinomannane, et aussi les phosphatidyl-*myo*-

inositolmannosides¹⁴. Cependant, les compositions spécifiques pour chaque genre peuvent différer. Par exemple, la chaîne des acides mycoliques de *Corynebacterium* est constituée de 22 à 36 carbones tandis que pour les *Mycobacterium*, le nombre de carbones est entre 60 à 100. Les acides mycoliques de *Corynebacterium* ne portent pas des modifications telles que des cyclopropanes et des groupes oxygénés tandis que les acides mycoliques de *Mycobacterium* les portent¹⁶.



Figure 4 : Représentation schématique de la structure de l'enveloppe du groupe CMNR.

2.1.3 Genre Mycobacterium

Le genre *Mycobacterium* compte 188 espèces¹⁷ dont la plupart sont non-tuberculeuses alors que ce genre avait d'abord été créé pour classer les agents responsables de la tuberculose et de la lèpre (*Mycobacterium leprae*).



Figure 5 : Arbre phylogénétique du genre *Mycobacterium* avec les cinq nouveaux groupes identifiés et la nouvelle nomenclature proposées par RS. Gupta *et al.* en 2018. Les espèces dont leur génome qui a déjà été analysé dans l'étude des auteurs sont écrites en gras, figure adaptée de Gupta *et al.* 2018¹⁸.
Une recherche sur les analyses phylogénomiques complètes et les analyses génomiques comparatives sur 150 génomes d'espèces de *Mycobacterium* ont été réalisées par RS. Gupta *et al.* en 2018 pour comprendre leurs interrelations¹⁸. Les arbres phylogénétiques construits montrent clairement l'existence de cinq groupes monophylétiques distincts au sein du genre *Mycobacterium* qui sont *Tuberculosis-Simiae, Terrae, Triviale, Fortuitum-Vaccae* et *Abscessus-Chelonae*. Les nouvelles nomenclatures ont été proposées par les auteurs. Ils proposent la division du genre *Mycobacterium* en un genre modifié, le *Mycobacterium*, pour le clade *Tuberculosis-Simiae* qui regroupe l'ensemble des principaux agents pathogènes humains et en quatre nouveaux genres. *Mycolicibacterium* correspond au clade *Fortuitum-Vaccae*, *Mycolicibacter* correspond au clade *Terrae*, *Mycolicibacillus* correspond au clade *Triviale* et *Mycobacterides* correspond au clade *Abscessus-Chelonae* (Fig 5). Cette division des espèces de mycobactérie en cinq groupes distincts permet de se concentrer sur les caractéristiques génétiques et moléculaires uniques qui différencient les membres de ces groupes.

Les mycobactéries sont des bacilles immobiles, aérobies et à la forme de bâtonnets assez longs et fins. La taille moyenne est de 0,2 à 0,6 μ m de largeur et 1 à 10 μ m de longueur. Les températures optimales pour la croissance varient selon les espèces et vont de 25 °C à plus de 50 °C. Ce sont des bactéries classifiées comme étant à Gram positif mais avec une enveloppe plus proche de celles à Gram négatif. Elles sont acido-alcoolo-résistantes (B.A.A.R : Bacilles Acido-Alcoolo-Résistants), une caractéristique spéciale de la paroi qui la rend distinguable par la coloration de <u>Ziehl-Neelsen</u>. Cette coloration permet de discriminer les mycobactéries par microscopie photonique. Elles se démarquent aussi par leur génome qui présente 66 % de GC en moyenne et un temps de génération très long (jusque 15-20 h contre 20 min pour *Escherichia coli*).

Les mycobactéries peuvent être classées en trois groupes principaux au niveau médical. Le complexe *Mycobacterium tuberculosis* contient plusieurs espèces responsables de la tuberculose ; *Mycobacterium leprae* cause la maladie d'Hansen ou lèpre ; et les mycobactéries non tuberculeuses (MNT) comprennent toutes les autres mycobactéries qui ne provoquent pas la tuberculose ou la lèpre. Pour la suite, le complexe <u>Mycobacterium tuberculosis</u> et les <u>mycobactéries non tuberculeuses</u> vont être développés.

2.1.3.1 Le complexe Mycobacterium tuberculosis

Le complexe *Mycobacterium tuberculosis* (CMT) partage une parfaite identité de séquence de l'ARNr 16S et 99,9 % d'identité nucléotidique au niveau du génome complet¹⁹. Ce complexe comprend *M. tuberculosis* (Mtb), *M. africanum*, *M. bovis*, *M. microti*, *M. canettii*, *M. caprae*, *M. pinnipedii*, *M. suricattae*, *M. mungi*, *M. dassie*, et *M. orygis* (Fig 6). Parmi ces espèces, *M. tuberculosis* est le membre le plus connu.

C'est en 1882 que Robert Koch a identifié *M. tuberculosis* également appelée Bacille de Koch (BK), comme agent causal de la tuberculose. Au sein de l'espèce *M. tuberculosis*, il existe plusieurs souches présentant une diversité génétique provenant de mutations ponctuelles ou de délétions. Ces différentes souches de *M. tuberculosis*, comme par la souche Beijing, sont importantes dans le cadre de la tuberculose humaine car elles peuvent présenter une hypervirulence et des résistances multiples aux traitements conventionnels²⁰.



Figure 6 : Représentation schématique de la relation évolutive du complexe de *M. tuberculosis.* Les membres du CMT adaptés aux humains et aux animaux peuvent être classés en fonction de l'absence ou de la présence de délétions des régions de différence. Les souches de la bactérie CMT adaptées à l'homme sont regroupées en 7 lignes et sont associées à des régions géographiques spécifiques²¹.

M. tuberculosis étant pathogène, un laboratoire de niveau P3 est nécessaire pour les recherches la concernant. *M. bovis* est le modèle d'étude de *M. tuberculosis* très généralement utilisée. Le séquençage complet du génome de *M. bovis* a été réalisé en 2003 et a montré 99,95 % de similitude en comparant les génomes de *M. bovis* et *M. tuberculosis*²². Cependant, bien que ces deux bactéries soient génétiquement très proches, la régulation de l'expression de leurs gènes peut différer fortement. En revanche, la croissance de *M. bovis* est aussi lente que celle de *M. tuberculosis*, c'est-à-dire que leur temps de génération est mesuré en heures et non en minutes. L'atténuation *in vitro* de *M. bovis* permet d'obtenir le bacille de Calmette-Guérin (BCG) qui est le nom donné à la famille de vaccins dérivés en 1921. Les quatre souches de BCG les plus utilisées (BCG Copenhague, BCG Tokyo, BCG Glaxo et BCG Pasteur) différant par l'absence de certaines régions génétiques, BCG Tokyo²³.

M. canettii et *M. africanum* sont étroitement apparentés à *M. tuberculosis et* peuvent également causer la tuberculose. *M. canettii* a été signalé pour la première fois en 1969, son temps de génération est plus court que celui des *M. tuberculosis*¹⁹. La tuberculose causée à la fois par *M. canettii* et par *M. africanum* semble être une maladie émergente chez les patients africains ou d'origine africaine²⁴.

2.1.3.2 Les mycobactéries non tuberculeuses

En regard des mycobactéries du complexe *Mycobacterium tuberculosis*, il existe des mycobactéries non tuberculeuses (MNT ou encore mycobactéries atypiques) aux caractéristiques diverses. Ces mycobactéries ne sont pas les agents étiologiques de la tuberculose ou la lèpre. Plus de 140 espèces de MNT ont été identifiées jusqu'à présent²⁵. Contrairement à *M. tuberculosis*, qui est un agent pathogène humain obligatoire sans réservoir environnemental, les MNT sont généralement isolées de sources environnementales telles que l'eau et le sol. Chez l'homme, ils sont en général considérés comme des pathogènes de faible virulence. La transmission de l'infection d'une personne à l'autre étant rare, l'isolement des personnes infectées n'est pas nécessaire. Certains MNT sont cependant des agents pathogènes susceptibles de provoquer des maladies graves potentiellement mortelles^{26,27}.

Les maladies humaines dues aux MNT ont été classées en quatre syndromes cliniques distincts : maladie pulmonaire, lymphadénite, cutanée et maladie disséminée. Parmi ces quatre classes, les infections pulmonaires sont les plus fréquentes (65-90 %)^{26,28}.

Les MNT ont été classées selon la classification de Runyon basée sur la vitesse de croissance, la morphologie des colonies et la pigmentation. Grâce à cette classification, quatre groupes ont été créés. Les groupes I, II et III sont considérés comme des cultures avec une croissance lente, nécessitant un temps de génération similaire à celui de *M. tuberculosis*, c'est-à-dire 7 jours ou plus pour pousser. Ils ont été également classés en fonction de leur couleur. Si le pigment est produit uniquement lors d'une exposition à la lumière, il s'agit de photochromogènes (groupe I) ; si elle est produite à l'obscurité, ce sont des scotochromogènes (groupe II); si les bactéries ne sont pas fortement pigmentées, elles ne sont non-photochromogènes (groupe III). Enfin, les groupes IV sont des cultivateurs rapides qui se développent en moins de 7 jours.



Figure 7 : Classification des mycobactéries non tuberculeuses (MNT), figure adaptée de Porvaznik *et al.*²⁵.

Cependant, la classification de Runyon est devenue moins pertinente ces dernières années en raison des progrès de la mycobactériologie, et notamment de techniques de culture plus rapides, de sondes ADN et grâce à la chromatographie en phase liquide à haute pression²⁹. Parmi les MNT, *M. smegmatis* est une mycobactérie très fréquemment utilisée en laboratoire en raison de sa croissance rapide et du fait qu'elle est non pathogène. Le temps de génération est d'environ 4 à 5 heures et une culture dans un laboratoire au niveau P1 est suffisante. Son

enveloppe est similaire à celle de *M. tuberculosis* et il existe plus de 2000 gènes orthologues connus avec *M. tuberculosis*^{30,31}. En plus, 12 gènes sur 19 gènes de virulence de *M. tuberculosis* décrits à ce jour partagent des homologues étroitement apparentés chez *M. smegmatis*³².

La souche *M. smegmatis* mc²155 est facile à transformer et à cultiver dans la plupart des milieux de laboratoire (synthétiques ou complexes), une durée de 3 à 5 jours seulement étant suffisante pour former des colonies visibles. Ces propriétés en font le modèle le plus utilisé pour l'étude de la physiologie de *M. tuberculosis* et d'autres agents pathogènes mycobactériens. Elle est également utilisée pour la culture de mycobactériophages. C'est donc une espèce particulièrement privilégiée pour la recherche fondamentale.

3. Cycle infectieux de la tuberculose

La tuberculose est transmise d'une personne à l'autre par voie aérienne : un individu infecté toussant, éternuant ou crachant transmet par l'expectoration des bacilles à un individu sain (Fig 9A). Les gouttelettes expirées peuvent rester dans l'atmosphère pendant plusieurs heures avec une dose infectieuse qui reste très faible : la Dl₅₀ est estimée inférieure à dix bacilles chez l'humain. Les bacilles sont ensuite conduits jusqu'aux alvéoles pulmonaires où ils interagissent avec les cellules du système immunitaire de l'hôte.



Figure 8 : Progression de la tuberculose, de l'infection à la tuberculose active (pulmonaire), modifié de Pai *et al*. (2016).

Cette exposition conduit généralement à deux résultats possibles : l'élimination ou la persistance de l'agent pathogène. Dans le premier cas, l'agent pathogène est éliminé en raison de réponses immunitaires innées. Pour ces personnes, les traitements ne sont pas nécessaires. Il est reconnu depuis longtemps que, même parmi les contacts familiaux rapprochés de patients atteints de tuberculose, près de la moitié des personnes exposées restent saines. Le constat qu'il existe une prédisposition génétique à rester toujours négatif au test Mantoux (le principe du test Mantoux est expliqué dans la <u>partie 5.1.1</u>) malgré une exposition importante, est une explication possible de la résistance naturelle de certaines personnes à la tuberculose. Dans le deuxième cas, l'agent pathogène n'est pas éliminé et peut persister à l'état de repos ou à l'état latent : on parle d'infection tuberculeuse latente (ITL)²(Fig 8).



Figure 9 : Cycle infectieux de la tuberculose, adaptée de Nunes-Alves et al. en 2014³³.

Après phagocytose par les macrophages alvéolaires, une réponse pro-inflammatoire localisée induite conduit au recrutement des cellules mononucléées à partir de vaisseaux sanguins. Les mycobactéries qui ont survécu vont ensuite se multiplier à l'intérieur de ces cellules. Chez les enfants et les personnes immunodéprimées, une tuberculose primaire peut se développer après la dissémination des bacilles dans l'organisme mais son occurrence ne dépasse pas les 5 % (Fig 9B). Dans une grande majorité des cas, une réponse immunitaire adaptative se met en place, conduisant à la formation d'un agrégat cellulaire organisé : le granulome (Fig 9C). Le granulome est une structure formée de différentes cellules du système immunitaire et servant normalement à contenir les bacilles et à attendre la défaillance du système immunitaire pour les disséminer. Mais sa fonction est détournée par *M. tuberculosis* pour en former une structure protectrice lui permettant de se multiplier³⁴. Au centre du granulome se trouvent les macrophages infectés et des cellules géantes (fusion de macrophages) entourés de macrophages spumeux (gorgés de lipides) et de macrophages différenciés en cellules épithéloïdes. En périphérie, le granulome est composé essentiellement d'un manteau de cellules lymphocytaires. Il évolue ensuite par un processus de néo-vascularisation et par la production d'une chape fibreuse composée principalement de collagène qui sépare le centre du granulome de sa périphérie³⁵(Fig 9D). Avec cette réponse tissulaire, l'infection est asymptomatique et est donc très difficile à éradiquer, et l'hôte ne transmet pas l'infection aux autres. Environ 10 % des individus immunodéprimés à cause du SIDA, du diabète, de médicaments immunosuppresseurs, de la vieillesse, de troubles génétiques ou d'autres facteurs (essentiellement tous les facteurs réduisant le nombre, ou altérant les fonctions des lymphocytes T CD4⁺) développent une tuberculose secondaire à ce stade (Fig 9E). Cela signifie que les bacilles qui se multiplient dans les granulomes sont réactivés³⁵. Les lymphocytes T CD4⁺ produisent de l'interféron-y (IFNy) qui agit en synergie avec le facteur de nécrose tumorale α (TNF- α) produit par le lymphocyte T ou le macrophage; l'IFN γ et le TNF- α induisent conjointement l'activité antimicrobienne des macrophages en vue de limiter la croissance de M. tuberculosis. Il en résulte que les mycobactéries ne peuvent plus être contenues par les granulomes et disséminent dans tout l'organisme, y compris le cerveau, après la rupture du granulome. L'infection redevient symptomatique et de nouveau contagieuse². On parle de tuberculose active. L'accumulation de matériel lipidique dans les poumons conduit alors à la multiplication des bacilles. Une obstruction des bronches et plus fréquemment une insuffisance respiratoire sont les principales conséquences de cette tuberculose secondaire³³.

4. Réponses immunitaires

La probabilité et l'intensité de la transmission aérogène de *M. tuberculosis* d'une personne à l'autre dépendent de quatre facteurs : la durée de l'exposition ; l'intensité de l'exposition ; les facteurs de l'hôte liés à la toux et aux expectorations ; et les caractéristiques de virulence liées à la souche de *M. tuberculosis*³⁶.



Figure 10: Réponses immunitaires protectrices pendant l'infection de *M. tuberculosis*: les mycobactéries sont initialement contenues et la maladie se développe plus tard à la suite d'une réactivation. Le granulome est le site d'infection, de persistance, de pathologie et de protection. Les lymphocytes T effectrices (les lymphocytes T CD4⁺ et CD8⁺, telles que les lymphocytes T $\gamma \delta$, et les cellules T double positives ou CD4⁺ / CD8⁺ simple-positives qui reconnaissent l'antigène dans le contexte de CD1) et les macrophages participent à la lutte contre la tuberculose. L'interféron- γ (IFN- γ) et le facteur de nécrose tumorale-a (TNF-a), qui ont été produit par les lymphocytes T, sont d'importants activateurs des macrophages. L'activation des macrophages permet la maturation des phagosomes et la production de molécules antimicrobiennes.

Les mycobactéries inhalées peuvent être phagocytées par des macrophages alvéolaires^{37,38}, des cellules épithéliales, des cellules dendritiques (CD) ^{39,40} et des neutrophiles. Les bacilles sont transportés dans les ganglions lymphatiques locaux par les macrophages alvéolaires et les CD, où les lymphocytes T sont amorcés et développés par clonage.

Différentes lymphocytes T participent à la réponse immunitaire protectrice (Fig 10) :

- Les lymphocytes T CD4⁺, également appelés lymphocytes T auxiliaires (Th1). Ils reconnaissent des peptides antigéniques dans le contexte de produits géniques codés par le complexe majeur d'histocompatibilité (CMH) de classe II.
- Les lymphocytes T CD8⁺ qui reconnaissent des peptides antigéniques dans le contexte du CMH de classe I.
- Les lymphocytes T γδ qui reconnaissent des ligands antigéniques spéciaux indépendamment des molécules de présentation spécialisées (notamment les phospholigands).
- Les lymphocytes T CD1 restreints qui reconnaissent des glycolipides abondants dans la paroi de la mycobactérie.

M. tuberculosis est capable de persister dans le compartiment du phagosome⁴¹. Elle arrête la maturation de ce dernier à un stade précoce et inhibe fortement la fusion des phagolysosomes. En effet, de nombreux agents pathogènes bactériens sont tués dans le phagolysosome car c'est un environnement hostile, tandis que le phagosome précoce est un compartiment moins hostile où *M. tuberculosis* peut s'accommoder. Cependant, l'arrêt de cette maturation des phagosomes par *M. tuberculosis* n'est pas complet car certaines bactéries sont tuées ou incapable de se répliquer à cause de mécanismes antibactériens. Les différents lymphocytes T produisant l'interféron γ (IFN γ) sont des lymphocytes T auxiliaires (Th1). L'INF γ est la cytokine jouant le rôle de médiateur central de l'activation des macrophages. Elle agit en synergie avec le facteur de nécrose tumorale α (TNF- α) pour activer les macrophages. Les lymphocytes T CD4⁺ produisent également de la lymphotoxine α (LT α) qui participe à la protection contre la tuberculose. Certaines lymphocytes T CD8⁺, lymphocytes T $\gamma\delta$ et lymphocytes T CD1 restreints sécrètent de la perforine et de la granulysine qui tuent directement les mycobactéries au sein des macrophages⁴².

Au cours de la tuberculose active, une réponse immunitaire pulmonaire locale se met en place et permet de recruter à nouveau des macrophages immatures⁴³ et d'augmenter les réponses des antigènes spécifiques Th1 de *M. tuberculosis*, avec la sécrétion d'une grande quantité

45

d'IFN-γ locaux en même temps^{44,45}. La vulnérabilité à la tuberculose est associée à l'infection par le VIH-1 (avec la diminution du nombre et de la fonction des lymphocytes T CD4⁺ et CD8⁺)⁴⁶, le blocage du TNF- α^{47} et l'anomalie des cytokines inflammatoires comme l'IFN-γ et l'interleukine-2(IL-12) héréditaires⁴⁸.

5. Diagnostic et traitement de la tuberculose

Le choix d'un outil de diagnostic de la tuberculose dépend de l'objectif du test : la détection de l'infection tuberculose latente, la détection de la tuberculose active ou la détection de la résistance aux médicament antituberculeux.

5.1 Diagnostic de l'infection tuberculeuse latente

Deux tests d'immunodiagnostic sont disponibles pour l'identification de l'ITL : le test TCT et le dosage IGRA.

5.1.1 Test TCT

Le test TCT (Test cutané à la Tuberculine), également appelé le test de Mantoux ou l'intradermo-réaction à la tuberculine (IDR), est un test réalisé *in vivo*. Il consiste à injecter de manière intradermique une dose standard de tuberculine. Chez les personnes présentant une immunité à médiation cellulaire vis-à-vis de ces antigènes, une réaction d'hypersensibilité de type retardée se produira dans les 48 à 72 h suivantes. En mesurant le diamètre de l'induration, il est possible de savoir si le résultat est positif (5-18 mm) ou négatif (< 5mm). Le test TCT permet de mettre en évidence la présence de *M. tuberculosis* de façon indirecte, car il détecte les protéines produites lors de l'infection⁴⁹. Ce test présente plusieurs avantages : le faible coût des réactifs et de l'équipement, les compétences basiques et les exigences de laboratoire simple, surtout pour les pays à faibles ressources. Cependant, il comporte deux limitations majeures. D'une part il peut induire des faux-positifs sur les personnes vaccinées au BCG. D'autre part sa valeur prédictive est limitée car la plupart des individus avec des résultats positifs au test TCT ne progressent pas vers une tuberculose active.

5.1.2 Le dosage IGRA

Le dosage IGRA (*In-vitro Interferon Gamma Release Assay*) est un test sanguin *in vitro* de réponse immunitaire à médiation cellulaire : il mesure la libération d'IFN-γ par les lymphocytes T suite à la stimulation par des antigènes codés par RD1. Les antigènes RD1 sont plus spécifiques de *M. tuberculosis* que les antigènes de PPD, ils ne sont pas codés dans le génome des souches du vaccin BCG et ni dans la plupart des espèces de mycobactéries non tuberculeuses. Cependant, le dosage IGRA a une faible valeur prédictive².

Après de nombreuses études, les chercheurs ont montré que le test TCT et le dosage IGRA sont des tests acceptables mais imparfaits pour le diagnostic de l'ITL^{50,51}. Ils présentent en effet une sensibilité réduite chez les patients immunodéprimés et aucun de ces tests ne permet de différencier avec précision l'ITL et de la tuberculose active.

5.2 Diagnostic de la tuberculose active

Pour la détection de la tuberculose active, cinq principales technologies sont utilisées : les techniques d'imagerie (la radiographie du thorax et le PET-CT), la microscopie des frottis d'expectoration, les tests moléculaires, les méthodes de culture et la coloration de Ziehl-Nelsen.

5.2.1 La radiographie du thorax

La radiographie du thorax est un test de triage ou de dépistage établi. Cependant les rayons X manquent de spécificité et chez les personnes immunodéprimées, les images radiologiques peuvent être totalement atypiques. Les examens radiologiques du thorax anormaux doivent être suivis de tests microbiologiques.

5.2.2 La microscopie des frottis d'expectoration

La microscopie des frottis d'expectoration est le test de dépistage de la tuberculose active le plus utilisé dans les pays à revenu faible ou intermédiaire⁵². Le principe de ce test est d'analyser les échantillons d'expectoration par microscopie pour déterminer la présence des bactéries. Cependant, une grande limite de cette méthode est son manque de sensibilité.

5.2.3 Les tests moléculaires

Le dosage Xpert MTB / RIF (Cepheid, États-Unis) est un test moléculaire basé sur la technologie automatisée GeneXpert (Cepheid). En utilisant l'analyse PCR en temps réel, l'identification de *M. tuberculosis* peut être réalisée en moins de deux heures⁵³. En raison de sa précision supérieure à celle de la microscopie des frottis d'expectoration, il a devenu le seul test rapide recommandé par l'OMS comme test de diagnostic de première intention chez tous les adultes ou enfants suspectés d'être atteints de tuberculose active².

5.2.4 Examen de culture

L'examen de culture est un test de référence. Il utilise un milieu solide, le *Löwenstein-Jensen*, qui est riche en ressources énergétiques et protéines nécessaires pour la prolifération des bactéries. Des milieux liquides peuvent aussi d'être utilisés, ils sont plus sensibles permettent de réduire le délai de culture de 6 à 2 semaines.

5.2.5 Coloration de Ziehl-Nelsen

La coloration de Ziehl-Nelsen est la méthode la plus souvent utilisée. En effet, les mycobactéries sont très difficiles à colorer par des méthodes classiques tel que la coloration de Gram à cause de la présence des acides mycoliques dans la paroi. Cette méthode se base sur le fait que les mycobactéries sont acido-alcoolo-résistantes (B.A.A.R). Le colorant nommé carbol-fuschine phéniquée est utilisé en raison de sa solubilité en lipide. Pour la décoloration des bactéries non-acido résistantes, une solution acido-alcoolique est alors utilisée. Seules les mycobactéries résistent à cette décoloration.

5.3 Diagnostic de la résistance aux médicaments antituberculeux

Pour la détection de la résistance aux médicaments antituberculeux, deux types de méthodes phénotypiques sont souvent utilisés. La première se base sur le milieu culture, c'est-à-dire la capacité des bactéries à se développer en présence de médicaments. La deuxième se base sur les critères moléculaires : les mutations génétiques chez *M. tuberculosis* qui lui confèrent une résistance aux médicaments antituberculeux et qui peuvent être détectées². Le dosage de Xpert MTB / RIF en tant qu'un outil de diagnostic de la tuberculose active peut aussi d'être utilisé pour la détection des MDR-TB. Le test d'hybridation inverse sur bandelette (ou *Line*

Probe Assay) est une méthode basée sur l'hybridation inverse de l'ADN, les mycobactéries résistantes ne peuvent pas être amplifiés par la réaction PCR. Ce test permet de tester la résistance des médicaments de première ligne tels que l'isoniazide et la rifampicine mais aussi des médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième ligne tels que les fluoroquinolones et les médicaments de deuxième les fluoroquinolones et les médicaments de deuxième les fluoroquinolones et les médicaments de deuxième deuxième les fluoroquinolones et les médicaments de deuxième deuxiè

5.4 Traitements antibiotiques

Le taux de mortalité par tuberculose est très élevé sans traitement. En effet, d'après les études sur l'histoire naturelle de la tuberculose, environ 70 % des personnes atteintes de tuberculose pulmonaire décèdent dans les 10 ans suivant le diagnostic en l'absence de traitement antituberculeux. Des traitements médicamenteux efficaces ont été développés pour la première fois dans les années 1940.

La tuberculose peut être traitée efficacement en utilisant des médicaments de première intention : l'isoniazide (INH), la rifampicine (RIF), la pyrazinamide (PZA), l'éthambutol (EMB) et la streptomycine (SM). Le traitement actuellement recommandé par l'OMS est une thérapie combinatoire de quatre médicaments de première intention pour une durée de six mois : l'isoniazide, la rifampicine, l'éthambutol et le pyrazinamide. Ce traitement combinatoire a pour principe de prévenir la résistance acquise aux médicaments antituberculeux et d'améliorer leur efficacité. Les bacilles tuberculeux subissent souvent des mutations chromosomiques aléatoires qui les ont rendus résistants aux médicaments utilisés pour traiter la tuberculose. Heureusement, ces mutations restent peu fréquentes⁵⁶.La résistance acquise aux médicaments de la tuberculose est presque toujours due à un traitement inadéquat. Le traitement combinatoire peut accélérer la réponse de la maladie et raccourcir la durée du traitement nécessaire⁵⁷. Les deux médicaments les plus utilisés sont l'isoniazide et la rifampicine. L'isoniazide est responsable de la destruction initiale d'environ 95 % des organismes pendant les deux premiers jours de traitement. La rifampicine est le médicament le plus important parce qu'elle peut réduire la durée du traitement et permet d'obtenir un meilleur résultat⁵⁷. Les recherches ont montré qu'au moins 95 % des patients ont été guéris en utilisant les thérapies combinatoires de la rifampicine et de l'isoniazide associées à une phase d'introduction d'éthambutol, de streptomycine, ou des deux pendant 6 mois.

Cependant, ces traitements perdent en efficacité en raison de l'apparition de souches résistantes. Ils nécessitent ensuite l'utilisation de médicaments de deuxième intention. Ces derniers sont difficiles à se procurer et sont beaucoup plus toxiques et coûteux que les médicaments de première intention. Pour la suite, seuls les cinq antibiotiques (Tableau I) de la première intention vont être décrits.

5.4.1 Isoniazide

L'isoniazide (INH) est l'un des médicaments antituberculeux les plus puissants et les plus spécifiques. C'est un élément clé de la thérapie antituberculeuse depuis son introduction 1952 car *M. tuberculosis* y est très sensible (CMI de 0,02 à 0,2 µg/ml).

L'INH inhibe la protéine codée par le gène *inhA*, il s'implique également dans les protéines codées par les gènes *katG* et *ahpC*. Le gène *katG* code pour une catalase-peroxydase bifonctionnelle et permet de convertir l'INH en sa forme active⁵⁸. L'INH activé va ensuite inhiber la synthèse des acides mycoliques essentiels en inactivant l'énoyl-[*acyl-carrier-protein*] réductase à NADH codée par le gène *inhA*⁵⁹. Dans *M. tuberculosis*, le gène *ahpC* code pour une alkyl hydroperoxydase réductase impliquée dans la résistance à l'oxygène réactif et aux intermédiaires azotés réactifs. C'est pour ça que la résistance à l'INH est un processus complexe car plusieurs gènes ont été identifiés comme responsables.

5.4.2 Rifampicine

La rifampicine (RIF) a été introduite en 1972 en tant que médicament antituberculeux et présente une excellente activité stérilisante. Les CMI de la RIF varient de 0,05 à 1 μ g/ml sur des milieux solides ou liquides. Elle est l'un des antibiotiques les plus efficaces et est aussi un élément clé du traitement antituberculeux.

La RIF inhibe le gène *rpoB* codant pour la sous-unité β de l'ARN polymérase⁶⁰ qui est l'enzyme responsable de la transcription et de l'expression des gènes de mycobactéries. Elle entraine donc l'inhibition de l'activité de la transcription bactérienne et la destruction de l'organisme. Cependant, la possibilité pour des bactéries de développer une résistance à la RIF est élevée. La grande majorité des isolats cliniques de *M. tuberculosis* résistants à la rifampicine présentent des mutations dans le gène *rpoB* qui code pour la sous-unité β de l'ARN polymérase. Cela se traduit par des changements de conformation qui déterminent une faible affinité pour le médicament, et par conséquent le développement d'une résistance⁶¹.

5.4.3 Pyrazinamide

La pyrazinamide (PZA) est un antituberculeux utilisé pour raccourcir la durée du traitement antituberculeux. Le temps de traitement peut être raccourci de neuf à six mois en utilisant la pyrazinamide avec la rifampicine⁶². Elle constitue l'un des médicaments essentiels du traitement de la MDR-TB. Une des caractéristiques clés de la pyrazinamide est sa capacité à inhiber les bacilles semi-dormantes résidant dans des environnements acides contrairement aux autres antibiotiques⁶³. La pyrazinamide est un précurseur du médicament qui doit être d'abord converti en sa forme active, l'acide pyrazinoïque par la pyrazinamidase (PZase) ⁶⁴. Cette acide pyrazinoïque inhibe la production d'énergie membranaire, l'activité de la RpsA (protéine ribosomatique S1), qui est impliquée dans la traduction et du PanD, qui est impliqué dans le pantothénate et la synthèse du co-enzyme A⁶⁵. Le PZA n'est actif contre *M. tuberculosis* qu'à pH acide (par exemple 5,5)⁶⁶. Cependant, son activité reste assez faible même à pH acide : les CMI de la PZA sont entre 6,25 et 50 µg/ml⁶².

La PZase est codée dans *M. tuberculosis* par le gène *pncA*. Des mutations dans le gène *pncA* peuvent entraîner une réduction de l'activité de la PZase, ce qui pourrait constituer le principal mécanisme de résistance à la PZA chez le MTB^{67,68}.

5.4.4 Éthambutol

L'éthambutol (EMB) est un médicament de première intention essentiel dans le traitement de la tuberculose. Jouant un rôle important dans la chimiothérapie de la tuberculose pharmacorésistante, il a été introduit en 1966. Il permet de renforcer l'effet d'autres médicaments, notamment les aminosides, les rifamycines et les quinolones.

Les CMI de l'EMB pour *M. tuberculosis* sont comprises entre 0,5 et 2 µg/ml. L'EMB est un agent bactériostatique qui est actif pour la croissance des bacilles et n'a aucun effet sur les bacilles qui ne se répliquent pas. L'EMB interfère avec la biosynthèse de l'arabinogalactane de la paroi cellulaire⁶⁹. Plus précisément, il inhibe la polymérisation de la chaîne arabinane de la paroi cellulaire de l'arabinogalactane et du lipoarabinomannane (LAM) et induit une accumulation de D-arabinofuranosyl-P-décaprénol qui est un intermédiaire de la biosynthèse de l'arabinane^{65,70}.

L'arabinosyl transférase, codée par embB, est une enzyme impliquée dans la synthèse de l'arabinogalactane. Elle est la cible de l'EMB dans *M. tuberculosis*⁷¹. Le gène *embB* est organisé en opéron avec *embA* et *embC* dans l'ordre embCAB. Des mutations dans l'opéron embCAB, en particulier embB et occasionnellement embC, sont responsables de la résistance à l'EMB⁶⁵. Mais il faut noter qu'environ 35 % des souches résistantes aux EMB (CMI <10 µg/ml) ne présentent pas de mutations embB⁶⁵, la mutation de la DPPR synthase (codée par *ubiA*) a montré la résistance à EMB également⁷², ce qui suggère qu'il pourrait exister d'autres mécanismes de résistance à l'EMB. Des études complémentaires sont nécessaires pour identifier les nouveaux mécanismes potentiels de résistance à l'EMB.

5.4.5 Streptomycine

La streptomycine (SM) est un antibiotique aminoside introduit en 1948 en tant que médicament antituberculeux⁷³. La SM tue les bacilles tuberculeux en croissance active avec une CMI de 2 à 4 µg/ml, mais il est inactif contre les bacilles non croissantes ou intracellulaires⁶³. La streptomycine (SM) inhibe la synthèse des protéines en se liant à la sousunité 30S du ribosome bactérien, provoquant une lecture erronée du message de l'ARNm⁶⁵. La résistance à la SM est due aux mutations de la protéine ribosomale S12 codée par le gène *rpsL* et de l'ARNr 16S codé par le gène *rrs*⁷⁴. Cependant, environ 20-30 % des souches résistantes au SM (CMI, 32 µg/ml) ne présentent pas de mutations dans les gènes *rpsL* ou *rrs*, ce qui indique un ou plusieurs autres mécanismes de résistance⁷⁵.

Une mutation dans *gidB*, codant pour une 7-méthylguanosine (m⁷G) méthyltransférase conservée spécifique de l'ARNr 16S, s'est avérée causer une faible résistance à la SM chez 33 % des isolats résistants de *M. tuberculosis*⁷⁶.

Antituberculeux	CMI	Gènes	Mécanisme	Cibles
(année d'introduction)	(µg/ml)	impliqués	d'action	
Isoniazide (1952)	0,02 - 0,2	katG	Inhibition de la	Catalase peroxydase
		inhA	synthèse d'acide mycolique et	Énoyl-[acyl-carrier-protein]
		ahpC	d'autres effets multiples	réductase Alkyl hydroperoxidereductase
Rifampicine (1972)	0,05 - 1	гроВ	Inhibition de la synthèse d'ARN	Sous-unité β de l'ARN polymérase
Pyrazinamide (1952)	6,25 - 50	pncA	Inhibition de la trans-traduction	pyrazinamidase
Éthambutol (1966)	0,5 - 2	embA	Inhibition de la	Arabinosyl transférase
		embB	synthèse de l'arabinogalactane	
Streptomycine (1948)	2 - 4	rpsL	Inhibition de la	Protéine ribosomale S12
		rrs	synthèse des	ARNr 16S
		gidB	proteines	7-méthylguanosine méthyltransférase

 Tableau I
 Antituberculeux de la première intention en présentant les gènes impliqués et leurs cibles.

6. La tuberculose multirésistante

La résistance aux médicaments antituberculeux a été démontrée peu de temps après l'introduction d'une thérapie antituberculeuse avec la description de la résistance à la streptomycine en 1947⁷⁷. Plusieurs cas de tuberculoses multirésistantes (MDR-TB), définis comme des maladies causées par des souches résistant au moins à l'isoniazide et à la rifampicine, ont été identifiés au début des années 1990⁷⁸. Au cours des années suivantes, la tuberculose pharmacorésistante et en particulier la tuberculose multirésistante, a été reconnue comme potentiellement catastrophique pour la santé publique mondiale.

La tuberculose ultrarésistante (XDR-TB) a été définie par l'OMS en 2006. La XDR-TB est très difficile à traiter car elle résiste non seulement à l'isoniazide et à la rifampicine, mais

également à de nombreux médicaments de deuxième intention⁷⁹. Le taux de mortalité de la MDR-TB est élevé tant chez les personnes VIH positives que chez les personnes non infectées⁸⁰. La résistance aux médicaments est divisée en 2 types au niveau clinique : la résistance primaire et la résistance acquise. La résistance primaire se produit chez des personnes qui n'ont jamais été traitées pour la tuberculose et qui auraient été infectées par une souche résistante de *M. tuberculosis*, tandis que la résistance acquise se développe pendant le traitement de la tuberculose. La tuberculose pharmacorésistante est donc le produit d'une utilisation inappropriée de médicaments antituberculeux soit par les patients, soit par les cliniciens. Les causes courantes de pharmaco-résistance sont la prescription inappropriée comme l'approvisionnement irrégulier en médicaments, la mauvaise qualité du médicament, etc.⁷⁷. Le développement de tests de diagnostic plus sensibles et plus rapides et la découverte de nouveaux médicaments antituberculeux sont des objectifs prioritaires de l'OMS. C'est pour cette raison que de nouvelles cibles thérapeutiques ont besoin être découverts, et enfin de synthétiser de nouveaux médicaments antituberculeux.

7. Conclusion du chapitre

La tuberculose est une maladie ancienne qui avait un taux de mortalité très élevés. Les traitements médicamenteux efficaces sont apparus pour la première fois dans les années 1940. L'utilisation de médicaments antituberculeux ont permis de réduire la charge de morbidité à très faible dans certains pays. Cependant, dans la plupart des pays à faibles ressources, la tuberculose reste toujours le problème majeur de la santé publique.

L'OMS a ainsi été créé la *End TB Strategy* en 2014. Les objectifs spécifiques fixés pour 2030 dans la stratégie de lutte contre la tuberculose sont une réduction de 90 % du nombre absolu de décès dus à la tuberculose et une réduction de 80 % de l'incidence de la tuberculose (nouveaux cas pour 100 000 habitants par an) par rapport aux niveaux de 2015. La charge de morbidité causée par la tuberculose est en train de diminuer dans le monde, dans toutes les régions de l'OMS et dans la plupart des pays, mais pas assez rapidement pour atteindre les premiers jalons (2020) de la stratégie de lutte contre la tuberculose. D'ici 2020, le taux d'incidence de la tuberculose doit être descendu à 4–5 % par an, et la proportion de personnes atteintes de tuberculose décédées des suites de la maladie doit tomber à 10 %.

C'est pour cette raison que des recherches fondamentales sont effectuées régulièrement afin de comprendre le mécanisme d'action des agents anti-TB courants. Et de trouver de nouvelles cibles pharmacologiques permettant la synthèse de nouveaux médicaments. La paroi cellulaire des mycobactéries est tout à fait particulière par rapport aux celles des autres bactéries, c'est pour ça que les attentions des chercheurs ont été menés principalement sur cette paroi spéciale. Dans le chapitre suivant, la structure, la biosynthèse ainsi que le remodelage de la paroi cellulaire de *M. tuberculosis* vont être décrire en détail.

II. Paroi cellulaire de Mycobacterium tuberculosis

L'enveloppe cellulaire de *M. tuberculosis* et des autres mycobactéries est essentielle à la physiologie de ces bactéries, comme le prouve indiscutablement le caractère létal de l'inhibition ou de la perturbation de la biosynthèse de certains de ses composants. *M. tuberculosis* est une bactérie à Gram positif mais elle se distingue des autres bactéries à Gram positif et à Gram négatif par la composition tout à fait particulière de son enveloppe.



Figure 11: Organisation de l'enveloppe cellulaire des mycobactéries. L'enveloppe comprend d'abord une membrane cytoplasmique majoritairement constituée de phospholipides, mais peut contenir dans sa couche externe des glycolipides et des lipoglycannes possédant une ancre lipidique de type phosphatidyl-myo-inositol⁸¹. Ensuite est retrouvée la paroi qui est essentiellement constituée du complexe mAGP comprenant successivement le peptidoglycane, l'arabinogalactane et les acides mycoliques. Ensuite, les acides mycoliques, les glycolipides et les lipoglycannes forment une membrane externe asymétrique ou mycomembrane. Enfin, une couche externe ou capsule peut être présente chez les espèces pathogènes⁸². Dans ce schéma, les éléments protéiques ne sont pas représentés.

L'initiation de la définition de la paroi cellulaire des mycobactéries a été réalisée principalement dans les années 1960 – 1970. En 1982, Minnikin *et al.* ont proposé pour la première fois un modèle structural de la mycobactérie, qui est maintenant accepté pour décrire l'architecture de leur paroi cellulaire⁸³. Le point le plus marquant est l'importance des lipides qui représentent environ 40 % du poids de la matière sèche contre 5 % pour les autres

bactéries à Gram positif et 10 % pour les bactéries à Gram négatif⁸². Cette particularité des mycobactéries est à l'origine de la très faible perméabilité de leur enveloppe, qui est 10 à 100 fois inférieure à celle de *Pseudomonas aeruginosa*⁸⁴. Cette structure inhabituelle empêche l'hôte de détruire l'enveloppe mycobactérienne, et son caractère imperméable protège la mycobactérie des dommages⁸².

Outre sa nature lipophile, l'enveloppe mycobactérienne possède une architecture complexe dont les principaux éléments sont : la membrane plasmique, une paroi composite constituée du peptidoglycane (PG) lié à l'arabinogalactane (AG) sur lequel sont attachés les acides mycoliques, et enfin une couche externe ou capsule (pour les espèces pathogènes). Les acides mycoliques liés à l'arabinogalactane forment une membrane externe (ou mycomembrane) asymétrique dont le feuillet interne est formé par les acides mycoliques (certainement dans une conformation repliée) et le feuillet externe par un ensemble de lipides, glycolipides et protéines (Fig 11 et 12). La composition exacte et l'architecture de cette mycomembrane demeure encore mal connue⁸².



Figure 12 : Représentation schématique de l'enveloppe cellulaire de *M. tuberculosis*. En présentant la membrane plasmique (MP), le peptidoglycanne (PG), l'arabinogalactane (AG) et la membrane externe (ME), adaptée de Kaur *et al.* en 2009⁸⁵.

Dans la paroi, la macromolécule la plus importante est le mycolyl-ArabinoGalactane-Peptidoglycane (mAGP) (Fig 13), qui est essentiel pour la viabilité. Cette structure est à l'origine de la résistance de la bactérie, elle assure à la fois l'architecture pariétale et la survie bactérienne, mais elle est également le site de susceptibilité et de résistance à de nombreux médicaments antituberculeux^{84,85}. Le mAGP est constitué de trois macromolécules liées de manière covalente : le peptidoglycanne (PG), l'arabinogalactane (AG) et les acides mycoliques^{86,87}. Le squelette du PG est constitué d'un enchaînement de résidus de Nacétylglucosamine et d'acide muramique N-glycolylés, liés en β-1,4. Ce squelette est substitué par des chaînes d'arabinogalactane sur 10 à 12 % des acides muramiques modifiés. L'ensemble des PG forme la couche interne de la paroi. Les acides mycoliques constituent la couche externe de la paroi, qui est une barrière imperméable due à la présence de ses nombreux lipides; cette barrière imperméable permet de bloquer les divers agents antimicrobiens. Les acides mycoliques sont des acides gras de 60 à 90 carbones α -alkylés et β hydroxylés et qui sont perpendiculaires par rapport au plan de la membrane plasmique. Dans la paroi, ces longues chaînes carbonées sont majoritairement liées à l'arabinogalactane de manière covalente mais aussi à des composés tels que les tréhaloses sous forme mono- et dimycolates (TDM, TMM). L'enveloppe contiens des lipides et des glycoconjugués liés de manière non-covalente et possédants des propriétés d'activation du système immunitaire associées à la pathogénicité de M. tuberculosis, tels que le phosphatidylmyoinositol (PIM), le lipomannane (LM), le lipoarabinomannane (LAM), le diacylglycérol (DAG), les glycolipides phénoliques (PGL), le sulfolipide (SL) , le tréhalose dimycolate (TDM), le phthiocérol dimycocerosate (PDIM), les tréhaloses di- et pentacylés (DAT / PAT) et les triacylglycérols (TAG), les quantités relatives qui varient entre les espèces⁸⁵⁻⁸⁷.

La structure de la membrane plasmique de l'enveloppe mycobactérienne n'a pas montré de différences en comparaison avec les autres membranes plasmatiques biologiques, non seulement au niveau de l'apparence en coupes ultra-minces, mais aussi de la caractérisation des fonctions métaboliques classiques et de la composition chimique connue. Aucune différence évidente n'a été trouvée entre celles des espèces à croissance rapide et celles à croissance lente examinées⁸².

La structure, la biosynthèse et le remodelage des trois macromolécules du complexe mAGP, c'est-à-dire le peptidoglycane (PG), l'arabinogalactane (AG) et les acides mycoliques vont être décrits en détail par la suite.

59



Figure 13 : Représentation schématique du complexe mAGP proposée par Barry et al. en 2007⁸⁸.

1. Structure de la paroi cellulaire de *Mycobacterium tuberculosis*

Comme la macromolécule mAGP est essentiel pour les mycobactéries, la structure du complexe mAGP de *M. tuberculosis* (Fig 13), et plus précisément le peptidoglycane, l'arabinogalactane et les acides mycoliques vont être décrire séparément.

1.1 Le peptidoglycane

Le peptidoglycane (PG) est une structure macromoléculaire complexe située à l'extérieur de la membrane plasmique de presque toutes les bactéries. Son arrangement en forme de maille confère une rigidité à la cellule et une stabilité osmotique aux bacilles Gram négatif et Gram positif⁸⁹. Le PG de *M. tuberculosis* a été classé dans la catégorie A1γ, selon le système de classification de Schleifer et Kandler⁸⁵, comme *E. coli* et *Bacillus* spp. Il est composé de résidus alternés de N-acétylglucosamine (GlcNAc) et d'acide N-acétylmuramique (MurNAc) liés par des liaison $\beta(1 \rightarrow 4)^{90}$ et substitués par des chaînes latérales peptidiques fortement réticulées (Fig 13, partie peptidoglycane). Cependant, à la différence du PG chez E. coli, les résidus MurNAc de *M. tuberculosis* et de *M. smegmatis* ont également été oxydés pour former une fonction N-glycolyle et ainsi l'acide N-glycolylmuramique (MurNGly)^{91,92}. Bien que la fonction précise de cette modification n'ait pas encore été élucidée, il a été supposé qu'elle permettrait de former des liaisons hydrogène supplémentaires, renforçant ainsi l'intégrité de la structure en forme de maille de la couche de PG⁹⁰, ainsi que la protection de l'organisme à la dégradation du lysozyme⁹³. Les chaînes latérales tétrapeptidiques sont attachées aux résidus muramyl et constituées de L-alanyl-D-isoglutaminyl-meso-diaminopimélyl-D-alanine. Elles sont réticulées avec des peptides des chaînes glycaniques voisines, via le résidu acide mesodiaminopimélyl (*m*-DAP) de la première chaîne et le résidu D-alanine de la deuxième chaîne par la liaison $(3 \rightarrow 4)$ (ou D,D). Cependant, le nombre de liaison $(3 \rightarrow 3)$ (ou L,D) entre deux résidus de *m*-DAP est très grand également, et il augmente pendant la phase stationnaire (jusqu'à 80 %). *M. tuberculosis* est capable de modifier ces liaisons $(3 \rightarrow 4)$ en $(3 \rightarrow 3)$ sans avoir obligatoirement la synthèse de novo du PG. Cette modification pourrait fournir, aux mycobactéries, une protection contre les endopeptidases⁹⁴. La proportion de réticulation est de 70 à 80 % chez Mycobacterium spp. contre 50 % chez E. coli 85. Une autre différence par

rapport au PG de *E. coli* est l'utilisation des résidus MurNAc pour la fixation de la chaîne galactane de l'arabinogalactane. La 6-OH de certains résidus MurNAc est capable de former une liaison phosphodiester et de se lier au *Linker* α -L-rhamnopyranose-(1 \rightarrow 3)- α -D-GlcNAc(1 \rightarrow P) de l'arabinogalactane⁹⁵.

1.2 L'arabinogalactane

L'arabinogalactane (AG) a été reconnu comme le principal polysaccharide de la paroi cellulaire des mycobactéries dès les années 1950⁸⁵. L'intégrité du complexe mAGP repose sur l'arabinogalactane qui permet de maintenir les deux couches de la paroi associées (Fig 14). En fait, l'arabinogalactane est un hétéropolysaccharide qui forme un lien physique entre le peptidoglycane et l'acide mycolique en produisant une région hydrophile visqueuse⁸⁵. Contrairement aux autres polysaccharides bactériens, la structure de l'arabinogalactane ne repose pas sur la répétition d'une unité courte, mais plutôt sur un *linker* au PG et d'une chaîne galactane liée au peptidoglycane par un *linker* disaccharidique et substituée par plusieurs chaînes d'arabinane⁹⁶. Chez les mycobactéries, l'arabinogalactane est lié à l'extrémité réductrice du PG de manière covalente via un *Linker*, c'est-à-dire α -L-rhamnopyranose-(1 \rightarrow 3)- α -D-GlcNAc(1 \rightarrow P), conduisant à la présence d'environ 1,3 molécules l'arabinogalactane / 10 unités répétées de PG, et entraînant une substitution de 10 à 12 % des résidus d'acide muramique par une liaison phosphodiester⁹⁵. D'après les calculs, la masse moléculaire de l'arabinogalactane est environ 15 kDa.

Le PG et l'arabinogalactane forment un macro polymère placé entre la membrane cytoplasmique et la couche externe d'acide mycolique des bacilles de la tuberculose. Certaines recherches ont montré que l'arabinogalactane constituait environ 35 % de la masse totale de l'enveloppe mycobactérienne⁹⁷⁻¹⁰⁰. À noter, les résidus d'arabinoses et de galactoses de l'arabinogalactane sont présents sous forme furanique, soit le D-galactofuranose (Galf) et le D-arabinofuranose (Araf), qui sont rencontrés moins fréquemment dans la nature. L'arabinogalactane contient environ 125 résidus de monosaccharides distribués entre une chaîne de galactane constituée de 30 résidus de D-Galf et trois domaines arabinane contenant chacun 31 résidus de D-Araf. L'extrémité réductrice de l'arabinogalactane est reliée au PG par un *Linker* α -L-Rhap-(1 \rightarrow 3)- α -D-GlcNAc-(1 \rightarrow P). Cette unité de *Linker* connecte le domaine

galactane de l'arabinogalactane à la position C-6 de résidus MurNGlyc engagés⁹⁵. Cette structure est ensuite prolongée par 30 résidus de Gal*f* avec alternance des liaisons en $\beta(1\rightarrow 5)$ et $\beta(1\rightarrow 6)^{96}$. Selon les études réalisées chez *M. tuberculosis* et *Corynebacterium glutamicum,* il apparait que sur cette chaine linéaire de Gal*f* se trouvent en position 8, 10 et 12, trois chaines d'arabinane d'environ 30 résidus chacune et liées à la chaîne de galactane en C-5 des résidus $\beta(1\rightarrow 6)$ de Gal*f*¹⁰¹. Les résidus d'Ara*f* sont liés en $\alpha(1\rightarrow 5)$ avec des points de branchements en C-3 formant des 3,5-Ara*f*. L'extrémité non réductrice des chaines d'arabinane forme un motif hexa-arabinofuranosyl [β -D-Ara*f*-($1\rightarrow 2$)- α -D-Ara*f*]₂-3,5- α -D-Ara*f*-($1\rightarrow 5$)- α -D-Ara*f*] avec les acides mycoliques attachés en position 5 des deux extrémités β -D-Ara*f* et des 2- α -D-Ara*f* en position antépénultième¹⁰².



Figure 14 : Structure du mycolyl-arabinogalactane (mAG) et son attachement au peptidoglycane et aux acides mycoliques, adapté de Jankute *et al.* 2015¹⁰²⁻¹⁰⁴.

Cette structure complète de l'arabinogalactane de *M. tuberculosis* a été quelque peu révisée par les mêmes auteurs, et cet arabinogalactane ne comporterait que deux chaînes d'arabinane par molécule d'arabinogalactane en moyenne avec une taille moyenne de 26 résidus d'arabinose. Les positions 8, 10 et 12 de la chaîne de galactane restent les seules à pouvoir porter une chaîne de galactane, et la taille maximale d'une chaîne de galactane reste de 31 résidus d'arabinose¹⁰⁴.

Des résidus de GalNH₂ et de succinate ont été identifiés sur la position C-2 de certains résidus intérieurs de 3,5- α -D-Ara $f^{102,105-107}$. Il a été estimé qu'un seul résidu de GalNH₂ est présent par molécule d'arabinogalactane, et il se trouve exclusivement dans les parois de mycobactéries à croissance lente, tandis que jusqu'à trois résidus de succinates sont présentés par arabinogalactane chez les mycobactéries à croissance lente et rapide. De plus, les résidus de succinate sont présentés seulement sur les chaînes arabinanes non-ramifiées par les acides mycoliques, et qu'aucune des chaînes arabinanes décorées n'avait les deux substituants en même temps¹⁰². Jusqu'à présent, la structure de l'arabinogalactane de *M. tuberculosis* est la seule avoir été caractérisée en détail, même s'il subsiste encore des zones d'ombre comme la répartition des mycolates sur les chaînes d'arabinane et les relations entre mycoylation/présence du résidu de GalNH₂ ou de succinate intercalair¹⁰². Il reste quand même attendu que l'arabinogalactane des mycobactéries du complexe M. tuberculosis soit de structures similaires. Par contre, si l'on compare la structure de l'arabinogalactane (et de la paroi par extension) de M. tuberculosis (in vitro) et celle de M. Leprae (in vivo), des différences notoires sont à relever. M. Leprae présente un arabinogalactane plus petit (59 résidus de monosaccharides contre 79 en moyenne) avec une chaîne de galactane plus courte (13 résidus de galactose contre 23). L'architecture des chaînes arabinanes reste similaire avec un double branchement. Par contre, il est observé un taux de branchement d'arabinogalactane sur le PG supérieur chez *M. leprae* (3,0 arabinogalactane/10 unités de PG contre 1,4 arabinogalactane/10 unités de PG), ce qui se traduit par un degré de mycoylation de la paroi supérieur chez *M. leprae* (20,8 mycolates/10 unités de PG contre 16,4/10 unités de PG)¹⁰⁴. En dernier point, l'hétérogénéité de structure de l'arabinogalactane de M. tuberculosis (taille, nombre de chaînes d'arabinane, présence du succinate ou de GalNH₂ intercalaire, ...) est indubitable au vu des travaux déjà publiés mais reste très mal connue à cette date.

64

1.3 Les acides mycoliques

Les acides mycoliques sont les principaux composants protecteurs de la paroi cellulaire qui protège le bacille tuberculeux des composés chimiques nocifs et du système immunitaire de l'hôte. Ils sont liés principalement aux résidus d'arabinose de l'arabinogalactane en position 5 (Fig 13), cependant ils peuvent estérifier les tréhaloses pour former des glycolipides de type mono- et dimycolates (TMM et TDM)¹⁰⁵.

Les acides mycoliques sont des acides gras à très longue chaîne (jusqu'à 90 carbones) α-alkylés et β-hydroxylés⁸² et qui font partie intégrante du complexe mAGP et contribuent à la fluidité et à la perméabilité de la paroi cellulaire¹⁰⁶⁻¹⁰⁸. Chez *M. tuberculosis*, les acides mycoliques sont des acides gras hydroxylés très hydrophobes. Comme montré dans la figure 15, les acides mycoliques de Mycobacterium sont constitués par deux chaînes alkylées, une chaîne principale, le méromycolate, qui contient de C₅₄ à C₆₃ acides gras, et une branche α (ou la chaîne courte) qui contient de C₂₂ à C₂₆ acides gras^{16,106,109}. Les variations structurales sont portées principalement par le méromycolate. Ces variations structurales peuvent être des cyclopropanes, une variation de la longueur de la chaîne, des insaturations et des fonctions supplémentaires telles que des fonctions cétone, méthoxyle, ester, hydroxyle et époxyde¹⁶. Ces groupements se situent en position proximale et/ou distale au niveau du méromycolate. Les acides mycoliques de *M. tuberculosis* se composent de trois types structuraux distincts principaux : les acides α -mycoliques contiennent des cycles cyclopropanes proximales et distales, en configuration cis uniquement ; les acides méthoxy mycoliques et les acides cétomycoliques contiennent des cycles cyclopropanes proximales en configuration cis ou trans avec une branche méthyle adjacente (Fig 15) 16,110,111 . L'acide α -mycolique est la forme la plus abondante (>70 %), tandis que les acides méthoxy et céto-mycoliques sont les composants mineurs (10 à 15 %)¹¹².



Figure 15 : Représentation structurale des acides mycoliques de *M. tuberculosis* avec ses trois types principaux : l'acide a-mycolique avec les cycles cyclopropanes en configuration *cis* ; l'acide méthoxy-mycolique avec le cycle cyclopropane proximale en configuration *cis* et *trans*, et l'acide céto-mycolique avec le cycle cyclopropane proximale en configuration *cis* et *trans*, adapté de Takayama *et al.* 2005¹⁰⁹.

Les acides mycoliques sont le composant principal de la paroi cellulaire mycobactérienne, elles jouent un rôle important en tant qu'une couche protectrice. Les acides mycoliques jouent également d'autres rôles importants en tant que composants structurels de la paroi cellulaire et de l'enveloppe¹⁰⁹. Par exemple, les cycles cyclopropanes dans les acides mycoliques de *M. tuberculosis* contribuent à l'intégrité structurelle du complexe de la paroi cellulaire¹¹³⁻¹¹⁵ et protègent le bacille du stress oxydatif (péroxyde d'hydrogène)¹¹⁶.

2. Biosynthèses du complexe mAGP

2.1 Le peptidoglycane

La biosynthèse du peptidoglycane contient trois parties principales : les étapes cytoplasmiques, les étapes associées à la membrane et enfin l'assemblage du PG dans le périplasme (Fig 16, partie synthèse).

2.1.1 Étapes cytoplasmiques

La biosynthèse des PG commence par la conversion du fructose-6-phosphate en UDP-Nacétylglucosamine (UDP-GlcNAc), qui est catalysée par les enzymes GlmS (*Rv3436*), GlmM (*Rv3441c*) et GlmU (*Rv1018c*). Parmi ces trois enzymes, seule GlmU a été caractérisée chez *M. tuberculosis*^{117,118}. GlmU est une acétyltransférase / uridyltransférase bifonctionnelle qui convertit la glucosamine-1-phosphate (GlcN-1-P) en N-acétylglucosamine-1-phosphate (GlcNAc-1-P), et catalyse ensuite la formation de l'UDP-GlcNAc à partir du GlcNAc-1-P et de l'uridine triphosphate¹¹⁷⁻¹¹⁹.

La conversion de l'UDP-GlcNAc en UDP-MurNAc est catalysée par la MurA (*Rv1315c*), une UDP-N-acétylglucosamine énolpyruvyl transférase, et la MurB (*Rv0482c*), une UDP-N-acétylénolpyruvyl-glucosamine réductase. Plus précisément, la MurA catalyse le transfert du phospho-énol-pyruvate (PEP) sur l'UDP-GlcNAc^{120,121}, et la MurB¹²² est une oxydoréductase dépendante du NADPH qui réduit le groupement énoylpyruvyl en D-lactate résultant en la formation du UDP-MurNAc. La NamH (*Rv3838*) peut hydrolyser l'UDP-MurNAc en UDP-MurGlyc, cela permet d'obtenir les deux substrats UDP-muramyl⁹².

L'assemblage de la tige peptidique de l'unité monomère PG (UDP-MurNGlyc ou UDP-MurNAcpentapeptide) est catalysé par une série de quatre ligases Mur ATP-dépendant : MurC (*Rv2152c*), MurD (*Rv2155c*), MurE (*Rv2158c*) et MurF (*Rv2157c*), catalysent l'addition de L-Alanine, D-Glutamate, *m*-DAP et D-Alanyl-D-Alanine, respectivement¹²³⁻¹²⁶. Le produit est nommé comme « le nucléotide de Park » : UDP-MurNAc/Glyc-L-Ala-D-isoglu-*m*-DAP-D-Ala-D-Ala est formé. Le D-Glu nécessaire à la synthèse de PG est fourni par un glutamate épimérase Murl (*Rv1338*). Le DapF (*Rv2726c*) convertit la L,L-DAP en *m*DAP. Le substrat di-peptidique de MurF (D-Ala-D-Ala) est synthétisé par l'alanine épimérase Alr (*Rv3423c*), suivi de la D-Ala-D-Ala ligase Ddl (*Rv2981c*)¹²⁷.

2.1.2 Étapes associées à la membrane

La première étape membranaire dans la biosynthèse du PG fait intervenir la phospho-Nacétylmuramoyl-pentapeptide transférase MraY (*Rv2156c*). MraY catalyse le transfert du nucléotide de Park sur le décaprényl phosphate, donnant le Lipide I (Fig 16, entouré en rouge). Suite à la synthèse du Lipide I, la N-acétylglucosamine transférase MurG (*Rv2153c*) catalyse le transfert d'un résidu de GlcNAc depuis l'UDP-GlcNAc sur le Lipide I, formant ainsi le Lipide II¹²⁷ (Fig 16, entouré en jaune).

Certaines mycobactéries peuvent avoir des étapes d'amidation après la formation du Lipide II¹²⁷. MurT (*Rv3712*) et GatD (*Rv3713*) sont responsables de l'amidation du résidu D-Glu dans la tige peptidique des Lipides I et II. Le résidu *m*-DAP peut subir également de l'amidation en utilisant l'AsnB (*Rv2201*). L'amidation du PG joue un rôle crucial dans la synthèse et la maturation du PG. Une hypothèse est que l'amidation pourrait faciliter la translocation du Lipide II à travers la membrane cytoplasmique car une réduction de la polarité y est associée¹²⁸.

La translocation du Lipide II à travers la membrane a fait l'objet de controverses, le rôle de la Lipide II *flippase* ayant été contesté entre FtsW (*Rv2154c*) et MurJ (*Rv3910*). Il a été initialement annoncé par Ruiz¹²⁹ que c'était MurJ chez *E. coli* qui en était responsable, mais Mohammadi *et al.*¹³⁰ ont identifié FtsW comme ayant une activité *flippase in vitro*. Cependant, le groupe de Ruiz¹³¹ a récemment montré des données *in vivo* chez *E. coli* qui confirment davantage l'identification de la *flippase* comme étant MurJ. Néanmoins, des recherches supplémentaires sur les deux enzymes sont nécessaires pour identifier complètement la Lipide II *flippase* mycobactérienne.



Figure 16 : Schéma de la biosynthèse, la dégradation et le recyclage du peptidoglycane des espèces *Mycobacterium* avec les différentes enzymes impliquées. Le symbole * marque les enzymes qui n'ont pas encore été identifiés expérimentalement. Elles comprennent les protéines qui sont putatives ou complètement inconnues. Adapté de Maitra *et al.* 2019¹²⁷.

2.1.3 Assemblage du peptidoglycane dans le périplasme

Une fois que le Lipide II est transféré dans le périplasme, il peut être utilisé comme substrat pour les réactions de polymérisation ultérieures et l'assemblage du PG. Deux enzymes bifonctionnelles PBP1 et PBP2 codées par le gène *ponA1* (*Rv0050*) et *ponA2* (*Rv3682*) respectivement, transglycosylent du Lipide II monomères en reliant le groupement GlcNAc au groupement muramyle de la chaîne PG¹³². PBP1 et PBP2 possèdent également une activité transpeptidase, qui peuvent former les liaisons transversales classiques (3→4) entre les résidus *m*-DAP et D-Ala¹⁰³. PBP1 et PBP2 contiennent toutes les deux un autre domaine, composé d'une hélice transmembranaire et d'un domaine PASTA (*Penicillin-Binding Protein and serine/threonine-kinase associated*). Jusqu'à présent, cinq paralogues ont été identifiés chez *M. tuberculosis* : Ldt_{Mt1} (*Rv0116c*), Ldt_{Mt2} (*Rv2518c*), LdtMt3 (*Rv1433*), Ldt_{Mt4} (*Rv0192*) et Ldt_{Mt5} (*Rv0483*). Cependant, seules les souches déficientes de *Ldt_{Mt1}* et *Ldt_{Mt2}* présentent des différences phénotypiques au niveau structural de PG¹³³. Les L,D-transpeptidases réalisent la conversion des liaisons (3→4) en (3→3) en clivant le résidu D-Ala de la chaîne donneuse restante, et laissant une chaîne tripeptide dans la phase stationnaire des mycobactéries¹³⁴.

2.2 L'arabinogalactane

La biosynthèse de l'arabinogalactane du complexe mAGP commence à partir d'un phosphate décaprényle (C₅₀-P) et se termine par la liaison des mycolates et la fixation au peptidoglycane. Elle se déroule en deux étapes consécutives. La première se déroule sur la face cytoplasmique de la membrane plasmique et procède à la synthèse de la chaîne de galactane. La seconde se déroule sur la face exoplasmique de la membrane plasmique et complète la biosynthèse de l'arabinogalactane jusqu'à son attachement au PG (Fig 17).

2.2.1 L'unité Linker

La biosynthèse de l'unité *Linker* commence par le transfert de GlcNAc-1-P de UDP-GlcNAc sur le support lipidique par une GlcNAc transférase WecA (Rv1302)^{135,136}, formant ainsi C₅₀-P-P-GlcNAc, sous le nom glycolipide 1 (GL-1)¹³⁵.

WecA de *M. tuberculosis* présente des homologies de séquence importantes avec WecA des bactéries à Gram-négatif comme *E. coli*. WecA d'*E. coli* permet de catalyser la première étape de la biosynthèse du lipopolysaccharide antigène-O¹³⁶. La caractérisation biochimique de cette enzyme montre que WecA est bien une phosphate décaprényle α -acétylglucosaminyltransférase responsable de la formation de GL-1¹³⁷, et que le gène *wecA* est essentiel à la croissance de *M. smegmatis*¹³⁶.

L'unité *Linker* disaccharique est ensuite complétée par l'action ultérieure de la rhamnosyltransférase WbbL codée par le gène *Rv3265c*, qui ajoute un résidu de rhamnose (Rha) à la position 3 du résidu de GlcNAc de GL-1, formant ainsi le glycolipide 2 (GL-2)¹³⁸. WbbL utilise le dTDP-rhamnose comme substrat pour la formation de GL-2. Son inactivation chez un mutant thermosensible de *M. smegmatis* a montré que WbbL est essentielle pour la formation de la liaison arabinogalactane-peptidoglycane et donc, elle est vitale pour les mycobactéries. Le dTDP-Rha permet de fournir le rhamnose (Rha) grâce à une réaction en quatre étapes. En effet, pour catalyser la conversion de l' α -D-glucose-1-P et de thymidine triphosphate (TPP) en dTDP-Rha, quatre enzymes sont employées : RmIA, RmIB, RmIC et RmID (codées par *Rv0334, Rv3464, Rv3465* et *Rv3266c* respectivement). RmIA (*Rv0334*) définit la séquence des réactions en convertissant le dTTP α -D-glucose 1-phosphate en dTDP-glucose PPI¹³⁹. Le produit obtenu subit trois réactions séquentielles catalysées par la dTDP-D-glucose 4,6-déshydratase (RmIB), la dTDP-4-céto-6-désoxy-D-glycose 3,5 épimérase (RmIC) et la dTDP-Rha synthase (RmID)¹⁴⁰⁻¹⁴². L'inactivation par mutation de *RmIB, RmIC* ou *RmID* dans *M. smegmatis* a été montrée comme létale pour les mycobactéries¹⁴³.

2.2.2 La chaîne galactane

Une fois que le *Linker* est formé, la synthèse de la chaîne galactane peut commencer. Le *Linker* joue un rôle d'accepteur dans l'addition de résidus de Gal*f*, 30 résidus de Gal*f* sont rajoutés séquentiellement et ils sont reliés de manière linéaire. Le substrat est l'UDP-Gal*f* qui provient de l'UDP-Gal*p* après action de l'UDP-Gal*p* mutase GL*f* (*Rv3809c*)^{144,145}. L'UDP-Gal*p* est synthétisée à partir d'UDP-glucose par l'UDP-Gal*p* épimérase (probablement codée par *Rv3634*¹⁴⁵). Deux galactofuranosyltransférases ont été impliquées dans la biosynthèse de la chaîne galactane.

GlfT1 (*Rv3782*) est reconnue comme la GalT responsable du transfert initial des deux premiers résidus de Gal*f*, en utilisant l'UDP-Gal*f* comme substrat pour former les C₅₀-P-P-GlcNAc-Rha-Gal*f* (GL-3) et C₅₀-P-P-GlcNAc-Rha-Gal*f*-Gal*f* (GL-4)^{146,147}. Tandis que la galactosyltransférase GlfT2 (*Rv3808c*) effectue les additions suivantes de résidus de Gal*f*. GlfT2 est une galactosyltransférase bifonctionnelle qui est capable de former alternativement des liaisons $\beta 1 \rightarrow 5$ et $\beta 1 \rightarrow 6$. Elle polymérise la majeure partie de la chaîne de galactane jusqu'au C₅₀-P-P-GlcNAc-Rha-Galf₃₀^{146,148}. GlfT2 présente d'autres caractéristiques intéressantes en plus de son activité bifonctionnelle : elle est capable de contrôler intrinsèquement la longueur de la chaîne du galactane en utilisant un mécanisme d'attachement de substrat¹⁴⁹.

L'élongation de la chaîne de galactane se fait dans la face cytoplasmique de la membrane plasmique alors que les étapes ultérieures ont lieu au niveau de la face exoplasmique. Les protéines impliquées dans ce passage ne sont pas encore connues malheureusement.

2.2.3 La chaîne arabinane

La connaissance des enzymes impliquées dans l'élongation de la chaîne arabinane est plus limitée. Des études *in vitro* ont montré qu'en présence d'un seul donneur connu d'Araf : le polyprenolmonophosphoryl- β -D-arabinose (DPA), les résidus d'Araf sont transférés directement sur l'intermédiaire précédent C₅₀-P-P-GlcNAc-Rha-Galf₃₀⁷⁰.

Le DPA est synthétisé à partir du 5-phosphoribosyl-1-pyrophosphate (pRpp) grâce à l'action successive d'une phosphoribosyl-1-pyrophosphate synthétase PrsA (*Rv1017c*), d'une décaprényl phosphoryl-5-phosphoribose synthétase UbiA (*Rv3806c*) et finalement d'une phospholipide phosphatase (*Rv3807c*) pour former le décaprényl phosphoribose (DPR). UbiA permet de catalyser le passage du pRpp en décaphénylphosphoryl-5-phosphoribose (DPPR). Elle a été montrée létale pour les mycobactéries et la délétion de *ubiA* chez *C. glutamicum* (NCgl2781) a montré une absence totale de la chaîne arabinane dans la paroi cellulaire, ce qui indique que le DPA est le donneur unique du résidu Araf dans la biosynthèse de l'arabinogalactane¹⁵⁰. La phospholipide phosphatase Rv3807c catalyse la déphosphorylation de la DPPR en décaprényl-5-phosphoribose (DPR), et l'orthologue de Rv3807c chez *M. smegmatis* (MSMEG_6402) a été montré comme un gène non essentiel¹⁵¹.

L'épimérisation du DPR en DPA est un processus en deux étapes pour lequel deux enzymes, DprE1 et DprE2, ont été impliquées. DprE1 (*Rv3790*) est une oxydoréductase qui permet
d'oxyder la position 2 du ribose de DPR pour former du décaprénylphosphoryl-2-céto-β-Dérythropentofuranose (DPK)¹⁵². DprE2 (*Rv3791*) réduit ensuite la C2-céto de la DPK, ce qui aboutit à la synthèse complète du donneur Araf unique, le DPA¹⁵². Une troisième enzyme Rv2073c serait impliquée dans la réaction d'épimérisation. Des études ont été réalisées sur DprE1 et DprE2, ainsi que leurs orthologues correspondants chez *M. smegamatis* (MSMEG_6382 et MSMEG_6385, respectivement) et *C. glutamicum* (NCgl0187 et NCgl0186, respectivement) ^{153,154}. Les résultats obtenus ont montré que seule la délétion du gène *dprE1* est létale, conduisant à un arabinogalactane avec un taux réduit d'arabinose et de rhamnose, mais un taux supérieur de mycolates de tréhalose par rapport à la souche parentale^{153,154}. Ces études montrent que DprE1 est une cible potentielle pour une nouvelle génération de médicaments antimicrobiens. Une attention particulière a donc été menée sur DprE1, principalement en raison de sa sensibilité à l'inhibition par une variété d'inhibiteurs. Ces inhibiteurs présentent également des activités anti-mycobactériennes puissantes, comme par exemples, les benzothiazinones (BTZs) qui montrent une activité très puissante car ils sont capables de tuer *M. tuberculosis* avec une puissance nanomolaire en ciblant DprE1¹⁵⁵.

L'addition du premier résidu d'Araf à partir du DPA sur la chaîne de galactane se fait aux positions 8, 10, 12 et est catalysée par une arabinofuranosyltransférase (ArafT) du locus emb, AftA (Rv3792)¹⁵⁶. De récents résultats obtenus sur la délétion de l'orthologue d'AftA chez *M. smegmatis, MSMEG_6386*, ont montré qu'elle était létale¹⁵⁷. Ensuite, pour en revenir à l'élongation de la chaine, EmbA (Rv3794) et EmbB (Rv3795) qui possèdent une activité α 1-5 transférase, agissent en hétérodimère pour catalyser la suite de la. Les enzymes Emb présentent également une activité importante dans la formation du motif hexa-arabinofuranosyl terminal, elles sont toutes les deux des α -(1→3) arabinosyltransférases¹⁵⁸.



Figure 17 : Biosynthèse de l'arabinogalactane, adapté de Jankute *et al.* 2012 et Jankute *et al.* 2015^{103,159}.

Malgré la formation de l'arabinogalactane soit essentielle à sa viabilité des mycobactéries, la délétion des orthologues de EmbA et EmbB chez *M. smegmatis* conduit à des mutants viables mais avec un arabinogalactane incomplet. En effet, la délétion conduisant une réduction d'arabinose importante produisant ainsi un arabinogalactane altérée¹⁶⁰.

Une nouvelle $\alpha(1\rightarrow 3)$ ArafT, AftC (*Rv2673*), a récemment été caractérisée chez *M. smegmatis*. Elle joue un rôle dans la branchement en $\alpha(1\rightarrow 3)$ du domaine arabinane interne à l'extrémité non réductrice de l'arabinogalactane^{161,162}. AftC est aussi inhibée par l'éthambutol¹⁶³. En plus, AftD (*Rv0236c*) de *M. tuberculosis* est proposée également pour une activité $\alpha(1\rightarrow 3)$ transférase qui pourrait avoir d'autres fonctions dans la biosynthèse de l'arabinogalactane¹⁶⁴. Enfin, le transfert des résidus de l'Araf terminal en $\beta(1\rightarrow 2)$ à partir du DPA au motif hexaarabinofuranoside à l'extrémité non réductrice de l'arabinogalactane est calalysé par AftB (*Rv3805c*). Ceci signifie la fin de la synthèse de l'arabinogalactane avant la décoration par les acides mycoliques^{165,166}.

2.2.4 Modifications de l'arabinogalactane

Deux modifications de l'arabinogalactane sont présentes dans sa région interne chez *M. tuberculosis* et probablement chez toutes les mycobactéries à croissance lente. Il s'agit du transfert d'un résidu de D-GalNH₂^{167,168} ou de succinate¹⁰² sur le résidu α -(3 \rightarrow 5) Araf. Les enzymes clées pour la synthèse et le transfert du résidu de D-GalNH₂ n'ont été que récemment découvertes¹⁶⁹. La formation de l'intermédiaire polyprenyl-P-D-GalNAc est catalysée par l'enzyme PpgS (*Rv3631*). Cet intermédiaire est ensuite désacétylé et transporté au travers de la membrane plasmique. Le transfert du résidu de D-GalNH₂ sur le C₂ d'un résidu d'Araf dans la région centrale branchée en 3 \rightarrow 5, est catalysé par une deuxième protéine Rv3779¹⁶⁹. L'inactivation de ces deux gènes empêche totalement la synthèse et le transfert de ce résidu chez *M. tuberculosis*¹⁰³.

Cependant, le moment auquel exact l'arabinogalactane est mycolylé, c'est-à-dire avant ou après son attachement au PG, reste inconnu. Grâce aux analyses enzymatiques directes et à la mutagenèse dirigée, les enzymes impliquées dans le transfert des résidus mycolyl sur l'arabinogalactane ont été identifiées. Le complexe antigène 85 est composé de trois enzymes FbpA (*Rv3804c*), FbpB (*Rv1886c*) et FbpC (*Rv0129c*), il a montré une activité mycolyltransférase¹⁶⁵. Ces enzymes catalysent le transfert de mycolates sur le tréhalose, pour former du tréhalose monomycolate et du tréhalose dimycolate¹⁶⁵. Ces résultats ont été confirmés par l'analyse du mutant déficient pour l'antigène 85C isolé de *M. tuberculosis*¹⁷⁰.

2.3 L'acide mycolique

La biosynthèse des acides mycoliques (Figure 18) commence par la biosynthèse des acides gras utilisant deux *Fatty Acid Synthases* (FASs), FAS-I et FAS-II¹⁶. FAS-I est un complexe multienzyme retrouvé chez les eucaryotes et les procaryotes avancés, tandis que le FAS-II est un système de synthétase retrouvé chez les plantes et les bactéries¹⁰⁹. FAS-I permet de synthétiser la chaîne courte (C₁₆₋₁₈ ou C₂₄₋₂₆) des acides gras¹⁶. Ces acides gras saturés peuvent ensuite former la chaîne α -alkylés (C₂₄) ou être prolongés par FAS-II pour former la chaîne méromycolates (C₅₆).

Les systèmes FAS-I et FAS-II rajoutent successivement une unité de deux carbones (acétate) du malonyl-coenzyme A (CoA) à un groupe acyl naissant. Le malonyl-CoA est fourni par la carboxylation de l'acétyl-CoA, catalysée par l'acétyl-CoA carboxylases, ACCases. Celles-ci sont constituées de plusieurs sous-unités, notamment les sous-unités AccA et AccD¹⁷¹. L'ACCase AccA3-AccD6 joue un rôle important dans la biosynthèse d'acide mycolique en fournissant du malonyl-CoA au complexe FAS-II¹⁷². A partir du malonyl-CoA et un transporteur acyl, holo-AcpM, le malnoyl-AcpM est enfin généré par une malonyl-AcpM transacylase FabD¹⁷³.



Figure 18 : Représentation schématique de biosynthèse des acides mycoliques de *M. tuberculosis*, adaptée de Jankute *et al.* 2015¹⁰³. La protéine InhA est inhibée par l'isoniazide (INH).

L'action de FAS-II est initiée par une réaction de condensation catalysée par FabH (β -cétoacyl-ACP synthase III) en utilisant du malonyl-AcpM et du palmitoyl-CoA, donnant un intermédiaire β -cétoacyl-AcpM. Celui-ci est réduit par MabA (β -cétoacyl-ACP réductase) pour générer du β -hydroxyacyl-AcpM, qui est ensuite déshydraté par HadAB / BC (β -hydroxyacyl-ACP déshydratase) pour donner un intermédiaire *trans*-2-énoyle-AcpM. Celui-ci est réduit par InhA (énoyl-ACP réductase) pour produire un acyl-ACP allongé par deux atomes de carbone. Le cycle FAS-II se poursuit ensuite par une condensation, catalysée par KasA / B (β -cétoacyl-ACP synthase), pour revenir au β -cétoacyl-AcpM. Ce cycle se répète jusqu'à ce que la chaîne acyle atteigne C₄₂₋₆₂, formant le méromycolate saturé à longue chaîne.

Diverses modifications sont apportées à la chaîne saturée en C_{42-62} suite à sa synthèse via FAS-II, telles que la *cis-/trans*-cyclopropanation et/ou l'addition de groupes céto ou méthoxy en utilisant des enzymes de type CmaA1-2, MmaA1-4 et PcaA^{113,114,174}. Il s'agit de modifications apportées à la chaîne du méromycolate qui modifient la fluidité et la perméabilité de la paroi cellulaire, tout en assurant une protection contre le système immunitaire de l'hôte pendant l'infection¹¹⁶. La chaîne de l'acide méromycolique (C₄₂₋₆₂) est donc activée pour la condensation de type-Claisen par la génération d'un méromycolyl-AMP par FadD32 (acyl gras AMP ligase)¹⁷⁵. Ce substrat méromycolyl-AMP est lié à la chaîne courte d' α -alkylé (C₂₂₋₂₄), cette réaction est catalysée par Pks13 pour produire l'acide α -alkyl- β -céto-mycolique.

Les étapes finales de la synthèse sont le transport et la liaison des acides mycoliques nouvellement synthétisés au complexe peptidoglycane-arabinogalactane. Cependant, ces étapes restent à clarifier. Jusqu'à présent, il est connu que la plus grande proportion d'acides mycoliques se trouve sous la forme d'esters liés dans le mAGP mais ils sont aussi retrouvés sous la forme de tréhalose mono- et di-mycolates (TMM et TDM)¹⁰⁹. L'acide α -alkyl- β -céto-mycolique, dont la formation est catalysée par Pks13, est transféré sur le tréhalose qui est ensuite réduit par l'enzyme CmrA pour obtenir le TMM. Le TMM est transporté du cytoplasme vers le périplasme par MmpL3. Il a été utilisé par le complexe Antigène 85 pour la biosynthèse du TDM et du mAG avec la libération du tréhalose libre dans l'espace pseudopériplasmique¹⁷⁶.

III. Les glycosides Hydrolases

Les glucides, ou glycanes, sont des composés organiques qui contiennent un groupe carbonyle (cétone ou aldéhyde) et au moins deux groupes hydroxyle. Ils sont répartis entre oses (monosaccharides) et osides, qui sont des polymères d'oses liés par des liaison osidique (oligosaccharides et polysaccharides). Les glucides constituent l'un des quatre composants principaux de la cellule avec les protéines, les acides nucléiques et les lipides. En plus de leur rôle connu au niveau du métabolisme énergétique, les glycanes présentent de nombreuses autres fonctions importantes, ils ont été montrés impliqués dans de nombreuses maladies et mécanismes biologiques tels que le cancer et l'inflammation.

Les glycosides hydrolases (GH) ou glycosidases forment une famille d'enzymes qui catalysent l'hydrolyse de liaisons glycosidiques, c'est-à-dire la liaison entre deux groupes hydroxyles de deux monosaccharides, en libérant au moins un composé osidique. Elles catalysent donc la dégradation d'oligosaccharides, de polysaccharides, de glycolipides, de glycoprotéines et de tout type de glycoconjugués. Pour les glycosidases, la molécule d'eau joue un rôle d'accepteur. Les glycosidases jouent un rôle important dans de nombreux processus biologiques, comme la modification post-traductionnelle des glycoprotéines, la digestion intestinale et le catabolisme lysosomal des glycoconjugués¹⁷⁷. Environ 1-3 % du génome d'un organisme code pour la dégradation des glucides, dont les GH ce qui montre l'importance du métabolisme des glucides pour les organismes vivants¹⁷⁸. Les liaisons glycosidiques entre les monosaccharides sont extrêmement stables, l'hydrolyse spontanée des liaisons glycosidiques est ainsi très difficile¹⁷⁹. La plupart des glycosidases sont des catalyseurs très efficaces : la vitesse d'hydrolyse peut être améliorée d'un ordre de grandeur de 10¹⁷ fois en présence du catalyseur en comparant avec la réaction non enzymatique, c'est-à-dire avec des constantes de vitesse catalytiques jusqu'à 1000 s^{-1 180,181}.

Les monosaccharides sont extrêmement divers au niveau stéréochimique : ils peuvent avoir divers groupes fonctionnels, différentes tailles de cycle (furanoside ou pyranoside) et peuvent être assemblés de manières différentes. Il existe plus de 10¹² isomères possibles pour un hexa-saccharide réducteur¹⁸². Les structures glucidiques possèdent donc un potentiel d'encodage d'information gigantesque *via* un « glycocode ». Cette diversité a été exploitée par les organismes vivants en utilisant des oligosaccharides et des polysaccharides pour de

79

nombreuses fonctions biologiques, de stockage, de structure jusqu'à des rôles de signalisation spécifiques. Les glucides ne diffèrent les uns des autres que par des variations mineures de stéréochimie, les GH ont acquis leur spécificité par le biais de différences structurelles subtiles et ont évolué à partir d'un nombre d'ancêtres limité¹⁸³.

1. Les glycosidases : généralité

Les glycosidases hydrolysent les liaisons glycosidiques en utilisant une grande variété de substrats, notamment les oligosaccharides, les polysaccharides, les glycoprotéines, les glycolipides. Le clivage de la liaison entre deux monosaccharides ou entre un sucre et une autre espèce peut se faire par hydrolyse pour générer un glycane libre (glycosidases), par phosphorolyse pour donner un glycane-1-phosphate (phosphorylases), par transglycosylation pour donner un autre type de glycoside (transglycosidases) ou par élimination pour donner des produits à base de monosaccharide insaturé (lyases)¹⁸³.

Les glycosides hydrolases sont caractérisées selon quatre principaux facteurs :

- 1. La spécificité anomérique (α ou β) ;
- La stéréochimie de la réaction (clivage du produit par inversion ou rétention par rapport à l'anomérie de configuration);
- L'emplacement du clivage par rapport au substrat (*exo-* si les monosaccharides sont hydrolysés à partir de l'extrémité non-réductrice d'une chaîne de sucre ou *endo-* si l'hydrolyse se produit à l'intérieur de la chaîne de polysaccharides) ;
- 4. La nature du sucre clivé.

Par exemple, selon cette nomenclature, une *exo*- β -galactosidase qui agit par rétention est une enzyme qui permet de cliver la liaison β -galactoside en libérant du β -galactose libre à partir de l'extrémité non-réductrice d'une chaîne d'oligo- ou de polysaccharides.

2. Mécanisme d'action des glycosidases

Deux types de mécanismes catalytiques sont observés chez les glycosidases, selon le changement possible ou non de la configuration du carbone anomérique d'un produit obtenu après réaction d'hydrolyse enzymatique. Nous pouvons distinguer le mécanisme d'inversion et le mécanisme de rétention (publié par Koshland en 1953)¹⁸⁴. Ces mécanismes se réalisent

en passant par un état de transition intermédiaire, dans lequel des ions oxocarbéniums se forment. Deux groupements carboxyliques provenant de l'acide glutamique ou aspartique du site actif de l'enzyme sont impliqués dans cette catalyse¹⁸⁵. Parmis ces deux résidus acides, l'un est un résidu nucléophile qui permet de réaliser l'attaque nucléophile et l'autre joue un rôle d'acide/base.

2.1 Mécanisme d'inversion

Pour le mécanisme d'inversion, la configuration du carbone anomérique d'un produit est inverse à celle de son substrat. Ce type d'enzyme a généralement besoin de deux résidus catalytiques, dont un portant un groupement acide carboxylique au niveau du site actif. Parmi ces deux résidus catalytiques, l'un réagit comme un catalyseur basique qui facilite l'attaque de l'eau sur le carbone anomérique tandis que l'autre réagit comme un catalyseur acide qui facilite le départ de l'aglycone. Cette réaction passe par un état de transition, dans lequel des ions oxocarbéniums se forment (Fig 19). La distance entre les résidus catalytiques doit donc être suffisamment importante, c'est-à-dire entre 6 et 12 Å.



Figure 19 : Mécanisme typique des glycosidases agissant par inversion de la configuration anomériques du substrat¹⁸⁶.

2.2 Mécanisme de rétention

Pour le mécanisme de rétention, la configuration du carbone anomérique du produit est la même que celle de son substrat. Cette fois-ci, la réaction s'effectue en passant par deux états de transition intermédiaires et non un seul. Cela signifie que le carbone anomérique est inversé deux fois par la formation d'ion oxocarbénium : il reste donc de configuration identique. Comme pour le mécanisme d'inversion, deux résidus catalytiques portant un

groupement acide carboxylique (acide aspartique ou glutamique) sont en jeu, avec de nouveau un résidu jouant le rôle de catalyseur acide et l'autre celui d'un catalyseur basique. La première étape est une étape de glycosylation, le catalyseur basique réagit comme un nucléophile, attaque le carbone anomérique et un intermédiaire covalent de type glycosylenzyme est alors formé. Le catalyseur acide ensuite agit comme un donneur de proton, permettant la protonation de l'oxygène glycosidique pour activer l'aglycone en tant que groupement partant. Une molécule d'eau peut ensuite remplacer l'oxygène de la liaison osidique rompue. La deuxième étape est une étape de déglycosylation : le même résidu réagit d'abord comme un catalyseur de base, donc déprotone la molécule d'eau, puis attaque le carbone anomérique et hydrolyse enfin l'intermédiaire covalent glycosyl-enzyme. Cette étape ré-inverse la configuration du carbone anomérique identique à celle du substrat. La distance entre les deux résidus catalytiques doit être de l'ordre de 5,5 Å¹⁸⁶.

Les mécanismes de rétention et d'inversion sont les deux mécanismes les plus couramment observés pour les glycosidases, ils sont impliqués dans la plupart des familles de GHs sauf pour les GH4 et GH109. Pour ses deux familles, l'hydrolyse se réalise via un mécanisme d'élimination. Le NAD⁺ est nécessaire en tant que cofacteur mais n'est pas consommé dans le cycle catalytique. De plus, un métal divalent, le Mn²⁺ en général, est requis pour stabiliser les charges négatives se développant dans les états intermédiaires et les états de transition¹⁸³.



Figure 20 : Mécanisme typique des glycosidases agissant par rétention de la configuration anomériques du substrat¹⁸⁶.

2.3 Nécessité d'un ion métallique

Le mécanisme catalytique des glycosidases repose entièrement sur une paire d'acide aminé (acide aspartique ou acide glutamique). Cependant, dans certains cas des glycosidases, un ion métallique est nécessaire pour la catalyse. Par exemple, les *exo*-mannosidases des familles GH 38, 47 et 92 nécessitent un ion de métal divalent, tel que le zinc ou le calcium.



Figure 21 : Mécanisme de réaction proposé par Zhu *et al.* pour une a-mannosidase de la famille GH92 agit par inversion, l'ion Ca^{2+} établit un pont entre les atomes d'oxygène de C2 et C3 tout en interagissant également avec l'eau attaquante¹⁸⁷.

Comme montré dans la figure 21, l'ion métallique permet de relier les atomes d'oxygène en C2 et C3 et est associé à la stabilisation de l'état de transition¹⁸⁷. Les β -galactosidases de la familles GH2 ont besoin de magnésium pour fonctionner. L'ion magnésium établit un contact direct avec le catalyseur acide / base et joue donc un rôle dans le déplacement du pKa du catalyseur acide / base ¹⁸⁸ ¹⁸⁹. Il a également été montré que le calcium est nécessaire pour la fonctionnement d'une α -glucosidase et d'une α -galactosidase de la famille GH97¹⁹⁰.

3. Élucidations du mécanisme de la glycosidase

3.1 Stéréochimie des glycosidases

Le monosaccharide en position terminale réductrice généré par hydrolyse s'équilibre spontanément sous les deux formes anomériques α et β . C'est pour cette raison que la détermination de la stéréochimie d'une glycosidase est compliquée. Le temps de mutarotation entre les deux formes est très variable en fonction du sucre, du tampon, du pH et de la température. Par exemples, pour le glucose, la demi-vie pour la mutarotation à pH 6,15 et à 35 °C est de 11 min¹⁸³.

Pour analyser la stéréochimie de la glycosidase, la Résonance Magnétique Nucléaire (RMN) est une des méthodes souvent utilisées car la réaction enzymatique peut être réalisée directement dans le tube RMN¹⁹¹. Il est relativement aisé de déterminer l'anomérie du premier produit formé en analysant le signal du proton anomérique des sucres libres. La chromatographie en phase liquide à haute pression (HPLC), la chromatographie en phase gazeuse (CPG) et la polarimétrie ont également été souvent utilisées¹⁸³. Cependant, la mutarotation s'effectue parfois trop rapidement pour que l'analyse par RMN soit réalisée correctement, en particulier avec les sucres contenant des groupes acides (acides carboxyliques ou phosphates). Dans ce cas, le méthanol peut être utilisé comme accepteur autre que l'eau, en produisant des produits méthyle-glycosides qui ne peuvent subir de mutarotation. La nature de ces produits peut ensuite être élucidée en déterminant la configuration anomérique par RMN ou CPG^{183,192,193}.

3.2 Identification des résidus catalytiques

Les résidus catalytiques d'une glycosidase peuvent être identifiés par alignement de séquence car pour certaines familles des glycosidases, leurs résidus catalytiques ont déjà été déterminés par une étude mécanistique à l'aide d'inhibiteurs ou de mutants. Bien que cette méthode puisse suggérer des résidus catalytiques, elle n'est pas décisive car dans la plupart des cas, plusieurs résidus d'acide glutamique ou aspartique conservés sont observés, et tous ne seront pas situés dans le site actif ou impliqué dans la catalyse. Des informations supplémentaires sur les rôles dans la catalyse sont généralement obtenus par mutagenèse dirigée suivi d'une analyse cinétique détaillée du mutant produit¹⁹⁴. Par exemple, pour identifier les résidus catalytiques de la β -mannosidase *Cm*Man5A, qui a été classée dans la famille GH5 chez CAZy, les auteurs sont d'abord réalisés un alignement de séquence primaires de *Cm*Man5A avec les autres glycosidases caractérisées de la famille GH5. Deux résidus candidats ont été retenus, le Glu215 et Glu330. Ensuite pour valider ces deux résidus, ils ont muté ces deux acides glutamiques en alanines puis calculer leur constantes *michealiennes*¹⁹⁵.

En l'absence de ces types de données, les structures de cristallographie aux rayons X peuvent donner des informations sur les identités de ces résidus. Si des informations structurelles 3D avec une résolution atomique sont disponibles, l'identité des résidus d'acides aminés du site actif peut souvent être déterminée par un examen détaillé de la région du site actif. L'identification concluante des résidus du site actif nécessite fréquemment la détermination de la structure d'un complexe enzyme / inhibiteur ou enzyme / substrat. Cependant, les structures 3D ne sont pas toujours disponibles pour la grande majorité des glycosidases. Même pour celles dont les structures sont connues, pour obtenir des informations structurelles sur les complexes ligand enzyme peut être difficile. Par conséquent, la vérification des résidus proposés nécessite des analyses cinétiques comme les constantes *michealiennes* sur les mutants¹⁹⁴.

3.2.1 Identification des résidus nucléophiles

L'identification des résidus nucléophiles des glycosidases peut être réalisée dans un premier temps par mutagenèse dirigée. En mutant les acides aminés candidats en alanine (Ala), le groupe carboxylique est remplacé par le groupe méthyle, une perte de K_{cat} de 10 000 fois est attendue pour le résidu nucléophile en comparant avec l'enzyme native¹⁸⁶. Par contre, en utilisant un substrat qui a un excellent groupe partant tel que le 2,4-dinitrophényl glycoside, une augmentation importante de l'activité catalytique devrait être observée pour le résidu nucléophile lors de l'addition d'un nucléophile anionique tel que l'azoture ou le formate¹⁹⁶ (Fig 22). Cet effet est dû au fait que l'anion nucléophile réagit plus rapidement avec la glycosylenzyme qu'avec l'eau en l'absence de catalyseur basique¹⁸⁶.



Figure 22 : Récupération de l'activité catalytique d'une glycosidase par l'azoture.

Le résidu d'acide glutamique (Glu) peut également être muté en acide aspartique (Asp) pour éloigner la fonction carboxylique d'environ 1 Å mais sans changer le mécanisme¹⁹⁶. Une baisse significative de l'activité catalytique devrait alors être observée mais sans la supprimer entièrement. Une autre méthode très souvent employée consiste à piéger directement le résidu nucléophile par la formation d'un intermédiaire glycosyl-enzyme stable, qui peut être ensuite analysé par spectrométrie de masse ou par RMN¹⁹⁴.



Figure 23 : Inactivation et réactivation d'une glycosidase agit en β -rétention par le 2-déoxy-2-fluoro glycosides. (A) Inactivation par la formation de la fluoro glycosyl-enzyme intermédiaire. (B) Réactivation par transglycosylation de l'intermédiaire fluoro glycosyl-enzyme¹⁹⁷.

Cet intermédiaire glycosyl-enzyme peut être formé à l'aide de 2-déoxy-2-fluoro glycosides pour les glycosidases qui agissent en β -rétention (Fig 23) et de 5-fluoro-glycosyl fluorures pour les glycosidases qui agissent en α -rétention¹⁸⁶ (Fig 24). Cette étape d'inactivation est basée sur des mécanismes enzymatiques. Le groupe fluor étant électronégatif, il permet de ralentir la réaction en déstabilisant l'état de transition d'ion oxocarbénium et de diminuer les vitesses de glycosylation et de déglycosylation. En tant qu'excellents groupements partants, le fluorure ou le dinitrophénol accélèrent l'étape de glycosylation sans affecter l'étape de déglycosylation. L'intermédiaire piégé s'accumule et est très stable, ne se retournant qu'extrêmement lentement par hydrolyse. La réactivation de cet intermédiaire a été mise en évidence en présence d'un sucre comme accepteur, par un processus de transglycosylation. L'étape de réactivation prend donc beaucoup plus de temps que l'étape d'inactivation, la demi-vie se mesurant typiquement en jours. La digestion protéolytique de l'enzyme marqué donne des peptides courts, dont l'un parmi tous les autres a le sucre-fluoro reste attaché. Seul ester dans ce complexe, la liaison glycosyl-enzyme est particulièrement sensible à la fragmentation par spectromètrie de masse. Ainsi, la localisation du peptide marqué peut être réalisée par cette dernière.





3.2.2 Identification des résidus acides / basiques

Le remplacement d'un acide aminé du catalyseur acide / basique général en alanine présente une diminution du K_{cat} d'environ 1 000 fois par rapport au type sauvage, en utilisant comme substrat naturel les disaccharides. Cependant, les substrats tels que les dinitrophényl glycosides qui ont un excellent groupement partant et qui n'ont pas besoin d'une assistance par un proton au départ ont moins d'effet sur¹⁸⁶. Mais en raison de l'accumulation de l'intermédiaire réactionnel, le K_M est généralement fortement réduit.

4. Classifications des glycosidases

Selon la classification de l'International Union of Biochemistry and Molecular Biology Enzyme

Commission (IUBMB-CE) qui est basée sur la spécificité du substrat, les glycosides hydrolases sont classées dans la famille EC 3.2.1., correspondant aux enzymes qui permettent de catalyser l'hydrolyse des liaisons O- ou S-glycosidiques. Et plus précisément, par exemple, les enzymes de la famille EC 3.2.1.146 (β -D-galactofuranosidase), EC 3.2.1.55 (α -L-arabinofuranosidase), EC 3.2.1.99 (arabinanase) et EC 3.2.1.- (β -xylanase) pourraient avoir une activité susceptible en utilisant l'arabinogalactane comme substrat. Cependant, cette classification n'est pas suffisamment fiable pour classer et prédire les relations structurales et fonctionnelles des glycosidases du fait que c'est un système basé sur les spécificités du substrat, les glucides présentent une diversité structurale plus importante que celle des protéines ou des acides nucléiques. De plus, l'existence d'une évolution divergente à partir d'un ancêtre commun vers de nouvelles spécificités et d'une évolution convergente vers des mécanismes enzymatiques similaires montre qu'il n'y a souvent aucune corrélation entre la numérotation EC d'une classe d'enzymes et les séquences des enzymes qui effectuent ces réactions¹⁹⁸.

En 1990, une classification des glycosidases basée sur la similarité des séquences primaires d'enzymes actives a été créée. La base de données CAZy (*Carbohydrate Active enZymes*, http://www.cazy.org) regroupe les séquences protéiques des glycoenzymes en cinq classes : les glycoside hydrolases (162 familles), les glycosyltransférases (106 familles), les polysaccharide lyases (36 familles), les carbohydrate estérases (16 familles) et les enzymes à activité auxiliaire (16 familles). Cependant de nombreuses séquences n'ont été classées dans aucune famille. L'un des principaux avantages de la classification CAZy par similarité de séquences, donc par la structure 3D, est les informations qu'il donne sur le mécanisme catalytique : les résidus de site actif qui caractérisent un mécanisme catalytique donné sont fortement conservés dans une famille donnée de GH (sauf pour certains cas exceptionnels). Tous les membres de la même famille GH ont le même mode de réaction stéréochimique^{199,200}.

Clan	Caractéristique structurale	Mécanismes	GHs	
		d'action	Familles	
GH-A	Tonneau (β/α) ₈	β-rétention	1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53,	
			59, 72, 79, 86, 113, 128, 147, 148, 157, 158	
GH-B	eta-Jelly roll	β -rétention	7, 16	
GH-C	eta-Jelly roll	β -rétention	11, 12	
GH-D	Tonneau ($\beta/lpha$) ₈	lpha-rétention	27, 31, 36	
GH-E	6-fold β -propeller	lpha-rétention	33, 34, 83, 93	
GH-F	5-fold β -propeller	Inversion	43, 62	
GH-G	Tonneau (α/α) ₆	α -inversion	37, 63, 100, 125	
GH-H	Tonneau ($\beta/lpha$) $_8$	α -rétention	13, 70, 77	
GH-I	α+β	β -inversion	24, 80	
GH-J	5-fold β -propeller	lpha-rétention	32, 68	
GH-K	Tonneau (β/α) ₈	β-rétention	18, 20, 85	
GH-L	Tonneau ($lpha/lpha$) $_6$	lpha-inversion	15, 65	
GH-M	Tonneau ($lpha/lpha$) $_6$	β -inversion	8, 48	
GH-N	Hélice-β	lpha-inversion	28, 49	
GH-O	Tonneau (α/α) ₆	β -rétention	52, 116	
GH-P	Tonneau ($lpha/lpha)_6$	β -rétention	127, 146	
GH-Q	Tonneau ($lpha/lpha$) $_6$	β -inversion	94, 149, 161	
GH-R	Tonneau (β/α) ₈	α-rétention	29, 107	

Tableau II Les différents clans de glycosidases selon la base de données C	AZy.
--	------

La similarité des séquences est souvent exprimée par une conservation de leur structure. Les membres d'une famille de GH donnée ont le même repliement structural. De plus, de nombreuses familles qui ont des séquences différentes partagent la même structure 3D, bien qu'elles agissent sur des substrats différents. Les familles GH montrant essentiellement la même topologie au niveau de la structure tridimensionnelle et conservant les mêmes sites actifs ont été regroupées en *clans*¹⁹⁹. Par exemple, le clan GH-A comprend une grande partie de GH qui agissent par rétention et qui montrent une activité sur les β -D glycosides. Leurs résidus catalytiques sont situés sur les mêmes endroits au niveau des extrémités des brins β d'un tonneau (β/α)₈. Le tableau II présente les différents caractères de glycosidase dans chaque clan.

La base de données CAZy donne accès à énormément d'informations provenant des résultats de séquençage génomique, ce qui explique pourquoi elle reste un outil très utile et précieux pour la prédiction de la structure 3D et de la fonction des glycosidases. Elle nous permet d'extrapoler facilement les caractéristiques structurales et mécanistique d'une glycosidase d'une famille bien caractérisée à une autre famille.

5. Structure des glycosidases

Les structures tridimensionnelles des protéines sont fortement conservées par rapport à leurs séquences et plusieurs familles basées sur des séquences similaires pouvent avoir des repliements semblables. Dans la classification CAZy, des enzymes ayant différentes spécificités de substrat se trouvent parfois dans la même famille, indiquant une évolution divergente pour obtenir de nouvelles spécificités, comme par exemple dans les familles GH1, GH13 et GH16. De même, les enzymes qui hydrolysent le même substrat peuvent être classées dans différentes familles. Par exemple, les cellulases se trouvent dans les familles GH5 et GH8. Depuis la structure 3D de la première enzyme identifiée par cristallographie est (le lysozyme par Johnson et Phillips en 1965), le nombre de structures 3D établies a explosé dans les années 1990 et a montré qu'une multitude de structures enzymatiques différentes peuvent catalyser la même réaction.

Les structures principales rencontrées pour les glycosidases sont présentées ci-dessous.

5.1 Tonneau (β/α)₈

Un tonneau (β/α)₈, aussi appelé tonneau *TIM*, est une des structures les plus présentes chez les enzymes. Environ 10 % de protéines structuralement caractérisées contiennent au moins un domaine en tonneau (β/α)₈²⁰¹. Il comporte huit hélices α et huit feuillets β , les brins β formant le cylindre à l'intérieur et les hélices α amphiphiles bordant l'extérieur (Fig 25a). Les sites actifs des glycosidases en tonneau (β/α)₈ sont toutes constituées par les boucles des extrémités C_t des brins $\beta^{201,202}$.

Chez les glycosidases, les clans A, D, H, K et R avec les familles 14, 29, 31, 67, 84, 89, 97, 99, 101, 106, 138, 140, 153 présentent cette structure. Pour le clan A, les résidus catalytiques acide/base sont situés sur le brin 4 et le résidu responsable de l'attaque nucléophile est situé sur le brin 7 du tonneau. Cependant certaines familles présentent un changement de leur tonneau *TIM* comme adaptation. Par exemple, les glycosidases des familles 38, 56 et 57 ne possèdent que 7 répétitions de l'unité β/α et sont sous forme tonneau (β/α)₇ au lieu de (β/α)₈.

Pour la famille 6, les glycosidases contiennent seulement 7 brins β . Enfin, la famille 25 est sous forme tonneau *TIM* mais avec un brin antiparallèle, elle comporte seulement 5 hélices donc une structure (β/α)₅(β)₃ est obtenue¹⁸³.



Figure 25 : Structures des glycosidases : (a)Tonneau (β/α)₈, (b) β -Jelly roll, (c) 5-fold β -propeller, (d)6-fold β -propeller, (e) 7-fold β -propeller, (f) Tonneau (α/α)₆, (g) Tonneau (α/α)₇, (h) hélices- β droites, (i) $\alpha+\beta$.

5.2 β -Jelly roll

Le repliement *Jelly roll,* également nommé *Swiss roll,* est un type de repliement formé par 4 paires de brins β antiparallèles (Fig 25b). Il a été montré par Jane S. Richardson en 1981.

Chez les glycosidases, les clans B et C utilisent ce repliement. Une caractéristique importante des *Jelly roll* est que le tonneau est fermé des deux côtés par des boucles « longues ». Le site actif est situé dans une crevasse de la surface, plus précisément, les résidus catalytiques d'acide sont situés sur le brin 9 pour le clan B et sur les brins 4 et 6 pour le clan C. Les glycosidases de la famille 45 ont un tonneau à 6 brins similaire à un *Jelly roll*, mais la différence

est que ce tonneau est composé à la fois de brins parallèles et antiparallèles. Les glycosidases de la famille 54 ont une structure similaire à celle du clan B.

5.3 β -propeller

Le β -propeller se caractérise par l'arrangement de 4 à 8 feuillets β sous forme de lame symétrique autour d'un axe central selon une configuration toroïdale. Les feuilles β sont normalement composées de 4 feuillets β antiparallèles tords de sorte à ce que les premiers brins et les quatrièmes brins soient presque perpendiculaires²⁰³. Les sites actifs sont situés dans une crevasse sur la face du feuillets β .

Chez les glycosidases, les clans F et J avec les famillets 130, 137 et 143 présentent cette structure avec 5 feuillets β (Fig 25c) et le clan E présente 6 feuillets β (Fig 25d). Les familles 74 et 145 présentent 7 feuillets β (Fig 25e).

5.4 Tonneau $(\alpha/\alpha)_6$

La structure tonneau $(\alpha/\alpha)_6$ se caractérise par 6 hélices α parallèles qui forment un fût interne et flanquées par 6 hélices externes (Fig 25f). Mais la famille 47 des glycosidases présentent une variation : un cylindre $(\alpha/\alpha)_7$ dans lequel 7 épingles à cheveux en hélice α forment les structures internes et les structures externes du cylindre (Fig 25g). Les positions des sites catalytiques des glycosidases n'ont pas de lien directe avec ces structures¹⁸³. Chez les glycosidases, les clans G, L, M, O, P et Q avec les familles 88, 95, 105, 126, 142, 144, et 162 présentent cette structure.

5.5 Hélice-β

L'hélice- β est un repliement qui forme par l'association de brins β parallèles et s'enroule comme un ressort hélicoïdal. Les structures les plus récentes obtenues par cristallographie ont montré l'existence d'hélices- β droites ainsi que d'hélices β gauches²⁰⁴.

Le clan N avec les familles 82, 90, 91, 136 et 141 présente ce repliement en hélices- β droites (Fig 25h). Les familles et les clans peuvent être distingués par le nombre de tours de brins β , le nombre de feuilles β formées et l'architecture du site actif¹⁸³.

5.6 **α**+β

La structure α + β est caractérisé par deux hélices et une feuille β qui contient 3 brins (Fig 25i). Le clan I avec les familles 19, 22 et 23 et certaines transglycosylases lytiques possèdent cette structure.

6. Topologie des sites actifs

Bien que de nombreux repliements ont été observés chez les glycosidases, seules quatre topologies de sites actifs ont été décrites (Fig 26). La topologie de sites actifs est dépendant de la fonction enzymatique mais indépendamment du type de mécanisme rétention ou inversion²⁰⁵.

6.1 Poche

La topologie de la poche se présente chez les *exo*-glycosidases comme la β -glucosidase, la β galactosidase, la neuraminidase et la β -amylase. C'est une topologie optimale pour couper les monosaccharides de l'extrémité non réductrice des oligo- et polysaccharides (Fig 26a). Les centres actifs ont été enfouis dans les poches généralement.

6.2 Crevasse ouverte ou sillon

La topologie de type crevasse ouverte ou sillon se présente chez les *endo*-glycosidases comme l'*endo*-cellulase, le lysozyme, l' α -amylase et la xylanase. C'est une structure ouverte qui permet de fixer plusieurs unités saccharidiques pour des substrats polymériques (Fig 26b).

6.3 Niche

La topologie de type niche a été observée chez certaines arabinanases GH43 libérant des trisaccharides. Dans ce cas, le site actif est situé dans un sillon bloqué d'une extrémité, et il permet d'avoir une activité *exo* ou *endo*-processive²⁰⁵ (Fig 26d). Parfois, la niche peut permettre une ingénierie plus facile des protéines, ce qui modifie le profil de produit de l'enzyme avec peu de changement sur l'efficacité catalytique.

6.4 Tunnel

La topologie de type tunnel a été observée dans le cas des cellobiohydrolases I et II de *Trichoderma reesei*²⁰⁵ par exemple. Cette topologie découle de la crevasse lorsque la protéine développe de longues boucles qui recouvrent partiellement cette dernière.

Elle empêche la sortie de la chaîne polysaccharidique, facilitant ainsi une réaction processive dans laquelle de multiples événements catalytiques se produisent sans libération du polysaccharide²⁰⁵ (Fig 26d).



Figure 26 : Topologies des sites actifs des glycosidases : (a) poche, (b) crevasse, (c) niche, (d) tunnel. Les positions des résidus catalytiques sont présentées en rouge²⁰⁵.

7. Les glycosidases de Mycobacterium tuberculosis

A ce jour, jusqu'à 40 % des gènes de *M. tuberculosis* codant pour une protéine n'ont pas de fonction attribuée ou celle-ci est considérée comme hypothétique²⁰⁶. Un grand nombre des glycosides hydrolases (GH) appartiennent à cette catégorie.

D'après la classification CAZy, 31 gènes codant pour des glycosidases provenant de 14 familles sont présents dans le génome de *M. tuberculosis* H37Rv (Tableau III). Ils répartissent en 4 groupes différents : enzymes ciblant les α -glucanes, les β -glycanases, les α -mannosidases et enzymes impliquées dans le remodelage du peptidoglycane. Cependant, seulement 13 gènes sur 30 ont été caractérisés biochimiquement.

Nom	GHs Familles	Nom du gène	Peptide signal	Orthologue chez <i>M.</i> SmeamatisMC ² 155				
Enzymes ciblant les α-glucanes								
Rv1327c*	GH13 – 3	glgE	Non	MSMEG_4916				
Rv1326c*	GH13 – 9	glgB	Non	MSMEG_4918				
Rv1562c*	GH13 – 10	treZ	Non	MSMEG_3184				
Rv1564c	GH13 – 11	glgX	Non	MSMEG_3186				
				MSMEG_6507				
Rv0126*	GH13 – 16	treS	Non	MSMEG_6515				
Rv1563c*	GH13 – 26	treY	Non	MSMEG_3192				
Rv2471	GH13 – 30	aglA	Non	MSMEG_4696				
Rv2402*	GH15	Rv2402	Non	MSMEG_4535				
Rv3031	GH57	Rv3031	Non	MSMEG_2349				
Rv2006	GH65	otsB	Non	-				
Rv3401	GH65	Rv3401	Non	MSMEG_1608				
				MSMEG_3954				
Rv1781c	GH77	malQ	Non	MSMEG_3673				
β-glycanases								
Rv0186	GH3	bglS	Non	MSMEG_2556				
D. 2000	CUE 12	D. 2000	0	MSMEG_5144				
KV3090	GH2 – 13	RV3090	Our	MSMEG_5877 MSMEG_6581				
Rv0062*	GH6	celA1	Oui	MSMEG 6752				
Rv1090*	GH12	celA2b	Non	-				
Rv0315*	GH16	Rv0315	Oui	MSMEI 0629				
	Enzymes	impliquées dans le re	modelage du peptido	glycane				
Rv0237*	GH3	lpqI	Oui	MSMEG 0361				
Rv2450c*	GH23	rpfE	Oui	-				
Rv1009	GH23	rpfB	Oui	MSMEG 5439				
Rv1022	GH23	lpaU	Oui					
Rv1230c	GH23	Rv1230c	Non	MSMEG 5067				
Rv0867c	GH23	rpfA	Oui	-				
Rv1884c	GH23	rpfC	Oui	_				
Rv2389c	GH23	rnfD	Non	_				
Rv3896c	GH23	Rv3896c	Non	_				
Rv2525c	GHnc	Rv2525	Qui	MSMEG 6815				
	cc	a-manno	sidases					
Rv0648*	GH38	Rv0648	Non	MSMEG 1361				
Rv0365c	GH76	Rv0365c	Non	MSMEG_1301 MSMEG_0740				
Rv0584	GH02	Rv0581						
Rv0304	GHpc	RyD20Ar	Non	_				
11003340	Unit	11003940	NOT	-				

Tableau IIILes GH chez *Mycobacterium tuberculosis* H37Rv selon la base de données CAZy (consulté le 20 Juin2019). On retrouve les 31 gènes avec leurs familles de GH, la présence ou non du peptide signal (c'est-à-dire protéine
sécrétée ou non) et ses orthologues chez *M. Smegmatis* MC² 155²⁰⁶.

* : gènes caractérisés.

7.1 Les α -glucanes

Le tréhalose est un disaccharide non réducteur constitué de deux résidus de glucose liés par une liaison α -(1 \rightarrow 1). Il se trouve dans les insectes, les plantes et de nombreux microorganismes, mais il est absent chez les mammifères. C'est une réserve de glucides principale, mais aussi un agent protecteur des protéines et des membranes contre de nombreux types de stress. Le tréhalose s'implique dans la biosynthèse de nombreux composants essentiels de la paroi cellulaire mycobactérienne, tels que les TMM et TDM décrits précédemment. Chez les *M. smegmatis*, les taux de tréhalose intracellulaire sont compris entre 1,5 et 3,0 % de la masse sèche totale des cellules²⁰⁶.

La plupart des enzymes codées par les 13 gènes caractérisés sont des glycosidases impliquées dans la biosynthèse du tréhalose (Fig 27). Le tréhalose peut être produit à partir du glycogène, qui est un polysaccharide composé entièrement de résidus de glucopyranoside avec des malto-oligosaccharides linéaires liés en α -(1 \rightarrow 4), les maltooligosaccharides se liés entre elles par des points de ramification en liaison α -(1 \rightarrow 6). La liaison terminale α -(1 \rightarrow 4) à l'extrémité réductrice du glycogène est convertie en liaison α -(1 \rightarrow 1) par une maltooligosyltréhalose synthase, TreY, de la famille GH13 appelée, pour donner du maltooligosyltréhalose. Le maltooligosyltréhalose est ensuite hydrolysé par une maltooligosyltréhalose, TreZ, de la famille GH13 pour donner du tréhalose²⁰⁶.



Figure 27 : Voies de biosynthèses et de dégradation du tréhalose chez *M. tuberculosis*, adaptée de Wyk et al. 2017²⁰⁶.

Le tréhalose peut aussi se former à partir du maltose. Pour cela, la liaison α -(1→4) est convertie en α -(1→1) par la maltose α -D-glucosylmutase TreS de la famille GH12²⁰⁷.

GlgE et GlgB sont deux enzymes essentielles à la synthèse du glycogène. GlgE est une maltosyltransférase de la famille GH13, qui utilisé l' α -maltose-1-phosphate comme substrat pour former des liaisons α -(1 \rightarrow 4) de la chaîne α -glucane²⁰⁸. GlgB catalyserait le clivage d'une liaison glucosidique α -(1 \rightarrow 4) et le transfert ultérieur d'oligosaccharide clivé pour la formation de liaisons α -(1 \rightarrow 6) au sein du glycogène linéaire et serait donc une enzyme de ramification du glycogène²⁰⁹.

M. tuberculosis possède également une α, α -tréhalase de la famille GH15, Rv2402. Elle catalyse l'hydrolyse du tréhalose en deux monomères de glucose, qui permet ainsi de recycler du tréhalose²¹⁰.

7.2 Les β-glycanases

Curieusement, *M. tuberculosis* possède plusieurs gènes codant pour des protéines qui pourraient être impliquées dans l'hydrolyse des polysaccharides, tels que la lignocellulose et

la cellulose. CelA1 est une endoglucanase de la famille GH6 qui est capable d'hydrolyser des substrats cellulosiques²¹¹. CelA2b est une cellulase de la famille GH12. Sachant que jusqu'à présent, aucune mycobactérie n'a été montré comme bactérie cellulolytique, c'est-à-dire qui peut utiliser de la cellulose naturelle comme source glucidique²¹².

7.3 Enzymes impliquées dans le remodelage du peptidoglycane

Les enzymes impliquées dans le remodelage du peptidoglycane, notamment l'enzyme *Resuscitation-promoting factor* (Rpf), ainsi que le LpqI vont être décrites plus précisément dans le chapitre suivant. Les enzymes RpfA-E ont été classées dans la famille GH23 et l'enzyme LpqI a été classée dans la famille GH3²¹³.

Bien que de nombreuses études *in vitro* et *in vivo* ont été réalisées concernant l'effet des Rpf sur la physiologie et la virulence des mycobactéries, leur caractérisation biochimique est encore loin d'être achevée. Il a été démontré que RpfB et RpfE interagissaient avec la peptidoglycane endopeptidase RipA, ce résultat démontrant ainsi que les Rpfs n'agissent pas seules²¹⁴. De plus, RpfE est également capable de diminuer fortement l'activité hydrolytique^{215,216}.

7.4 Les α -mannosidases

Chez *M. tuberculosis*, la mannosylation (c'est-à-dire l'ajout d'unités de résidu mannose sur des protéines et des lipides) joue un rôle important dans la biologie de la paroi cellulaire mycobactérienne. La mannosylation est effectuée par une grande gamme de mannosyltransférases²¹⁷. Une grande partie de ces enzymes sont essentielles à la biogenèse et à l'intégrité de l'enveloppe cellulaire, grâce à la biosynthèse de composants associés à la paroi cellulaire, tels que PIMs, LM et LAM.

Trois gènes ont été supposés coder pour des enzymes avec une activité α -mannosidasique, Rv0648, Rv0365c et Rv0584, en se basant sur le fait qu'elles appartiennent à des familles GH (GH38, GH76 et GH92 respectivement), qui ciblent uniquement les substrats contenant du mannose²⁰⁶. Parmi ces trois gènes, il a été démontré expérimentalement que seul le Rv0648 agissait comme une a-mannosidase pouvant hydrolyser des substrats synthétiques de type aryl- α -mannoside²¹⁸.

8. Conclusion du chapitre

Pour conclure, les glycosidases sont des enzymes omniprésentes dans *a priori* tous les organismes vivants. Dans le domaine thérapeutique, une attention toute particulière s'est portée sur les glycosidases ces vingtaines années. Plusieurs inhibiteurs ont été conçus dans le cadre du cancer, des infections virales, des maladies neurodégénératives, des maladies autoimmunes et du diabète, par exemple, l'inhibiteurs N-butyl-deoxynojirimycine (NB-DNJ) est utilisé pour la maladie de Gaucher et le GLYSET® est utilisé pour la traitement de diabète 2²¹⁹. Malheureusement très peu d'attention est portée aux glycosidases mycobactériennes. En raison de la structure unique de la paroi et surtout de l'architecture lipidique complexe des mycobactéries, les recherches thérapeutiques se sont principalement focalisées sur le métabolisme lipidique plutôt que glucidique. Cependant, de nombreuses glycosidases impliquées dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse de la paroi, particulièrement les enzymes impliqués dans la biosynthèse du tréhalose et les Rpfs peuvent jouer un rôle crucial dans la virulence, la persistance et la biologie de *M. tuberculosis*, et peuvent être prises en compte en tant que cibles thérapeutiques²⁰⁶.

IV. Remodelage de la paroi

La paroi de *M. tuberculosis* joue un rôle crucial dans les interactions hôte/pathogène et dans l'immunomodulation des défenses de l'hôte. Dans ce cadre, un remodelage de cette paroi apparaît comme un mécanisme essentiel que la bactérie va utiliser pendant sa croissance et sa division, mais aussi pour s'adapter aux changements externes (c'est-à-dire l'environnement de l'hôte) et détourner les défenses de l'hôte pendant l'infection.



Figure 28 : Le remodelage de la paroi de *M. tuberculosis* est lié à la survie de la bactérie pendant l'infection, La prolifération *in vivo* et l'utilisation des sources de carbone de l'hôte pendant la croissance dans les macrophages conduit à une augmentation de la production des acides mycoliques des lipides de la paroi comme le phthiocérol-dimycocérosate (PDIM). De même, le PG est remodelé avec le maintien des ponts transpeptidiques 3-3 qui favorisent l'infection chronique et la résistance aux antibiotiques. En condition de croissance lente ou en état non prolifératif, la paroi s'épaissit probablement suite à une synthèse accrue de ses composants. *M. tuberculosis* diminue aussi la quantité d'acides mycoliques immunostimulants pour éviter sa détection par le système immunitaire de l'hôte et augmente la proportion de ponts transpeptidiques dans le PG pour le rendre plus rigide. Schéma adapté de Kieser et Rubin 2014²²⁰.

Le remodelage est considéré comme des modifications structurales dynamiques se produisant après la biosynthèse de la macromolécule concernée, dans le but de répondre à une problématique physiologique. Pour cela, les trois éléments clés de la paroi, PG/arabinogalactane/acides mycoliques, peuvent subir des modifications diverses : substitution par des groupements divers du PG et de l'arabinogalactane, nature des liaisons entre les monomères du PG, variations de la structure des acides mycoliques.... La combinaison de ces modifications induit des modifications physico-chimiques de la paroi qui suggère qu'elles peuvent faire partie d'une réponse adaptative de la bactérie notamment lors de l'infection (Fig 28). L'étude a déjà démontré que la composition de la paroi est modifiée au cours de l'infection²²⁰.

1. Dégradation, remodelage et recyclage du peptidoglycane

Le peptidoglycane est une structure dynamique constamment synthétisée et remodelée (Fig 17). Les PG hydrolases forment un grand groupe d'enzymes comprenant des glycosidases, des amidases, des endopeptidases et des carboxypeptidases, qui sont capables de digérer complètement le polymère PG en fragments solubles (Fig 16, partie dégradation). En raison de leur capacité sélective à cliver le PG, les hydrolases sont impliquées dans de nombreux processus cellulaires comprenant la croissance cellulaire, la maturation du PG, le repliement du PG, la division cellulaire, la réanimation ou la pathogénicité¹²⁷.

Les glycosidases telles que les *N*-acétyl- β -D-glucosaminidases et les *N*-acétyl- β -dmuramidases clivent le squelette du polysaccharide de PG, tandis que les *N*-acétylmuramyl-l-Ala amidases hydrolysent la liaison amide entre la MurNAc et le L-Ala de la chaîne peptidique. Le peptidoglycane des mycobactéries est largement modifié pour interagir avec l'environnement de l'hôte et de s'adapter. Le nombre de liaisons (3 \rightarrow 3) entre les résidus de *m*-DAP augmente lorsque *M. tuberculosis* entre en dormance, due à un réarrangement de PG par des L,D-transpeptidases (Ldt) non classiques²²¹. Les transpeptidases non canoniques LdtA et LdtB catalysent la réticulation prédominante du peptide (3 \rightarrow 3) dans le peptidoglycane. L'expression de LdtB culmine en fin de phase exponentielle et en phase stationnaire, alors que l'expression de LdtA n'augmente qu'en fin de phase stationnaire. Les mutants LdtB ne peuvent pas persister et leur nombre diminue au cours de l'infection²²². Ces résultats indiquent qu'un peptidoglycane riche en ponts transpeptidiques (3 \rightarrow 3) semble être nécessaire pour le maintien de la bactérie chez l'hôte et sa capacité à détourner son système immunitaire²²⁰.

En plus, plusieurs étapes de la biosynthèse du peptidoglycane sont régulées positivement (*upregulated*) *in vivo*, comme les gènes responsables de la production de précurseurs (tels que *murE* et *ddl*) ²²³. De plus, PBP1 est nécessaire à la survie de *M. tuberculosis* au cours de

l'hypoxie, à l'infection et à la croissance de *M. smegmatis* dans des conditions de réplication réduite^{222,224-226}. En outre, le remodelage du peptidoglycane par des PBPs de faible poids moléculaire, tels que DacB1 et DacB2, qui hydrolysent les résidus terminaux D-alanine des sous-unités de l'acide peptidyl-N-acétyl muramique, est nécessaire à la survie chez l'hôte. Le remodelage du peptidoglycane par DacB1 peut maintenir les liaisons (3→3) indispensables à la persistance *in vivo*, car l'élimination de la D-alanine terminale inhibe la réticulation (3→4) ²²⁰. DacB2 de *M. tuberculosis* est une D,D-carboxypeptidase qui élimine les résidus D-Ala en position 5 des pentapeptides²²⁷. La surexpression de *dacB2* entraîne une modification de la morphologie et une formation de biofilm défectueuse, tandis que sa suppression entraîne une augmentation de la survie intracellulaire du mutant dans des cellules THP-1²²⁸. Il a également été signalé que DacB2 possédait une activité de D,D-endopeptidase²²⁹. Donc il peut affecter la morphologie des cellules en hydrolysant la liaison (3→4), fournissant donc des substrats tétrapeptidiques pour les L,D-transpeptidases, en éliminant le résidu D-Ala terminal.

Il a aussi été suggéré par Votyakova²³⁰ et Mukamolova²³¹⁻²³³ que la présence d'un facteur favoriserait la récupération pour les bactéries à la dormante, le Resuscitation Promoting Factor (Rpf) chez Micrococcus luteus. Il présente une activité glycosidasique, qui clive la chaîne glycannique du PG, et ainsi être responsable de la libération de fragments du PG. Cette découverte a permis d'identifier cinq Rpfs (RpfA à RpfE) chez Mycobacterium spp.²³⁴⁻²³⁶. Ces résultats ont permis de bien avancer non seulement dans la compréhension du remodelage du PG pendant la formation des cellules filles, mais également dans les processus à l'origine de la dormance et de la réactivation du bacille pour la tuberculose latente/active. Mycobacterial RipA (Rv1477) et RipB (Rv1478) sont des D-Glu-mDAP endopeptidases qui clivent entre le D-glutamate et le *m*DAP de la chaîne peptidique²³⁷. Les *N*-acétylmuramyl-L-Ala amidases hydrolysent la liaison entre le résidu MurNAc et le premier L-Ala de la chaîne peptidique. Elles sont aussi impliquées dans la dégradation et le renouvellement du PG, la séparation cellulaire, la résistance aux antibiotiques et la formation de spores¹²⁷. Les amidases identifiées chez *M. tuberculosis* sont CwlM (*Rv3915*) et Rv3717, elles jouent un rôle régulateur dans la synthèse et le maintien du PG^{238,239}. Pris ensemble, ces résultats suggèrent que les cellules doivent subir un remodelage du peptidoglycane pour s'adapter à divers stress.

Le recyclage du PG a été bien étudié chez *E. coli, Pseudomonas aeruginosa* et *Bacillus subtilis*²¹³. La voie de recyclage du PG (Fig 16, partie recyclage) entraîne généralement

la dégradation du polymère PG en ses constituants monomères de GlcNAc-(1→4)-1,6anhydro-MurNAc-peptide, des dissacharides de GlcNAc-(1→4)-1,6-anhydro-MurNAc et peptides^{127,240}. Il a été montré très récemment l'existence d'une lipoprotéine LpqI (*Rv0237*) chez *M. tuberculosis* et de *M. bovis BCG*. LpqI est une exo-β-N-acétylglucosaminidase qui permet de cliver des fragments du PG. LpqI peut hydrolyser les disaccharides GlcpNAc-MurpNAc quand la chaîne peptidique est retirée, ce qui permet la libération de MurpNAc libre. La MurpNAc est ensuite suivie d'un clivage de D-lactyl-éther pour libérer du lactate, qui peut être utilisé par la cellule dans des conditions d'aérobies²¹³.

Les monosaccharides libérés sont finalement phosphorylés et le MurNAc-6-phosphate est converti en glucosamine-6-phosphate par les enzymes cytoplasmiques MurQ et NagA (*Rv3332*)^{213,241}. En même temps, les chaînes peptidiques sont dégradées en composants plus petits par la murine peptide ligase Mpl (*Rv3712*) et sont de nouveau transférés dans la biogenèse du PG¹²⁷.

2. Les Acides mycoliques

Les acides mycoliques sont aussi particulièrement touchés par ces modifications (Fig 17). Les trois formes d'acides mycoliques (c'est-à-dire les acides α -mycoliques, les acides méthoxy-mycoliques et les acides céto-mycoliques) sont synthétisés avec différentes longueurs de chaîne acyl, différents nombres de saturation et sont également modifiés par la cyclopropanation comme mentionné précédemment. Les bacilles isolés des patients atteints d'une infection chronique ne sont plus acido-résistants, ce qui suggère que des modifications structurelles et/ou de composition majeures de la paroi se produisent pendant les périodes prolongées de croissance *in vivo*²⁴². Les cellules isolées de cochon d'inde infectées présentent des différences de composition en acide mycolique par rapport aux cellules cultivées *in vitro*²⁴³. De même, l'analyse transcriptionnelle de bactéries isolées de diverses régions des poumons a révélé des changements substantiels dans l'expression des gènes de biosynthèse de la paroi cellulaire et des protéines membranaires par rapport aux cultures *in vitro*²²³, notamment la régulation positive (*upregulation*) des gènes de synthèse des lipides. Par exemple, *pcaA*, *desA1* et *desA3*, qui sont impliqués dans la cyclopropanation et la désaturation des acides mycoliques, sont régulés positivement et sont nécessaires à la virulence *in vivo*²⁴⁴.

De plus, la longueur et l'abondance des acides mycoliques augmentent pendant la croissance de *M. tuberculosis in vivo*²⁴⁵. Le tout conduisant à une augmentation de l'épaisseur de la paroi. Ces modifications confèrent de multiples avantages à la bactérie. Tout d'abord, une paroi cellulaire plus épaisse limite le transit de molécules toxiques (tels que des antibiotiques) à travers la membrane cellulaire. En plus, les acides mycoliques peuvent manipuler le système immunitaire²⁴⁶ et absorber les radicaux oxydants¹¹³. Au cours de la croissance lente et lors des états non réplicatifs, ainsi que dans les environnements hypoxiques, *M. tuberculosis* présente une baisse de l'abondance des acides mycoliques immunostimulateurs pour éviter potentiellement la reconnaissance immunitaire.

Pendant l'hypoxie, la production de lipides diminue ainsi que la composition en lipides change. Comme beaucoup de ces lipides sont des molécules pro-inflammatoires^{244,246}, des changements dans leur abondance affecteront également les interactions entre l'agent pathogène et l'hôte. Par exemple, la suppression de *kasB*, qui est un composant du complexe acide mycolique synthase FAS-II, atténue la virulence dans une souche de *M. tuberculosis*²⁴⁷. Les souris immunocompétentes infectées par un mutant $\Delta kasB$ semblent rester saines jusqu'à 600 jours après l'infection, tandis que les souris infectées par le type sauvage ou par une souche complétée par *kasB* meurent de l'infection dans les 356 jours. La persistance des bactéries $\Delta kasB$ est probablement due à leurs chaînes d'acide mycolique raccourcies et altérées²⁴⁷ qui doivent réduire leur immunogénicité.

3. Catabolisme de l'arabinogalactane

A première vue, l'arabinogalactane semble être le constituant de l'enveloppe cellulaire qui subit le moins de modifications structurales. Pourtant, de nombreux éléments semblent indiquer que l'arabinogalactane lui aussi est capable de subir ce type de modifications après sa biosynthèse par un catabolisme spécifique. Cette hypothèse d'un catabolisme de l'arabinogalactane est premièrement apparue avec la mise en évidence, uniquement chez *M. smegmatis*, d'une activité endoarabinanasique capable de cliver la partie terminale de l'arabinane²⁴⁸. Malheureusement, cette enzyme n'a jamais été isolée et identifiée. Il a été montré par le Dr. Méry de l'équipe du Dr. Guérardel que deux types d'oligosaccharides peuvent être formés par cette endo-arabinanase (Fig 29). La première est les petits fragments de la chaîne arabinane comprend de 5 à 8 arabinofuranoses (Ara₅₋₈) et la deuxième est

d'oligosaccharides plus grandes, c'est-à-dire de 18 à 22 arabinofuranoses (Ara₁₈₋₂₂) (résultats non publiés).



Figure 29 : Représentation schématique de la chaîne arabinane de l'arabinogalactane en présentant les différents sites de coupure et les fragments formés Ara₅₋₈ et Ara₁₈₋₂₂, adapté de Jankute *et al.* 2015.

Ensuite, l'équipe du Dr. Guérardel a montré l'existence dans la paroi de *M. Bovis* BCG d'un arabinoglycérolipide, le DMAG (dimycolyl diarabinoglycérol) ou 5-O-mycolyl- β -Araf- $(1\rightarrow 2)$ -5-O-mycolyl- α -Araf- $(1\rightarrow 1)$ -Gro²⁴⁹ (Fig 30). En comparant la structure du DMAG avec la partie terminale non réductrice du mAGP, il apparaît que le DMAG pourrait tout à fait provenir de la transglycosylation de ce motif glycolipidique sur une molécule de glycérol. Une telle activité de transglycosylation est tout à fait plausible pour une arabinanase si elle agit en présence d'un accepteur potentiel – autre que l'eau - pour l'oligosaccharide libéré, soit ici le glycérol. Cette hypothèse a été confirmée récemment dans l'équipe en démontrant que l'activité arabinanasique présente chez *M. smegmatis* est capable de transférer les oligosaccharides libérés (Ara₆, Ara₇, Ara₈, Ara₁₇, Ara₁₈, Ara₁₉ et Ara₂₀) sur le glycérol (résultats non publiés).



Figure 30 : Le motif 5-*O*-mycolyl- β -Ara*f*-(1 \rightarrow 2)-5-*O*-mycolyl- α -Ara*f* est retrouvé à l'extrémité terminale non réductrice du mAGP et du DMAG²⁴⁹. L'identification du DMAG suggère fortement qu'existence d'une ou plusieurs endo-arabinanase(s), qui permets de réaliser cette transglycosylation en transferant les fragments arabinanes terminales du complexe mAGP sur un accepteur glycérol.

Pour en finir avec les modifications structurales que peut subir la paroi mycobactérienne chez un hôte, il a été observé aussi une augmentation du nombre d'acides mycoliques et de chaînes d'arabinogalactane par rapport au PG lors d'une culture *in vivo*¹⁰⁴. Une dernière preuve de l'importance de la structure de l'arabinogalactane sur la physiologie de la mycobactérie, notamment lors de l'infection, est le fait que l'absence de l'unique résidu de galactosamine des chaînes d'arabinane entraîne chez les cellules dendritiques une expression accrue des marqueurs de maturation, une augmentation de la sécrétion de l'IL10 et une activation de NFkB par la voie TLR2²⁵⁰. L'hypothèse présentée est que la présence ou non de ce résidu de galactosamine module la présentation des molécules immunomodulatrices à la surface de la mycobactérie en introduisant des charges positives qui peuvent se repousser et relaxer la structure de l'arabinogalactane, modifiant par la même la présentation des antigènes de surface. Il est alors possible d'établir un lien direct entre la structure de l'arabinogalactane et la pathogénicité de la mycobactérie.

4. Conclusion du chapitre

La paroi cellulaire des mycobactéries présente une complexité remarquable et constitue une barrière perméable contre les médicaments ainsi que contre l'attaque immunitaire de l'hôte. Dans la paroi, le complexe le plus important est le mycolyl-ArabinoGalactane-Peptidoglycane (mAGP). La structure et la biosynthèse du complexe mAGP ont bien été élucidées aujourd'hui, bien qu'il reste quelques éléments à étudier plus précisément. Cependant, la compréhension d'un remodelage ou d'un catabolisme du mAGP sont plus limitées, surtout pour la pièce centrale du complexe mAGP, l'arabinogalactane, sachant que le remodelage de la paroi permet aux mycobactéries de s'adapter à leur environnement. Lorsque l'on énumère toutes les modifications subies par le PG et les acides mycoliques, tout particulièrement lors de l'infection, il apparait surprenant voire illogique que l'arabinogalactane ne soit pas aussi concerné par ce remodelage. Mais malheureusement, le remodelage de l'arabinogalactane est un sujet qui n'a pas encore été réellement abordé car ses acteurs potentiels (c'est-à-dire glycosidases, ...) n'ont pas encore été identifiés. C'est pour ça que je me suis intéressé pendant ma thèse à mettre en évidence le catabolisme de l'arabinogalactane en cherchant à identifier des glycosidases pouvant y être associé car ce sont elles qui seront vraisemblablement les molécules effectrices principales de ce phénomène.
Partie II : Objectifs

Chez M. tuberculosis, les modifications structurales de la paroi sont des événements essentiels que la bactérie utilise pendant sa croissance et sa division mais aussi pour s'adapter aux changements externes (c'est-à-dire l'environnement de l'hôte) et détourner les défenses de l'hôte pendant l'infection. Pour cela, les trois éléments clés de la paroi, peptidoglycane/arabinogalactane/acide mycoliques, peuvent subir des modifications diverses : substitutions sur le peptidoglycane et l'arabinogalactane, nature des liaisons entre les monomères du peptidoglycane, variations de la structure des acides mycoliques, ... Les acides mycoliques sont particulièrement touchés par ces modifications (surexpression des gènes liés à la synthèse des lipides, modification de la nature et de la longueur des chaînes, ...), le tout conduisant à une augmentation de l'épaisseur de la paroi et donc à une restriction accrue du passage de molécules. Le peptidoglycane est lui aussi modifié lors du passage dans l'hôte. Les gènes liés à plusieurs étapes de la biosynthèse du peptidoglycane sont surexprimés comme, par exemple, *LdtB* qui catalyse la formation des ponts peptidiques 3-3 nécessaires au le maintien de la bactérie chez l'hôte et à sa capacité à détourner le système immunitaire²²⁰. Lorsque l'on énumère toutes les modifications subies par le peptidoglycane et les acides mycoliques, tout particulièrement lors de l'infection, il apparait surprenant voire illogique que l'arabinogalactane ne soit pas aussi concerné par ces modifications. Pourtant, de nombreux éléments semblent indiquer que l'arabinogalactane lui aussi est capable de subir des modifications de sa structure après sa biosynthèse. Par exemple, une activité arabinanasique capable de cliver la chaîne arabinane a été observée chez *M. smegmatis*²⁴⁸, mais malheureusement l'enzyme responsable n'a jamais été identifiée. Ensuite, il y a l'identification du DMAG (dimycolyl diarabinoglycérol) dans la paroi de *M. Bovis* BCG, dont la structure peut être vue comme la transglycosylation de l'extrémité non réductrice de l'arabinogalactane (avec ses acides mycoliques) sur une molécule de glycérol²⁴⁹, transfert qui peut tout à fait être réalisé par une glycosidase utilisant le glycérol comme accepteur. Ces premières observations impliquent l'existence de glycosidases agissant sur l'arabinogalactane et pointent du doigt l'existence d'un catabolisme spécifique de ce polysaccharide

Pris ensemble, ces différents éléments montrent que des modifications structurales de la paroi cellulaire, que l'on peut présenter comme un remodelage, permettent aux mycobactéries de s'adapter à leur environnement. Cependant, si pour le peptidoglycane et les acides mycoliques, les observations expérimentales attestent de ce phénomène, le cas de l'arabinogalactane est un sujet qui n'a pas encore été réellement abordé.

112

Dans le groupe du Dr. Yann Guérardel (Glycobiologie des Interactions Hôte-Pathogène, UGSF-UMR8576), la recherche d'un catabolisme de l'arabinogalactane des mycobactéries est soutenue, depuis 2015, par l'ANR *MyCat : Le catabolisme de la paroi mycobactérienne : vers le développement de nouveaux inhibiteurs.*

L'objectif de cette thèse consistait donc à prouver l'existence de ce catabolisme centré sur l'arabinogalactane. Pour cela, il a été choisi de commencer par identifier des glycosidases spécifiques de l'arabinogalactane des mycobactéries. Une approche employée a consisté à rechercher dans le génome de *M. tuberculosis* et *M. smegmatis* des gènes pouvant coder pour les arabinanases et/ou galactosidases, en se basant sur des homologies de séquences avec des arabinanases. Cette recherche est basé sur la base de la similarité connue entre les activités D-galactosidase et L-arabinase²⁵¹. Chez *M. tuberculosis*, trois gènes ont été retenus : Rv0186 (CAB09737.1), Rv0237 (CAA17329.1) et Rv3096 (CAB08388.1). Ces trois gènes ont ensuite été clonés dans E. coli par l'équipe du Dr. Laurent Kremer (FRE3689, Montpellier). Mon travail sur ce projet a commencé par l'expression dans E. coli, la purification et la détermination de l'activité des trois protéines concernées. Seule Rv3096 (nommé GlfH1 dans la suite du manuscrit) a montré une activité glycosidasique envers l'arabinogalactane et il ne s'agissait pas d'une arabinanase mais d'une exogalactofuranohydrolase qui dégrade la chaîne galactane de l'arabinogalactane de façon récurrente en libérant du galactose libre. Les travaux ont donc principalement porté sur l'expression, la purification et la caractérisation du mode d'action de GlfH1. Ce travail sur GlfH1 permet d'établir les fondations d'un catabolisme de l'arabinogalactane des mycobactéries.

Ce manuscrit sera décrit en quatre parties suivantes : la première partie présentera la découverte de la protéine GlfH1, cette partie permettra d'avoir une vision globale de l'identification bioinformatique de cette galactosidase jusqu'à l'obtention de la protéine GlfH1. La deuxième partie contiendra la caractérisation de GlfH1, elle nous permettra de connaître les caractères enzymatiques principales, mais aussi la caractérisation du site actif de GlfH1. Pour les glycosidases, la caractérisation du site actif est un critère essentiel pour la découverte ou conception d'inhibiteurs spécifiques. La troisième partie se consacrera à l'étude de la fonction de GlfH1, surtout par le biais de son activité vis-à-vis de différents substrats, synthétiques ou naturels. Et la quatrième partie présentera l'implication de GlfH1 dans un modèle d'infection chez amibe.

Partie III : Résultats et discussion

I. Découverte de la protéine GlfH1

1. Mise en évidence d'une nouvelle galactosidase chez *M*. *tuberculosis*

1.1 Identification de GlfH1

L'idée que les mycobactéries possèdent un ensemble de glycosidases spécifiques destinées agissant sur les polysaccharides endogènes de la paroi cellulaire a été introduite par l'observation initiale d'une activité arabinasique chez *M. smegmatis,* qui est capable de dégrader de l'arabinogalactane²⁴⁸. Par conséquent, l'équipe du Dr. Guérardel a entrepris de rechercher des enzymes potentielles capables de dégrader l'arabinogalactane mycobactérien et introduire la notion de catabolisme de ce polysaccharide. Ce sujet a donc été commencé en recherchant d'abord des galactosidases chez *M. tuberculosis* capables de couper la chaîne galactane de l'arabinogalactane présent dans leur enveloppe cellulaire. Cependant, notre recherche initiale dans la base de données CAZy n'a pas permis d'identifier directement une galactofuranosidase parmi les 31 glycosidases prédites dans le génome de *M. tuberculosis*.



Figure 31 : Structure du L-Arabinofuranose et du D-Galactofuranose.

Sur la base de la similarité connue entre les activités des D-galactosidases et des Larabinases²⁵¹ mais aussi de la similarité structurelle entre le L-Arabinofuranose et le D-Galactofuranose (Fig 31), j'ai utilisé les α -L-arabinofuranosidases comme patron pour la recherche de β -D-galactofuranosidases (EC 3.2.1.146) potentielles. À ce jour, les α -Larabinofuranosidases (EC 3.2.1.55), les endo- β -xylanases (EC 3.2.1.-), les arabinanases (EC 3.2.1.99), les exo- α -L-1,5-arabinanases (EC 3.2.1.-) et les β -L-arabinofuranosidases (EC 3.2.1.-) ont bien été caractérisées dans de nombreuses familles GH dans la banque de glyco-enzymes CAZy.

Parmi toutes ces familles, seule la famille GH3 possède deux candidats, Rv0186 (NP_214700.1) et Rv0237/LpqI (YP_177702.1)²¹³ qui présentent des ORF annotés dans le génome de *M. tuberculosis* H37Rv. Un troisième candidat pour une potentielle galactofuranosidase, Rv3096/GlfH1 (NP_217612.1), est aussi retrouvé dans la famille GH5. En effet, alors que la famille GH5 contient principalement des β -xylanases (EC 3.2.1.-), celles-ci partagent souvent des similarités de séquence et d'activités avec les α -L-arabinofuranosidases²⁵²⁻²⁵⁴. J'ai d'abord produit ces trois protéines candidates sous forme de protéines recombinantes dans *E. coli* en vue de leur isolement. Cependant, la protéine Rv0186 n'a pas réussi à être exprimée et purifiée, la protéine Rv0237 n'a pas montrée d'activité galactofuranosidase vis-à-vis de l'arabinogalactane. Par contre, la GlfH1 a montré une activité galactofuranosidase vis-à-vis de l'arabinogalactane de *M. bovis* BCG (Fig 32). C'est donc cette protéine qui a été retenue pour la suite de ce travail.



Figure 32 : Chromatogramme représentant l'hydrolyse de l'arabinogalactane de *M. bovis BCG* par GlfH1 (noir) et analysé par HPAEC.

Le gène *glfH1* est composé de 1140 pb codant pour une protéine de 379 résidus. Selon le serveur SignalP-5.0²⁵⁵, cette séquence contient un peptide signal de 27 résidus d'acides aminés. Le poids moléculaire de GlfH1 sans le peptide signal a est de 40 382 Da. L'analyse de séquence par SSDB a montré qu'il contient un motif Pfam Cellulase de la famille GH5 couvrant les résidus 150 à 282 (Fig 33) et qui concorde avec son classement dans la banque CAZy.

GlfH1



Figure 33 : L'annotation de la séquence GlfH1 montre la présence d'un peptide signal (1-27) et d'un motif Pfam Cellulase (150-282).

GlfH1 a donc été classée dans la famille GH5_13 dans la base de données CAZy. De manière surprenante, cette famille ne contient que des glycosidases qui utilisent des polysaccharides tels que du glucose ou le mannose comme substrats. Ce sont principalement des mannosidases (EC 3.2.1.25), des mannanases (3.2.1.78, 3.2.1.100), des cellulases (3.2.1.4, 3.2.1.91), des xylanases (EC 3.2.1.8), des glucosidases/glucanases (3.2.1.21, 3.2.1.58, 3.2.1.74, 3.2.1.75, 3.2.1.104, 3.2.1.54) mais une seule catégorie de galactanases (3.2.1.164). Toutes les enzymes de cette famille CAZy agissent par rétention de configuration.

Un alignement multiple des séquences de 13 glycosidases mycobactériennes différentes provenant de la famille GH5_13 a été généré par l'algorithme MUSCLE (Fig 34). Ces 13 séquences présentaient une grande similarité avec un maximum de 36 % dans le test de dissimilarité de Hamming (56 % des résidus sont conservés entre toutes les séquences de l'alignement) (Fig 36).

	1 10	20	30	40	50	60	70	80
AYJ05285.1	QPPRTSPQASR	SPERANRWYOA	ODWPVGANYIT	NAINQLEM	FOPDTFDPRRI	DTELGWARRNG	FNAVRVFLHI	DLLWEQ
ASW92947.1	QPPRTSPQASR	ISPE <mark>RA</mark> NRWYQA	QG <mark>WPVG</mark> ANYIT	S <mark>NAINQ</mark> LEM	IFQADTFDPRRII	DTELGWAQSNG	F N <mark>A V</mark> R V F L H I	DL <mark>LW</mark> AQ
AFC43445.1	QPPRTSPQASR	SPERANRWYQA IDVERANAWYOA	QGWPVGANYIT:	SNAINQLEM	IFQADTFDPGRII	DTELGWAQSNG	FNAVRVFLHI	
AKC40943.1	APGO	PVERANAWYOA	OGWLVGTNFIT	SNAINOILEM	FAPGTYDARRI	DSELGACRLLG	FNTVRVFLHI	
AN006525.1	ASGQ	PVERANAWYQA	QGWLVGTNFIT	SNAINQLEM	IFGPGTYDARRII	DSELGACRLLG	FN <mark>TV</mark> RVFLHI	DLLWAQ
AIU17506.1	AP <mark>G</mark> Q <mark></mark>	IPVERANAWY <u>O</u> A	QG <mark>WLVGT</mark> NF IT	S <mark>NAINQ</mark> LEM	IF SAGTYDPRRII	DSELGACRLLG	FN <mark>TV</mark> RVFLHI	DL <mark>LW</mark> AQ
AKS36268.1	CEDCD	PVERANAWY QA	OGWLVGTNFIT:	SNAINQIEM SNAINOIEM	IF SPGTYDPRRII	DSELGACRLLG	FNTVRVFLHI	
NP 217612.1	EEPGR	SADRAHRWYOA	HGWLVGANYIT	SNAINOLEM	FOPGTYDPRRI	DNELGLARFHG	FNTVRVFLHI	DLLWAO
BBA89654.1	VA <mark>G</mark> R <mark>W</mark>	ISAE <mark>RA</mark> NTWYQT	QG <mark>WIVG</mark> ANYIT	A N A I N Q L E M	IF ÕP ATYDPRRII	DR <mark>ELG</mark> LARLIG	FN <mark>SM</mark> RVFLHI	DQLWAŠ
ABL05635.1	· · · · · · · · VAGR	SAERANTWYOT	QGWIVGANYIT	ANAINQLEM	IF QPATYDPRRII	DRELGLARLIG	FN <mark>SM</mark> RVFLHI	DQLWAS
consensus>70	VAGR	I., #RAn.WYOA	aaw.VG.N%IT	SNAINOLEN	IF.p.T%DprRI	D.ELGrG	FN.VRVFLHI	DlLWag
			15					1
	90	100	110	120	130	140	150	160
AV.105285.1	DHRGEOGRIAR	VDTAARHGTKP	LEVIEDSCWDP		PREGINNSGWV	SPGAARLDDH	GVI.HTIRGV	
ASW92947.1	DHRGFQGRLARF	VDIAAHHGIKP	LFVLFDSCWDP	FPQPGPQRA	PRPGIHNSGWV	SPGAARLDDR	GYVRTLRGYV	T GVLT
AFC43445.1	DHR <mark>GFQ</mark> G RL ARF	VDIA ARHGIKP	L F V <mark>L</mark> F D S C W D P	F PR PG POPA	PRPGIHNSGWV	<mark>DSPGA</mark> ERLGDR	G Y V R T <mark>L R G Y V</mark>	V T G V L T
VEG20668.1 AKC40943 1	DRAGFONRLSOF	VGIAAROGIKP	LFVFFDSCWDP	NPQLGAORA	PTPGVHNSGWV	DSPGAHRIDDP.	RYRGMLKDY	VVGVMS
AN006525.1	DRAGFODRLSOF	VGIAAROGIKP	LFVFFDSCWDP	NPOLGAORA	PTPGVHNSGWV	SPGAHRIDDP	RYRGVLSDY	TGVMS
AIU17506.1	DRA <mark>GFQ</mark> N <mark>RL</mark> AQF	V <mark>SIA</mark> SRQ <mark>GIK</mark> P	L F V <mark>L</mark> F D S C W D P	L <mark>P</mark> K P G A Q R A	אַר <mark>אד</mark> אַר <mark>ד איז איז איז איז איז איז איז איז איז איז</mark>	Q S P G A Q R I D D P :	R <mark>Y</mark> R P V <mark>L R D Y V</mark>	V V G V M S
AKS36268.1	DRAGFONRLAOF	VSIAAROGIKP	LFV <mark>F</mark> FDSCWDP	LPRPGAORA	PTPGVHNSGWV PPAGWHNSGWV	OSPGAORIDDP:	RYRPVLHDYV	
NP 217612.1	DAPGFOTRLAOF	VAIAARYHIKP	LFVLFDSCWDP	LPRPGRORA	PRAGVHNSGWV	SPGAERLDDR	RYASTLYNY	TGVLG
BBA89654.1	DQR <mark>GFQ</mark> TRLAQF	V<mark>AIA</mark>ARHGIKP	L F V <mark>L</mark> F D S C W D P	FPKLGQ <mark>Q</mark> RA	P <mark>RP</mark> G <mark>V</mark> HNSGWV	<mark>OSPGA</mark> EHL <mark>GD</mark> P	S <mark>Y</mark> QAV <mark>L</mark> HG <mark>Y</mark> V	V T <mark>G V</mark> L N
ABL05635.1	DORGFOTRLAOF	VAIAARHGIKP	LFVLFDSCWDP	FPKLGQQRA	PTPGVHNSGWV	OSPGAEHLGDP	SYQAVLHGY	V T G V L N
consensus>70	D., GFO. RLagF	V.IAar.gIKP	LFV.FDSCWDP	.P.,G.OrA	P. pG! HNSGWV	OSPGA.rD.	Y	V.GVŚ.
		····· · ···						
	170	180	190	200	210	220	230	240
AYJ05285.1	OFRIDDRILGWD	LWNEPDNPADA	ASVERTOKLD	RVAELLPOV		LTSCVWOGRW	ADPARRSVI	GTOLD
ASW92947.1	Q F R N D D R I L G W D	LWNEPDNPADT	Y A S VE R K D K L D	LVANLLPQV	FEWARLVDPRQI	PLTSGVWHGEW	ADPAS <mark>RS</mark> VI	AGIQLD
AFC43445.1	Q F R N D D R I L G W D	LWNEPDNPA<mark>D</mark>T	Y A S V E R K D K L D I	LVANLLPQV	FEWARLVDPRQ	PLTS <mark>GVW</mark> HGEW	ADPAR <mark>RS</mark> VI	AGIQLD
VEG20668.1	QFRNDDRILGWD	LWNEPDNPAKO	YRKVERKDKID VRKVERKDKID	AVSGLLPQV	F SWARSVNPVQI	PLTSGVW0GSW	DRGSRSEM	AGFOLD
ANO06525.1	OFRNDNRVLGWD	LWNEPDNPAKO	YRKVERKDKID	AVSGLLPOV	FNWARSVNPVOI	PLTSGVWOGSW	.DRGSRSAMA	AGFOLD
AIU17506.1	QFR <mark>NDQRV</mark> LGWD	DLWNEPDNPA <mark>R</mark> Q	Y RK <mark>VER</mark> SDKLD	A <mark>V</mark> GA <mark>LLPQ</mark> V	FGWARSVNAAQ I	PLTS <mark>G</mark> VW <mark>0</mark> GSW	. ERGR <mark>RS</mark> EMA	ASF <mark>QLD</mark>
AKS36268.1	QFRNDQRVLGWD	LWNEPDNPA <mark>KQ</mark>	YRKVERSDKID	AVGSLLPQV	FGWARSVNASQI	PLTSAVWOGSW	EQGRRSEMA	ANFQLD
NP 217612.1	OFRNDDRWLGWD	DI.WNEPDNPARV	VRKVERKDKLE	RVAELLPOV	FRWARTVDPVQ		GDPGRRSTIS	SATOLD
BBA89654.1	QFRNDNRVLGWD	LWNEPDNPAKV	Y R K V E R K D K L E	RVAELLPQV	FQWAREVDPSQI	PLTSGVWQGNW	SDPGK <mark>RS</mark> TI <i>I</i>	ASIQLD
ABL05635.1	QFRNDNRVLGWD	DLWNEPDNPA <mark>K</mark> V	Y R K V E R K D K L E	RVAELLPQV	FQWAREVDPSQI	PLTS <mark>GVWQ</mark> GNW	SDPGKRSTI	ASIQLD
consensus>70	OFRND#R!LGW	LWNEPDNPAN	YrkVERKDKLEJ	.VLLPOV	FUWAREVDPSU	PLTSGVWOGNW	$ + . \sigma . RS a$	
	2							
	250	260	270	280	290	300	310	320
AYJ05285.1	NSDVITEHCYGE	PAAFEKRIADL	VPLERPILCTE	MARPLGST	VOTILPIAKRAI		GKTOTEPWI	DSWEHP
ASW92947.1	N S D V I T F H C Y G E	PAE <mark>FE</mark> R <mark>RI</mark> DQL	VPL <mark>G</mark> RPI <mark>L</mark> CTE	Y M <mark>AR</mark> PL <mark>GS</mark> I	VÕNI <mark>LPIAKR</mark> A(G <mark>VGA</mark> FN <mark>WG</mark> FVA	GKTÕT <mark>YF</mark> PWI	D S W D E P
AFC43445.1	NSDVITFHCYGE	PAAFERRIAEL	VPLGRPILCTE	YMARPLGSI	VQNILPIAKRT(GVGAFNWGFVA	GKTQTYFPWI	D S W D H P
AKC40943.1	NSDVISFHSIAG	PTEFEARINEL	APLGRPILCTE	LARDOGST	VEGVLPVAKRHI	VGAYSWGLVA	GKTOTYFPWI	DSWDKP
AN006525.1	N S D I I S F H S Y A G	GEFEARINEL.	APLGRPILCTE:	Y L A R D Q G S I	VEGV <mark>LP</mark> VAKRHI	N VGAYSWGLVA	GKTÕT <mark>YF</mark> PWI	D S W D K P
AIU17506.1	NSDVISFHSYAC	PDEFEARIAEL	EPLGRPILCTE:	Y LARSEGS 1	LEGVLPVAKRHI	VGAYSWGLVA	GKTQTYFPWI	DSWDKP
AKS36268.1 CCC45453.1	NADVITEHSYA	PAEFEGRIAEL	APLORPILCTE:	TLARSOGST	VEGTLPTAKRHI	VGATSWGLVA	GKTOTYLPWI	DSWDRP
NP_217612.1	NADVITFHSYAA	PAEFEGRIAEL	APLORPILCTE	Y L A R S Q G S I	VEGILPIAKRHI	VGAFNWGLVA	GKTQTYLPWI	DSWDHP
BBA89654.1	NADVITFHSYAA	APAGFEARIDEL.	APLGRPIICTE:	Y L A R S Q G S S	VEGV <mark>LP</mark> I <mark>AKR</mark> RI	N VGAFNWGLVA	GKTQT <mark>YL</mark> PWI	D S W D H P
ABL05635.1 AXN43512.1	NADVITFHSYAP	PAGFEARIDEL.	APLERPIICTE: APLERPIICTE:	Y LARSODSS Y LARSOCSS	VEGVLPIAKRE VEGVLPIAKRE	NVGAFNWGLVA NVGAFNWGLVA	GKTQTYLPWI	DSWDHP
consensus>70	N.D!I.FHsYa.	PFE.RI.#L	.PLgRPI1CTE	Y\$AR.qgSt	v#g!LP!AKR.	nVGA%.WGlVA	GKTQT%.PWI	DSW#.P
			-					
	330	340	350	3 6 Q				
AYJ05285.1	DPAM. PREWFH	IDLLDPDGRPFR	DSEIQTILE <mark>L</mark> SI	DLAMHLQPF	PPAG			
ASW92947.1	NPDPAMPOEWFH	I <mark>DLL</mark> G <mark>PD<mark>GRP</mark>FR</mark>	DTEIQTILELSI	DLAMHLQPP	PAG			
AFC43445.1 VEG20668.1	VTKT. PDVWBS		DARYOTLREIS	ULAMHLRPF TRV	PAG			
AKC40943.1	YTKIPDVWFS	DLLQPDGRPYR	DAEYQTLRKLT	rrv				
AN006525.1	YTKIPNVWFS	DLLQPDGRP YR	DAEYQTLRKLT	ARV	•••			
AIU17506.1	YTKV. PNVWFH		DSEYATLRKLT	ARV	•••			
CCC45453.1	YRAP. PKVWFH	IDLLHPNGRPYR	DGEVQTIRKLN	GMPSQD				
NP_217612.1	YRAPPKVWF	IDLLHPNGRPYR	DGEVÕTIRKLNO	GMPSQD				
BBA89654.1	YTKP. PKVWFS	DLLOPEGRPYR	ESEIQSIQS <mark>L</mark> TO	GARTQD	•••			
AXN43512.1	YTKP. PKVWPS	DLLOPDGRPYR.	ESETOTIOSLT	GARTOD				
consensus>70	yP.vWF.	DLL.P#GRP%r	#.E.qtL.	•••••	•••			

Figure 34 : Alignement multiple des séquences de GlfH1(NP_217612.1) avec d'autres glycosidases mycobactériennes de la famille GH5_13. Les séquences sont identifiées par leur numéro d'accession GenBank (NCBI). La séquence consensus est décrite comme suit : majuscule représente l'identité, minuscule représente le niveau consensus> 0.7, ! représente I ou V, \$ représente L ou M, % représente F ou Y, # représente l'un des NDQEBZ.

Ensuite, cet alignement de 13 séquences de la famille GH5_13 a été utilisé pour construire un arbre phylogénique selon la méthode *PHYLIP Neighbor-Joining* (Fig 35). Le *clustering* des séquences a aussi été confirmé par la méthode *PhyML Maximum Likehood* (Fig 36).



Figure 35 : Analyse phylogénétique des glycosidases mycobactériennes de la sous-famille GH5_13 en utilisant la méthode *PHYLIP Neighbor Joining* avec un *boostrapping* de 100 réplicats.

Le regroupement obtenu des séquences choisies corrobore avec la classification de *Mycobacterium* récemment proposée par Gupta *et al.*¹⁸ et qui est basée, entre autres, sur l'analyse des séquences de 1941 protéines conservées dans le genre. Les séquences de GlfH1 analysées sont regroupées dans trois des quatre nouveaux clades proposés par Gupta, c'està-dire le clade *M. abscessus-chelonae*, le clade *M. fortuitum-vaccae* et le clade *M. tuberculosis-simiae*.



Figure 36 : Analyse phylogénique de glycosidases mycobactériennes de la sous-famille GH5_13 à l'aide de la méthode *PhyML Maximum Likehood* avec la matrice de substitution de Dayhoff et un *boostrap* de 100 réplicats

L'identification de nouveaux orthologues pour l'addition de nouvelles séquences de GlfH1 permettra certainement de montrer des membres de GlfH1 dans les quatre clades proposés par Gupta, à savoir en plus la clade *Terrae*.

1.2 Production de GlfH1

1.2.1 Expression de GlfH1 chez *E. coli*

La protéine GlfH1 de *M. tuberculosis* a été produite chez *E. coli* BL21 à l'aide du vecteur d'expression pET32a sous forme d'une protéine de fusion (Fig 37). Deux étiquettes permettent une purification par chromatographie d'affinité : l'étiquette poly-histidine par chromatographie d'affinité sur métal immobilisé (*Immobilized Metal Affinity Chromatography*, IMAC) et le *S-Tag* par reconstruction de la ribonucléase S. La séquence primaire de la protéine de fusion GlfH1 comprend aussi une étiquette Trx (thioredoxine d'*E. coli*) *destinée* à améliorer la solubilité de la protéine recombinante et faciliter la mise en conformation correcte de la protéine recombinante ainsi que deux sites de clivage par une protéase à spécificité étroite (pour excision des étiquettes), un pour la thrombine (LVPR \downarrow GS) et un pour l'entérokinase (DDDDK \downarrow).



QGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQL100KEFLDANLAGSGSGHMHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPD150LGTDDDDKAMADHRRTALKLPLLLAAGTVLGQAPRAAAEEPGRWSADRAH200RWYQAHGWLVGANYITSNAINQLEMFQPGTYDPRRIDNELGLARFHGFNT250VRVFLHDLLWAQDAPGFQTRLAQFVAIAARYHIKPLFVLFDSCWDPLPRP300GRQRAPRAGVHNSGWVQSPGAERLDDRRYASTLYNYVTGVLGQFRNDDRV350LGWDLWNEPDNPARVYRKVERKDKLERVAELLPQVFRWARTVDPVQPLTS400GVWQGNWGDPGRRSTISAIQLDNADVITFHSYAAPAEFEGRIAELAPLQR450PILCTEYLARSQGSTVEGILPIAKRHNVGAFNWGLVAGKTQTYLPWDSWD500HPYRAPPKVWFHDLLHPNGRPYRDGEVQTIRKLNGMPSQD*Poids moléculaire calculé = 60477 Da

Figure 37 : Représentation schématique à l'aide du logiciel pDRAW32 de la séquence codante du vecteur pET32a- GlfH1. Sont indiqués sur la séquence en acide aminés de la protéine de fusion : l'étiquette thioredoxine Trx-Tag (rouge foncé), les 6 histidines du His-Tag (orange), le S-Tag (violet), le site de clivage de l'entérokinase (vert), le site de clivage de la thrombine (rouge) et la séquence correspondant à GlfH1 (bleu). Le poids moléculaire apparent proposé par le logiciel pour la protéine de fusion est de 60,5 kDa.

L'expression de la protéine GlfH1 dans *E. coli* BL21 (cultivées dans le milieu LB classique avec induction par l'IPTG à 37 °C ou à 25 °C) a été examinée par SDS-PAGE (Fig 38). Pour la culture à 37 °C, la présence d'une bande intense de poids moléculaire apparent correspondant à la protéine de fusion GlfH1 (Fig 38, piste 2) est observée. GlfH1 est donc correctement induite

et exprimée. Par contre, si la bande correspondant à la protéine de fusion est très visible dans le lysat cellulaire total de la souche BL21 pET32a-Rv3096, elle ne l'est pas de façon aussi évidente dans la faction soluble du lysat cellulaire (Fig 38, piste 6).

Cette première tentative de production de GlfH1 dans *E. coli* a montré que la protéine de fusion était principalement produite sous forme de corps d'inclusion. Les principales causes de la formation de corps d'inclusion sont les vitesses de traduction et de mise en conformation des protéines dans *E. coli* (10 fois supérieures à celles des cellules eucaryotes), le caractère hydrophobe de la protéine recombinante et l'absence des protéines accessoires adéquates facilitant la mise en conformation des protéines naissantes²⁵⁶. Par exemple, sur un panel de 110000 protéines testées, seulement 50 % des protéines procaryotes et 10 % des protéines eucaryotes ont été exprimées principalement sous forme soluble dans *E. coli*²⁵⁷.



Figure 38 : Analyse par SDS-PAGE (10 %) des extraits cellulaires protéiques de souches d'*E. coli* BL21 pET32a-Rv3096 cultivées en milieu LB. **1.** Lysat cellulaire total sans induction, **2.** Lysat cellulaire total avec induction à 37 °C, **3.** Lysat cellulaire total avec induction à 25 °C, **4.** Marqueur de poids moléculaire *Prestained PAGE RULER plus* (Thermoscientific), **5.** Fraction soluble de l'extrait cellulaire avec induction à 37 °C, **7 :** Fraction soluble de l'extrait cellulaire avec induction à 37 °C, **7 :** Fraction soluble de l'extrait cellulaire avec induction à 37 °C, **7 :** Fraction soluble de l'extrait cellulaire avec induction à 25 °C.

Donc, pour minimiser au maximum ce phénomène, j'ai dû modifier les conditions initiales de production de GlfH1 en jouant d'abord sur la température et le type d'induction d'expression de GlfH1. Donc, la température de production de GlfH1 a été fixée à 25 °C et non plus 37 °C. Des essais supplémentaires réalisés à 16 °C n'ont pas montré d'amélioration significative de la quantité récupérée de GlfH1 sous forme soluble. Ensuite, l'induction n'a pas été réalisée classiquement grâce à l'IPTG mais en utilisant un milieu auto-inducteur (AIM). Le milieu AIM est formulé de façon à permettre la culture des souches *d'E. coli* destinées à la production de

protéines recombinantes par induction par l'IPTG mais avec une levée de la répression n'utilisant pas l'IPTG. Le principe des milieux AIM est basé sur une utilisation différentielle des sources de carbone du milieu permettant une croissance des cellules à haute densité et une induction automatique de l'expression de la protéine à partir du promoteur *lac*. Une concentration limitée de glucose est métabolisée préférentiellement au début de la croissance, ce qui empêche l'absorption du lactose jusqu'à ce que le glucose soit épuisé, ce qui arrive généralement dans la phase logarithmique tardive. Suite à l'épuisement du glucose, le lactose est utilisé et entraîne la levée de la répression du promoteur T7lac du vecteur permettant ainsi l'expression de la protéines cible par l'ARN polymérase T7. De cette façon, l'expression de la protéine cible se met en place progressivement et non pas de façon brusque comme c'est le cas pour l'induction par l'IPTG.

1.2.2 Extraction des protéines

Comme GlfH1 semble former majoritairement des corps d'inclusion dans *E. coli*, j'ai utilisé plusieurs détergents pour tenter de stabiliser la fraction soluble de GlfH1 et éviter qu'elle ne s'agrège aussi lors de l'extraction des protéines solubles.



Figure 39 : Analyse par SDS-PAGE (T=10 % homogène) des fractions solubles et insolubles de souches d'*E. coli* BL21-pET32a- GlfH1 cultivées en milieu audoinducteur (AIM) à 25 °C, obtenu par *sonication*, avec ou sans la présence de détergents. Le gel supérieur (A) correspond à la fraction soluble et le gel inférieur (B) à la fraction insoluble. 1. et 2. après *sonication* sans détergent; 3. après *sonication* en présence de CHAPS ; 4. après *sonication* en présence de DDM ; 5. Marqueur de poids moléculaire *Precision Plus Protein Standards All Blue* (Bio-Rad) ; 6. après *sonication* en présence de LDAO ; 9. après *sonication* en présence de Tween20.

D'abord, le culot de culture de la souche BL21 d'*E. Coli* transformées (cultures varient de 500 ml à 2 l) a été lysé par *sonication* avec 10 cycles de 5 secondes ou 10 cycles 15 secondes (avec 10 secondes de pause entre chaque cycle) et en présence de 0,5 % (p/v) de plusieurs détergents : CHAPS (zwitterionique), DDM (non ionique), OBG (non ionique), HM (non ionique), LDAO (zwitterionique) et Tween20 (non ionique). Après centrifugation, le surnageant a été séparé du culot insoluble. Ces deux fractions ont d'abord été analysées par SDS-PAGE (Fig 39).

Cette première analyse montre que la lyse en présence des détergents DDM et LDAO permet de récupérer une quantité supérieure de protéines solubles dont GlfH1. Elle permet aussi de montrer que la majeure partie de GlfH1 se retrouve dans la fraction insoluble sous forme de corps d'inclusion.

L'activité de GlfH1 a été mesurée sur ces deux fractions en utilisant le substrat synthétique pNP-Galf à pH 4,5 et 37 °C et avec une incubation pendant 15 min (suite à la mise au point des conditions expérimentales qui va être détaillée dans le chapitre suivant).



Figure 40 : Activité totale (AT) et activité spécifique (AS) de GlfH1 exprimée dans *E. coli* en fonction de la méthode de lyse (*sonication* par cycles de 5 ou 15 secondes ou presse de French) et du détergent employé. Chaque catégorie représente le détergent employé à l'exception de la dernière catégorie où la lyse a été réalisée par presse de French (1650 psi, 2 cycles), Pour chaque catégorie, les deux valeurs de droite correspondent à l'activité totale de GlfH1 (AT, échelle de gauche) et les deux valeurs de droite (hachures) à l'activité spécifique (AS, échelle de droite). De même, pour les couples de valeurs d'activité totale et d'activité spécifique, la mesure de gauche correspond à une lyse par *sonication* avec cycles de 5 sec et la valeur de droite à lyse par *sonication* avec cycles de 15 sec.

Dans la figure 40, j'ai fait varier les conditions (*sonication* par cycles de 5 ou 15 sec ou presse de French) en fonction des détergents utilisés. À partir du surnageant obtenu après centrifugation et contenant les protéines solubles extraites, j'ai calculé l'activité totale (AT) qui me permet d'évaluer la quantité de GlfH1 fonctionnelle récupérée et l'activité spécifique (AS) qui me permet d'évaluer l'enrichissement en GlfH1 par rapport aux autres protéines extraites. Une première constatation est l'importance de la durée des cycles de *sonication*. En effet, il apparaît que des cycles de 15 sec amènent systématiquement à une quantité

inférieure de GlfH1 fonctionnelle récupérée. Pour la suite, les cycles de *sonication* seront fixés à 5 sec avec 10 sec de repos sur glace entre chaque cycle. En ce qui concerne le choix du détergent, le LDAO affiche à la fois la plus grande quantité de GlfH1 obtenue et aussi la meilleure activité spécifique. Il sera donc retenu pour les purifications à venir. Concernant la lyse par presse de French, si la quantité totale de GlfH1 récupérée est bonne, l'activité spécifique est une des plus faibles obtenues, ce qui risque de compliquer l'étape de chromatographie IMAC à venir.

1.3 Isolement de GlfH1

La purification de GlfH1 produite dans *E. coli* a donc été réalisée à 25 °C et en milieu AIM classiquement sur un volume de 2 l. La protéine de fusion GlfH1 comporte deux étiquettes d'affinité, *His-Tag* et *S-Tag*, et c'est l'isolement par chromatographie IMAC avec l'étiquette *His-Tag* qui a été retenue. Après isolement de la protéine, il a été prévu dans un premier temps de ne pas éliminer la partie de GlfH1 comportant les différentes étiquettes à l'aide des sites de coupures pour protéases à spécificité étroite, thrombine ou entérokinase, pour conserver l'étiquette Trx qui améliore la solubilité de GlfH1. L'obtention des protéines solubles d'*E. coli* a été réalisée selon les conditions mises au point dans le paragraphe précédent.



Figure 41 : Purification de GlfH1 par chromatographie IMAC. Pour tous les chromatogrammes, le suivi de l'élution des protéines a été réalisé par mesure de l'absorbance à 280 nm. A) chromatographie IMAC sur colonne de 5 ml. La courbe en rouge correspond à la concentration en imidazole. La flèche rouge indique la protéine associée éluée après incubation pendant 1 h dans le tampon de lavage. B) Idem au point A) sauf que la chromatographie IMAC de concentration a été réalisée avec une colonne de 1 ml. C) Chromatographie de gel filtration pour dessalage de la protéine GlfH1. La courbe en rouge correspond à la mesure de la conductivité.

Le protocole de purification de GIfH1 comprend trois étapes de chromatographie : deux chromatographies IMAC sur support Ni-NTA (GE Healthcare Lifesciences) et une chromatographie de gel filtration pour dessalage. La première chromatographie IMAC (Fig 41A) est l'étape principale pour isoler GIfH1. La seconde (Fig 41B) sert à concentrer la fraction GIfH1 obtenue sans utiliser de membrane d'ultrafiltration car il s'avère qu'il est très délicat de concentrer une fraction contenante GIfH1 sans perdre totalement la protéine par adsorption. C'est pour cela qu'il a été choisi de concentrer la fraction contenant GIfH1 par chromatographie IMAC en passant d'une première chromatographie avec une colonne de 5 ml à une seconde chromatographie avec une colonne de 1 ml et un volume d'élution beaucoup plus réduit. Cette étape permet de concentrer rapidement et sans perte l'échantillon d'un facteur ≥ 5. Enfin, l'étape finale de gel filtration (Fig 41C) est indispensable pour éliminer l'imidazole provenant de l'élution en IMAC et obtenir GIfH1 directement dans son tampon de stockage.

A l'issue de la première chromatographie IMAC et malgré un profil d'élution de GlfH1 sous forme d'un pic étroit et parfaitement symétrique, l'analyse de la fraction éluée par l'imidazole par SDS-PAGE a montré la présence de deux protéines en rapport 1 : 1. La première correspond à GlfH1 (Fig 42, piste 3, cadre rouge) à la masse moléculaire apparente attendue de 60 kDa et la seconde (Fig 42, piste 3, cadre vert) possède une masse moléculaire apparente d'environ 70 kDa. De multiples essais de purification ont toujours montré le même résultat avec la co-élution de ces deux protéines en rapport 1 :1. Selon la bibliographie, l'hypothèse la plus vraisemblable est que GlfH1 soit restée associée à une protéine chaperonne d'E. coli comme par exemple DnaK dont la masse moléculaire est de 70 kDa²⁵⁸. Comme généralement les protéines chaperonnes perdent leur affinité pour les protéines qu'elles assistent lorsqu'elles fixent de l'ATP, il a été choisi de réaliser, conjointement à l'étape de lavage, une étape d'incubation du support IMAC avec un tampon contenant 5 mM d'ATP et 10 mM de MgCl₂. Cette étape a été testée avec des incubations allant de 30 min à 3 h (Fig 42, pistes 4 à 6). Elle a permis de faire baisser sensiblement le ratio de la protéine associée sans toutefois l'éliminer totalement. Les temps d'incubation testés ne montrent pas de différences significatives et il a été choisi de les fixer à 1 h pour les purifications ultérieures. Les étapes suivantes de concentration par chromatographie IMAC et de gel filtration permettent au final d'éliminer totalement cette protéine associée. Comme le montrent les chromatogrammes présentés dans la figure 41A et 41B, à la reprise de la chromatographie après l'incubation en présence d'ATP, le pic correspondant à la protéine associée est visible et celle-ci a pu être récupérée pour identification par empreinte peptidique massique.



Figure 42 : Analyse par SDS-PAGE (T=10 % homogène) des extraits protéiques de souches d'*E. coli* BL21 pET32a-Rv3096 cultivées en milieu autoinducteur (AIM) à 25 °C après purification IMAC. 1. lysat cellulaire total de la souche BL21 DE3 pET32a-Rv3096 après culture ; 2. fraction soluble après lyse par *sonication* ; 3. fraction éluée par l'imidazole lors de la chromatographie IMAC, 4. fraction éluée par l'imidazole lors de la chromatographie IMAC avec étape de lavage en présence d'ATP (30 min) ; 5. fraction éluée par l'imidazole lors de la chromatographie IMAC avec étape de lavage en présence d'ATP (1 h) ; 6. fraction éluée par l'imidazole lors de la chromatographie IMAC avec étape de lavage en présence d'ATP (3 h) ; 7. fraction correspondant à la protéine inconnue décrochée pendant l'étape de lavage ; 8. fraction éluée par l'imidazole lors de la chromatographie IMAC de concentration ; 9. fraction GlfH1 après le dessalage par gel filtration ; 10. marqueurs de poids moléculaire *Precision Plus Protein Standards All Blue* (Bio-Rad).

1.4 Empreinte peptidique massique (PMF)

L'identification des deux protéines éluées par l'imidazole à la suite de la chromatographie IMAC a été réalisée par empreinte peptidique massique. Pour cela, les bandes d'intérêt ont été découpées et digérées à la trypsine pour obtenir un mélange de peptides trypsiques de la protéine à identifier. Les peptides obtenus ensuite été analysés par spectrométrie de masse MALDI-TOF. A partir d'un spectre MS contenant les rapports m/z de chaque peptide du mélange, les spectres mono-isotopiques sont sélectionnés par le logiciel de traitement de l'appareil puis comparés aux banques de données en ligne. La requête d'identification a été réalisée grâce au serveur Mascot v2.51 (<u>http://www.matrixscience.com</u>) avec comme paramètres la prise en compte d'une coupure manquée par la trypsine, la carbamidométhylation comme modification fixée et la banque NCBInr pour l'interrogation. La banque nr est compilée par le NCBI en tant que banque d'interrogation pour les recherches BLAST. Elle contient des séquences non redondantes provenant des banques GenBank CDS *translation*, PDB, Swiss-Prot, PIR et PRF. Elle est, à ce jour, la banque la plus exhaustive pour les séquences de protéines.

1.4.1 GlfH1

Le PMF a d'abord été réalisé sur la bande correspondant normalement à GlfH1 afin de confirmer son identité (Fig 43). Évidemment, la partie de la séquence correspondant aux étiquettes fusionnées à GlfH1 n'a pas été prise en compte dans l'analyse.

Comme résultat de l'interrogation de la banque NCBInr, La protéine GIfH1 a bien été identifiée avec une couverture de séquence par les peptides trypsiques de 84 % et un score de 100. Pour information, le score représente la valeur -10 x Log(P) avec P étant la probabilité que le résultat obtenu soit dû au hasard. Un score de 70 correspond à une confiance supérieure à 95 %. Avec un score de 100, l'identification de GIfH1 est donc sans ambiguïté.

1.4.2 Protéine associée

En ce qui concerne la protéine associée, le résultat obtenu n'a pas confirmé l'hypothèse qu'il ait pu s'agir d'une protéine chaperonne (Fig 44). En effet, la protéine identifiée est une protéine bifonctionnelle d'*E. coli*, l'UDP-acide glucuronique oxydase/UDP-4-amino-4-deoxy-*L*arabinose formyltransférase (ou ArnA) avec 67 % de couverture de séquence et un score de 90. L'identification est de nouveau sans ambiguïté.



Figure 43 : Spectre de masse MALDI-TOF des peptides trypsiques de GlfH1. Sur le spectre sont indiquées les signaux les plus intenses ayant servis à l'identification de la protéine. Dans l'encart est indiquée la séquence primaire de GlfH1 (sans les étiquettes) et la zone de couverture des peptides identifiés (en rouge).



Figure 44 : Spectre de masse MALDI-TOF des peptides trypsiques de la protéine inconnue associée à GlfH1 lors de son isolement. Sur le spectre sont indiqués les signaux les plus intenses ayant servi à l'identification de la protéine. Dans l'encart est indiquée la séquence primaire de l'épimérase/déshydratase NAD dépendante d'*E. coli* et la zone de couverture des peptides identifiés (en rouge).

Par contre l'identité de la protéine ArnA est surprenante car elle n'a aucun lien avec les protéines associées à la mise en conformation des chaines naissantes. ArnA est donc une

enzyme bifonctionnelle qui catalyse la décarboxylation oxydative de l'UDP-acide glucuronique (UDP-GlcUA) en UDP-4-deoxy-L-arabinose (UDP-Ara4O) et aussi l'addition d'un groupement formyle UDP-4-amino-4-deoxy-L-arabinose (UDP-L-Ara4N) pour former l'UDP-L-4-formamidoarabinose (UDP-L-Ara4FN). Cet arabinose modifié sera ensuite attaché au lipide A et est nécessaire à la résistance de la bactérie à la polymyxine et aux peptides cationiques antimicrobiens^{259,260}. Mais comme cela est souvent rencontré pour les protéines interactantes, cette protéine bifonctionnelle chez *E. coli* est présente chez *M. tuberculosis* sous forme de deux orthologues distincts²⁶¹(principe de la pierre de Rosette). Et l'un des deux orthologues est *Rv3464* qui code pour la protéine RmlB impliquée dans la biosynthèse de l'arabinogalactane des mycobactéries. Il faudra peut-être envisager que cette interaction hétérologue chez *E. coli* est le reflet d'une interaction homologue chez *M. tuberculosis* impliquant des protéines associées au métabolisme de l'arabinogalactane. Le point négatif de cette hypothèse est que RmlB est une protéine cytosolique (pour ce que l'on en sait actuellement) et que GlfH1 est certainement exoplasmique de par sa fonction et la présence d'un peptide signal putatif.

II. Caractérisation de GlfH1

1. Paramètres enzymatiques de GlfH1

Les paramètres enzymatiques de GlfH1 ont été déterminés en utilisant un substrat synthétique, le *p*NP-Gal*f*. Ce substrat a été synthétisé par le Dr. Sydney Villaume de l'équipe du Pr. Stéphane Vincent. Il convient de noter que le GlfH1 utilisé pour la caractérisation est sous sa forme complexe, c'est-à-dire en présence du peptide signal et des étiquettes.

1.1 Validation du substrat synthétique : le pNP-Galf

Pour déterminer les conditions expérimentales de base, il a d'abord fallu tester le substrat visà-vis de l'activité enzymatique de GlfH1. La détermination s'est faite grâce à une gamme de GlfH1 allant de 0 à 500 ng, 1 mM de substrat (*p*NP-Gal*f*) pour un volume final de 100 µl de tampon phosphate 20 mM pH 7,0. Le tampon est déposé en premier, puis le *p*NP-Gal*f*. Après un temps de pose de 5 min, l'enzyme est ajoutée. Le tout est incubé à 37 °C pendant 15 min, puis la réaction est arrêtée par 50 µl d'une solution carbonate de sodium 0,5 M (pH=11,5). L'absorbance est ensuite mesurée à 405 nm. La figure 45 montre que GlfH1 utilise bien le *p*NP-Gal*f* pour son activité hydrolytique avec relargage du *p*NP en conséquence.



Figure 45 : Représentation graphique de l'activité enzymatique en μ mole/min en fonction de différentes concentrations de GlfH1 en utilisant le substrat synthétique *p*NP-Gal*f*. ------ : intervalle de confiance à 95 %.

1.2 Calcul de la vitesse enzymatique

Afin de pouvoir convertir l'absorbance à 405 nm donnée par le pNP libéré, une titration du pNP a été réalisée dans un tampon dont la composition correspond à celle obtenue après arrêt de la réaction par le carbonate de sodium 0,5 M (pH=11,5). En effet, l'absorbance à 405 nm du pNP n'apparaît qu'en milieu basique.



Figure 46 : Représentation graphique de l'absorbance à 405 nm en fonction de la concentration en pNp. A) gamme large de substrat pNP, B) gamme étroite de substrat pNP.

Jusqu'à une concentration de 0,5 mM, le dosage colorimétrique du *p*NP à 405 nm montre une relation totalement linéaire avec une équation Y=4,960*X (intervalle de confiance à 95 % pour la pente : 4,904 <pente<5,017) (Fig 46). Cette équation sera utilisée pour calculer les vitesses enzymatiques à partir de la lecture de l'absorbance à 405 nm.

1.3 Détermination des conditions de vitesse initiale

Pour les expériences suivantes et surtout pour pouvoir déterminer les constantes cinétiques de GlfH1, il faut s'assurer que l'expérience sera bien réalisée en condition de vitesse initiale, c'est-à-dire que la quantité de substrat consommé reste négligeable par rapport à son influence potentielle sur la vitesse enzymatique pour une concentration d'enzyme et un temps d'incubation donnés. Les conditions de vitesse initiale dépendent principalement de la concentration en enzyme, de la concentration en substrat et de la cinétique d'hydrolyse. L'expérience a été réalisée en faisant varier la concentration en substrat et le temps de réaction pour une quantité donnée en enzyme (25 ng). Les résultats (Fig 47) montrent que pour des concentrations en *p*NP-Gal*f* allant jusqu'à 5 mM et des temps de réaction allant

jusque 30 min, la libération de *p*NP reste linéaire. Par la suite, un temps de réaction de 15 min sera utilisé car l'absorbance obtenue est proche de 0,7 et est donc idéale pour un dosage colorimétrique.



Figure 47 : Détermination des conditions expérimentales de vitesse initiale, représentation graphique de l'absorbance à 405 nm avec différentes concentrations du substrat *p*NP en fonction du temps.

1.4 Détermination du pH optimal

Pour déterminer le pH optimal d'activité de GlfH1, des solutions tampons à différents pH ont été préparées. Plusieurs substances ont été utilisées afin de conserver un pouvoir tampon adéquat (tolérance de ± 1,5 unité pH autour du pK de la substance tampon) : tampon citrate (mélange d'acide citrique 20 mM et de citrate de sodium 20 mM) à pH 3,0, pH 3,5, pH 4,0, pH 4,5, pH 5,0, pH 5,5, pH 6,0. Tampon tris-maléate à pH 6,0, pH 6,5, pH 7,0, pH 7,5, pH 8,0, pH 8,5, pH 9,0. Une fois les tampons préparés, la réaction se fait avec 1 mM de substrat (*p*NP-Gal*f*), 10 ng de GlfH1 et un qsp à 100 µl avec les tampons de pH souhaité. L'incubation est réalisée à 37 °C pendant 15 min puis la réaction est arrêtée par 50 µl d'une solution carbonate de sodium 0,5 M. L'absorbance est mesurée à 405 nm.

Dans la figure 48, j'ai observé qu'à pH 6 avec deux tampons différents, l'activité enzymatique est presque identique, montrant que les substances tampons utilisées n'ont pas d'effets différentiels sur l'activité de GlfH1. Les résultats obtenus me permettent aussi de conclure que le pH optimal de GlfH1 est de 4,5. La suite de nos expériences sera donc réalisée à pH 4,5 dans un tampon citrate 20 mM.



Figure 48 : Représentation graphique de l'activité enzymatique en µmole/min en fonction du pH.

1.5 Stabilité

La stabilité de GlfH1 à différents pH a aussi été étudiée. Pour cela, 25 ng de GlfH1 sont ajoutés à 20 µl de différents tampons dont le pH varie de 3 à 9 puis incubés à 37 °C pendant 30 min. Ensuite, le substrat et le tampon citrate pH 4,5 (qsp 100 µl) sont ajoutés et la réaction enzymatique est réalisée comme précédemment.

Sur la figure 49, j'observe que, dans ces conditions expérimentales, GlfH1 conserve une activité enzymatique stable pendant une durée de 30 min sur une gamme de pH s'étendant de 4,5 à 7,0. L'activité de GlfH1 chute très rapidement en dessous de pH=4,5 et plus progressivement pour les pH supérieurs à 7,0.



Figure 49 : Représentation graphique de la stabilité de l'activité enzymatique de GlfH1 en fonction du pH. La mesure de l'activité résiduelle a été réalisée dans les conditions standards avec une incubation de 30 min au pH indiqué en abscisse.

1.6 Détermination des constantes michaeliennes : K_M et k_{cat}

Les manipulations précédentes m'ont permis de déterminer les conditions pour l'étude de l'activité de GIfH1 comme étant :

- GlfH1 : 25 ng ;
- [pNP-Galf] : 0 à 5 mM ;
- tampon : citrate 20 mM pH 4,5 avec 20 mM de CaCl₂;
- durée : 15 min.

La détermination de Km et Vm se fait grâce à une gamme de substrat (pNP-Galf) allant de 0 mM à 5 mM, 25 ng de GlfH1 et un qsp de 100 µl de tampon citrate à pH 4,5 avec CaCl₂ 20 mM. Le tampon est déposé en premier, puis le pNP-Galf. Après un temps de pose de 5 min, les extraits de la protéine sont ajoutés. Le tout est incubé à 37 °C pendant 15 min, puis la réaction est arrêtée par 50 µl d'une solution carbonate de sodium pH 11,5. L'absorbance est mesurée à 405 nm. La gamme de pNP-Galf (0 à 5 mM) a été choisie par rapport aux valeurs moyennes de Km observées dans la littérature pour des galactosidases utilisant un substrat synthétique avec pNP partant. Comme il s'agit d'un substrat disponible, au départ de l'étude, en quantité limité, j'ai décidé de *limiter* la gamme à environ 5 fois la valeur du Km attendu contrairement aux 10 fois recommandées pour une étude de ce type.

L'allure de la courbe correspond bien à une cinétique *michaelienne* (Fig 50). Les K_M , k_{cat} et k_{cat}/K_M ont été établis à 0,6 ± 0,1 mM, 11 s⁻¹ et 1,8 × 10⁴ M⁻¹.s⁻¹ respectivement avec un

intervalle de confiance de 95 %. Ce dernier est tout à fait en accord avec les données de la littérature pour une galactofuranohydrolase utilisant un substrat couplé au pNP^{262} .



Figure 50 : Représentation graphique de l'activité enzymatique en μ mole/min en fonction de la concentration en *p*NP-Gal*f*. ------- : intervalle de confiance à 95 %.

1.7 Rôle du calcium

L'effet de plusieurs cations divalents (Mn²⁺, Mg²⁺, Ca²⁺, Ni²⁺ et Zn²⁺) vis-à-vis de l'activité de GlfH1 a été évalué en les utilisant sous forme de chlorure à une concentration finale de 20 mM. Une chélation par l'EDTA a aussi été utilisée pour savoir si des cations divalents n'ont pas seulement un rôle activateur mais sont absolument nécessaires à l'activité de GlfH1. L'expérience a été réalisée dans les conditions déterminées précédemment soit pH 4,5 ; 37 °C, 15 min.

Dans la figure 51A, j'ai observé qu'en présence de calcium, l'activité enzymatique est augmentée de 1,3*10⁻³ µmole/min à 1,8*10⁻³ µmole/min, soit un gain d'activité de 35 %. Le calcium a donc bien un effet positif sur l'activité enzymatique de GlfH1. Par contre en présence de Mn²⁺, Mg²⁺, Ni²⁺ et Zn²⁺, l'activité enzymatique est diminuée. Par exemple, l'activité de GlfH1 a diminué en présence de Ni²⁺ et de Zn²⁺ de 76 % et 86 % respectivement.

Les ions Ca²⁺ montrent donc un impact positif sur l'activité de. Pour ensuite clairement démontrer le caractère indispensable des ions Ca²⁺, GlfH1 a d'abord été incubée durant 5 min à 37 °C en présence de chélateurs des cations divalents. Deux chélateurs ont été utilisés : l'EDTA et l'EGTA. L'EDTA est un chélateur très fréquent, qui a un fort pouvoir chélatant pour les cations métalliques. EGTA est un chélateur moins fréquent mais avec une sélectivité 1000

fois supérieure pour les ions Ca²⁺ comparé aux ions Mg²⁺ par exemple. Suite à cette première étape, l'activité résiduelle de GlfH1 a été mesurée soit simplement selon le protocole standard (Fig 51B, +EDTA et +EGTA), soit en ajoutant 50 mM de CaCl₂ dans le tampon standard (Fig 51B : EDTA \rightarrow Ca²⁺ et EGTA \rightarrow Ca²⁺) pour complémenter la perte des cations pendant la préincubation en présence de chélateurs.

L'expérience standard a été réalisée avec 200 ng de GIfH1, 10 μ M d'EDTA ou EGTA et un qsp à 100 μ l avec un tampon citrate 20 mM pH 4,5 avec 20 mM de CaCl₂. Après pré-incubation en présence d'EDTA ou d'EGTA, l'activité enzymatique de GIfH1 est clairement diminuée : -84 % en présence d'EDTA et -50 % en présence d'EGTA. Ce résultat reste toutefois logique en considérant que l'EDTA possède une affinité pour les ions Ca²⁺ tandis que l'EGTA présente une meilleure spécificité mais une affinité inférieure. Par contre, il n'est pas possible de savoir si l'activité résiduelle est due au fait que tout le calcium lié à GIfH1 n'a pas été chélaté ou si GIfH1 conserve une activité résiduelle en absence de calcium. Fait intéressant, la complémentation du tampon de réaction en ions Ca²⁺ permet de restaurer complétement l'activité de GIfH1 et prouve donc le caractère indispensable des ions Ca²⁺.



Figure 51 : Pourcentage d'activité résiduelle de GlfH1 en présence de différents chélateurs ou cations divalents. A) GlfH1 a été pré-incubée pendant 5 min à 37 °C en présence d'EDTA ou de différents cations divalents (n=1). L'activité résiduelle a ensuite été mesurée selon le protocole standard. B) +EDTA et +EGTA : GlfH1 a été pré-incubée pendant 5 min à 37 °C en présence d'EDTA ou d'EGTA avant de mesurer l'activité résiduelle de GlfH1 selon le protocole standard (n=3). EDTA \rightarrow Ca²⁺ et EDTA \rightarrow Ca²⁺ : GlfH1 a été pré-incubée pendant 5 min à 37 °C en présence d'EDTA ou d'EGTA. L'activité résiduelle a ensuite été mesurée selon le protocole standard (n=3). EDTA \rightarrow Ca²⁺ et EDTA \rightarrow Ca²⁺ : GlfH1 a été pré-incubée pendant 5 min à 37 °C en présence d'EDTA ou d'EGTA. L'activité résiduelle a ensuite été mesurée selon le protocole standard mais avec un tampon contenant en plus 50 mM de CaCl₂.

L'impact des ions Ca²⁺ sur l'activité de GlfH1 est surprenant car le mécanisme catalytique des glycosidases repose entièrement sur une paire de chaines latérales de résidus d'acide aspartique/glutamique. Cependant, cette importance du calcium a déjà été mise en évidence pour l'endo-1,5- α -L-arabinanase BsArb43B de *Bacillus subtilis*²⁶³ de la famille GH43, plusieurs exo-mannosidase des familles GH38, 47 et 92 et deux glucosidases (BtGH97a et BtGH97b) de la famille GH97¹⁹⁰. Cependant, cette nécessité du calcium n'avait pas encore été observée chez les β -Gal*f*-ases.

1.8 Spécificité de GlfH1 vis-à-vis du monosaccharide libéré

Pour déterminer la spécificité de GlfH1 vis-à-vis du monosaccharide libéré, j'ai utilisé douze différents substrats synthétiques (Fig 52). Le *p*NP- β -D-Gal*f* est le substrat habituel de GlfH1 et sert de contrôle positif. Le *p*NP- β -D-Gal*p* est utilisé pour confirmer la spécificité de GlfH1 pour la forme furanique. Le *p*NP- α -D-Gal*f* est utilisé pour confirmer la spécificité de GlfH1 pour la liaison β . Le *p*NP- α -L-Ara*f* a été choisi car les rares exogalactofuranohydrolases identifiées ont souvent montré qu'elles possédaient aussi une L-arabinasique²⁶⁴.

Parmi les douze substrats synthétiques de *p*NP-glycosides testés, seul le *p*NP- β -D-Gal*f* a été hydrolysé par GlfH1. Point important, GlfH1 n'a montré aucune activité envers *p*NP- α -L-Ara*f* alors que ces activités sont très souvent associées à d'autres galactofuranosidases²⁶⁴.

pNP-monosaccharide	Activité Relative (%)		
₽NP-β-D-Gal <i>f</i>	100 %		
<i>p</i> NP-β-D-Gal <i>p</i>	0 %		
<i>p</i> NP-α-D-Gal <i>f</i>	0 %		
pNP-₿-D-Glcp	0 %		
<i>p</i> NP-α-L-Ara <i>f</i>	0 %		
<i>p</i> NP-β-D-GalNAc	0 %		
<i>р</i> NP-ß-D-GlcNAc	0 %		
<i>p</i> NP-β-D-Man <i>p</i>	0 %		
<i>p</i> NP-α-D-Man <i>p</i>	0 %		
pNP-B-D-Xylp	0 %		
pNP-a-D-Xylp	0 %		
<i>p</i> NP-α-D-Lac <i>p</i>	0 %		

Figure 52 : Activité de GlfH1 vis-à-vis de douze substrats synthétiques en utilisant les conditions standard d'hydrolyse, c'est-à-dire 25 ng de GlfH1 avec 1 mM de substrat et un qsp de 100 μ l de tampon citrate à pH 4,5 avec CaCl₂ 20 mM incubé à 37 °C pendant 15 min, puis la réaction est arrêtée par 50 μ l d'une solution carbonate de sodium pH 11,5. L'absorbance est mesurée à 405 nm.

1.9 Inhibition par l'éthambutol

-

L'éthambutol est une drogue antituberculeuse perturbant la synthèse de l'arabinogalactane en ciblant les arabinosyltransferases codées par les gènes *embA*, *embB* et *embC*^{160,163}. Cette molécule est aussi capable de perturber le fonctionnement de plusieurs autres protéines effectrices comme par exemple le cytochrome P450²⁶⁵. J'ai donc cherché à savoir si l'éthambutol montre un effet sur l'activité de GlfH1.

La détermination d'un effet inhibiteur de l'éthambutol s'est d'abord faite en utilisant une gamme d'éthambutol allant de 0 mM à 100 mM dans les conditions expérimentales standards. Dans la figure 53A, l'activité enzymatique est progressivement inhibée par des concentrations croissantes d'éthambutol. Il y a donc bien un effet inhibiteur de cette molécule sur l'activité enzymatique de GlfH1.

Le type d'inhibition de GlfH1 par l'éthambutol a ensuite été déterminé en calculant les constantes *michaeliennes* de GlfH1 avec des concentrations en éthambutol variant de 0 mM à 30 mM (Fig 53B).



Figure 53 : Inhibition de GlfH1 par l'éthambutol. A) inhibition de GlfH1 par une gamme de concentration croissante en éthambutol. B) Détermination du type d'inhibition et du K_i de l'éthambutol sur GlfH1. Les constantes *michaeliennes* de GlfH1 ont été déterminées pour des concentration en éthambutol allant de 0 à 30 mM.

L'analyse des données d'inhibition a été réalisée avec le logiciel Prism v7.0 (Graphpad). La structure moléculaire de l'éthambutol éloignant l'hypothèse d'une inhibition compétitive, les modèles d'inhibition incompétitive, non-compétitive et mixte ont été testées à partir des résultats expérimentaux (Tableau IV).



Figure 54 : Modèle d'inhibition en mode mixte et non compétitif. Dans le cas d'une inhibition non compétitive, $\delta = 1$ car l'inhibiteur présente la même affinité pour l'enzyme et le complexe enzyme/substrat²⁶⁶.

Et il apparaît que le modèle de l'inhibition non compétitive est celui qui correspond le mieux aux valeurs observées. Le fait que le choix le plus délicat soit entre le modèle non compétitif et le modèle mixte est logique car l'inhibition non compétitive est un cas particulier d'inhibition mixte où l'affinité de l'inhibiteur pour l'enzyme seule ou pour le complexe enzyme/substrat demeure identique (Fig 54)²⁶⁶.

	Non comp./mixte	Non comp./incomp.	Non comp./comp.
Mode	non comp.	non comp.	non comp.
Probabilité (%)	69,11	88	> 99,99
Mode	mixte	incompétitif	compétitif
Probabilité (%)	30,89	12	< 0,01
Ratio des probabilités	2,24	7,33	-
Choix	non compétitif	non compétitif	non compétitif

Tableau IVChoix du mode d'inhibition de GlfH1 par l'éthambutol. L'hypothèse d'une inhibition non compétitivea été testée contre les autres modèles à l'aide du logiciel Prism v7.0.

En présence d'EMB, le K_M est resté constant alors que le V_{max} a diminué de $1,34 \times 10^{-3} \mu$ mole.min⁻¹ à 5,01 × $10^{-4} \mu$ mole.min⁻¹. L'inhibition a ensuite été déterminée comme non compétitive avec un K_i de 23,6 ± 1,7 mM.

Malgré le fait que le K_i d'EMB ne soit que de l'ordre du mM, il peut néanmoins donner des indices pour la conception d'inhibiteurs potentiels de GlfH1 et sur le mécanisme d'inhibition de l'EMB. Environ 35 % des souches de *M. tuberculosis* résistantes à l'EMB ne possèdent pas de mutations au niveau des arabinosyltransférases, ce qui suggère l'existence d'autres mécanismes de résistance impliqués⁶⁵. Par exemple, il a été démontré que la mutation de la DPPR synthase (codée par *ubiA*) était également responsable de la résistance à l'EMB⁷². Il reste toutefois à vérifier si GlfH1 peut représenter une cible primaire *in vivo*.

1.10 Trois formes différentes de GlfH1

Comme décrit dans les parties 1.1, la protéine GlfH1 contient un peptide signal de 27 acides aminés et différentes étiquettes comme *His-Tag* et *S-Tag*. Pour que la caractérisation des propriétés de GlfH1 soit indiscutable, il a fallu démontrer que ses paramètres enzymatiques sont similaires avec ou sans le peptide signal et les étiquettes.

Pour cela, le peptide signal a été enlevé par clonage et les étiquettes ont été éliminées en utilisant l'entérokinase. Les paramètres enzymatiques comme le pH optimal, les constantes *michaeliennes*, l'effet du calcium et l'inhibition d'éthambutol de GlfH1 ont ensuite été mesurés sur trois formes de la protéine : la forme totale de la protéine exprimée (avec le peptide signal et les étiquettes), la forme sans le peptide signal et la forme sans les étiquettes. Comme montré dans le tableau V, les trois formes de GlfH1 ont des caractères enzymatiques
similaires. Leur pH optimal est à 4,5, ils ont tous besoin de calcium pour fonctionner et ils sont tous inhibés par l'éthambutol.

	GlfH1	GlfH1 sans peptide signal	GlfH1 sans étiquettes
<i>K_M</i> (mM)	0,6 ± 0,1	2,5 ± 0,1	0,6 ± 0,03
K_{cat} (s ⁻¹)	11	78	30
$K_{cat}/K_{M}(M^{-1}.s^{-1})$	$1,8 \times 10^{4}$	3,1 x 10 ⁴	5 x 10 ⁴
pH optimal	4,5	4,5	4,5
Nécessité du Ca ²⁺	Essentiel	Essentiel	Essentiel
Inhibition par l'éthambutol	inhibé	inhibé	inhibé

 Tableau V
 Paramètres enzymatiques de GlfH1, GlfH1 sans peptide signal et GlfH1 sans étiquettes.

2. Étude du site actif de GlfH1

2.1 Stéréochilmie du mécanisme catalytique : rétention ou inversion ?

Un point crucial pour comprendre le fonctionnement d'une glycoside hydrolase est de savoir si elle agit par rétention ou par inversion, c'est-à-dire si elle conserve ou inverse l'anomérie du monosaccharide réducteur libéré. C'est un résultat difficile à obtenir car la fonction réductrice libérée va spontanément s'épimériser. Pour cela plusieurs stratégies sont envisageables. Celle que j'ai employé consiste à faire travailler GlfH1 en présence d'un accepteur autre que l'eau pour former un glycoside dont l'anomérie bloquée est ensuite à déterminer. J'ai d'abord réalisé plusieurs essais pour trouver un accepteur potentiel du galactose qui ne compromette pas l'activité de GlfH1. Le méthanol a été retenu car GlfH1 peut travailler en présence de 20 % (v/v) de méthanol en ne perdant que 20 % de son activité sur une incubation standard de 15 min (Fig 55).



Figure 55 : Activité enzymatique résiduelle de GlfH1 en présence de méthanol mesurée dans les conditions standards.

La manipulation a d'abord consisté à utiliser GlfH1 avec le substrat synthétique pNP-Galf et 20 % de méthanol pour former le 1-0-méthyl-Galf. Ensuite, il a fallu déterminer l'anomérie de ce composé. La RMN était la technique la plus adaptée pour cela mais il aurait fallu purifier ce composé en quantité suffisante, ce qui était problématique au vu de la quantité limitée de substrat synthétique. J'ai choisi de déterminer cette anomérie par CPG et par comparaison des temps de rétention avec ceux obtenus pour du galactose méthylés en C₁. Pour cela, l'hydrolyse de *p*NP-Galf par GlfH1 a été effectuée en présence de 20 % de méthanol afin de bloquer la configuration du monosaccharide libéré. Pour l'échantillon standard, le galactose libre a été traité avec du méthanol à 20 % dans les mêmes conditions. Les méthyl-galactosides libérés ont été dérivatisés par per-triméthyl-silylation avant d'être analysés par GC/MS.

Les résultats obtenus sont présentés dans la figure 56. Le chromatogramme en 56A correspond au galactose ayant subi d'abord une méthanolyse pour méthyler le C1 puis une dérivation par triméthylsylilation pour l'analyse par CPG. Les quatre isomères possibles (αf , αp , βf and βp)²⁶⁷ sont séparés et assignés sur la base de données disponibles dans le groupe. Le chromatogramme présenté en 56B montre qu'un produit formé lors du travail de GlfH1 en présence de méthanol correspond parfaitement au β -Galf, classant GlfH1 parmi les glycosides hydrolases agissant par rétention. Cela permet aussi de postuler du fait que les deux résidus catalytiques seront séparés d'une distance voisine de 5,5 Å¹⁸⁶.





Pour que ce résultat soit indiscutable, j'ai ensuite confirmé la structure 1-O-méthyl-Gal*f* formée par GC/MS. L'identification des formes individuelles de pyranose et de furanose était basée sur le rapport connu des quatre isomères du galactose²⁶⁷, les temps de rétention

chromatographiques ainsi que sur les profils de fragmentation en masse EI-MS de Per-TMS-1-O-métalytyl-galactosylé résidus²⁶⁸. De plus, la forme furane est caractérisée par un rapport m/z = 217, ce qui correspond au fragment Me₃SiOCH = CH-CH = OSiMe₃. Comme montré dans la figure 57, pour le signal à 25,99 min, le m/z est principalement de 217.

La protéine GlfH1 a été classée dans la famille GH5_13 dans la base de données CAZy. Les glycosidases de la famille GH5_13 agissent toutes par rétention. Le résultat expérimental obtenu ici correspond parfaitement à cette donnée théorique.



Figure 57 : Analyse GC / MS du produit de la réaction de GlfH1. A) Chromatogramme du galactose libre méthylé en C1 comme standard. B) chromatogramme de méthyl-galactoside formé par GlfH1 en utilisant *p*NP-Gal*f* comme substrat. C) Spectre de masse à impact électronique pour le signal à 25,99 min du chromatogramme A. D) Spectre de masse à impact électronique pour le signal à 25,99 min du chromatogramme B.

2.2 Identification des résidus catalytiques

J'ai ensuite cherché à déterminer la nature des chaines latérales impliquées dans le mécanisme catalytique. Pour cela les V_m et *K*_M à différents pH ont été calculés. En représentant le logVm/*K*_M en fonction du pH, le changement d'ionisation des chaînes latérales impliquées dans le mécanisme catalytique provoque un changement de pente de + 1 à 0 ou de 0 à - 1²⁶⁶. A la suite de ces expériences, j'ai observé un changement de pente de + 1 à 0 à pH 4,4 (Fig 58B), cela correspond au pK_a expérimental moyen d'une chaine latérale d'acide glutamique. C'est tout à fait logique car une glycosidase a toujours besoin de 2 chaines latérales acides dans le mécanisme catalytique. Mais j'ai aussi observé un deuxième changement de pente de 0 à -1 (Fig 58B) à pH 5,6, ce qui correspond au pK_a expérimental moyen d'une chaine latérale d'acide glutamique d'histidine. Ce résultat peut apparaître surprenant car il n'y a pas de glycoside hydrolase qui utilise directement un résidu d'histidine dans son mécanisme catalytique. Par contre, un résultat similaire a déjà observé dans le cas de l'endo-1,5-α-l-arabinanase BsArb43B de *Bacillus subtilis*²⁶³. Les auteurs proposent que ce résidu d'histidine servirait à stabiliser l'ion Ca²⁺ nécessaire au fonctionnement de cette enzyme, ce qui est tout à fait en accord avec les résultats obtenus l'étude de pour GlfH1.



Figure 58: A) Détermination des constantes *michaeliennes* de GlfH1 à différents pH. B) Représentation du log(Vm/ K_M) (déterminés en A) en fonction du pH. Un changement de pente (+1 à 0 ou 0 à -1) représente le changement d'ionisation d'un type de chaine latérale impliquée dans le mécanisme catalytique.

2.3 Détermination du site catalytique de GlfH1

2.3.1 Identification du site catalytique par alignement multiple de séquences

Grâce aux résultats obtenus sur l'alignement de séquences multiples de 13 glycosidases mycobactériennes différentes de la famille GH5_13, il est possible de proposer des résidus d'acides aminés candidats pour le site catalytique.

Parmi ces 13 séquences, 15 résidus d'acide aspartique et sept résidus d'acide glutamique sont conservés, et donc des candidats pour la paire de résidus d'acide catalytique nécessaires à l'activité de la glycosidase. De même, trois résidus d'histidine sont conservés et pourraient être impliqués dans la stabilisation d'un ion Ca²⁺ essentiel pour le fonctionnement hydrolytique de GlfH1, comme démontré dans la partie 1.7.

De plus, dans la famille GH5_13, la structure cristallographique de plusieurs glycosidases a été résolue et les deux résidus d'acides aminés catalytiques ont été identifiés^{195,269,270}. Il est à noter que les protéines de la famille GH5_13 présentent toujours deux résidus d'acide glutamique en tant qu'acides aminés catalytiques dans la base de données CAZy.

	170 	240 	270
GITHI (Mycobacterium tuberculosis)	WDLWNEPDNPA	DVITFHSYAA	
CsMan5 (Chrysonilia sitophila, PDB ID: 4AWE)	WQLANEPRPGN	DFGTMHLYPDW	PLVLEEYGMWT
RmMan5B (<i>Rhyzomucor miehei</i> , PDB ID: 4LYP)	WQIANEPQ	DYTTCHCWVEN	PIVMEEFGMAR
ManBK (Aspergillus niger, PDB ID: 3WH9)	WELANEPRCPS	DFGTLHLYPDS	PCLLEEYGVTS

Figure 59 : Alignement de séquences multiples de GlfH1 avec d'autres glycosidases de la famille GH5_13. Les étoiles indiquent les résidus catalytiques ou essentiels affectés expérimentalement pour l'activité de CmMan5A¹⁹⁵, CsMan5²⁷⁰, RmMan5B²⁶⁹ et ManKB.

À partir de là, un alignement multiple de MUSCLE avec la séquence de GlfH1 utilisant ces protéines permet de mettre en évidence deux candidats, E170 et E268, respectivement en tant que résidus nucléophiles acide/base et catalytiques (Fig 59). Le résidu H242 est également conservé parmi ces séquences. Ce résidu est essentiel pour l'activité de certaines GH5 glycosidases²⁷¹ et, dans le cas de GlfH1, pourrait servir à stabiliser l'ion Ca²⁺ nécessaire à son activité.

2.3.2 Modélisation par homologie de GlfH1

La modélisation par homologie de GlfH1 a surtout été réalisée afin de proposer des acides aminés candidats pour un rôle catalytique avant de tenter de les vérifier par mutagenèse dirigée. Cette modélisation par homologie a été réalisée avec l'aide de plusieurs services disponibles sur Internet : i-Tasser, Phyre2, Swiss-Prot-2 et Modeller (Fig 60). Les quatre modèles obtenus ont montré une architecture similaire en *TIM-barrel* α_8/β_8 avec un *rmsd* moyen de 1,767 Å (calculé avec UCSF Chimera v1.10.2). Comme il n'y a aucune information structurale concernant GlfH1, le fait que les quatre services employés montrent un repliement similaire est encourageant. De plus, l'analyse des *templates* retenus par ces services montre qu'ils ne sont constitués que de glycosides hydrolases et, très majoritairement, de la famille GH5. J'ai choisi de poursuivre ce travail avec le modèle fourni par i-Tasser (<u>http://zhanglab.ccmb.med.umich.edu/I-TASSER/</u>) qui a obtenu les meilleurs scores lors des cinq derniers *CASP Experiments*.



Figure 60 : Superposition des quatre modèles 3-D de GlfH1 obtenus avec i-Tasser (bleu), Phyre2 (beige), Modeller (vert) et Swiss-prot (lila). Le *rmsd* moyen entre les quatre structures est de 1,767 Å (calculé avec UCSF Chimera v1.10.2).

Le modèle obtenu avec i-Tasser a ensuite subi un affinage de sa structure par dynamique moléculaire avec le service FG-MD (<u>http://zhanglab.ccmb.med.umich.edu/FG-MD/</u>) pour obtenir le modèle de travail final.



Figure 61: Le modèle par homologie de GlfH1 s'inscrit parfaitement dans la mesure de volume obtenue par *Dynamic Light Scattering* (DLS). Les deux sphères reprèsentent le volume mesuré par DLS avec sa marge d'erreur (5.543 nm<dia.< 5.765 nm).

La taille de la protéine GlfH1 a été vérifiée par diffusion dynamique de la lumière. Notre modèle par homologie corrèle parfaitement avec la mesure du volume de GlfH1 par *Dynamic Light Scattering* (DLS) (Fig 61), ce qui confirme que la protéine GlfH1 est bien sous forme monomérique dans la solution de stockage.



Figure 62 : Modèle par homologie de GlfH1 obtenu avec i-Tasser et affiné par MG-MD. Sont indiqués les trois résidus retenus pour la détermination du site catalytique par mutagenèse dirigée.

Sur la base des résultats expérimentaux obtenus, les résidus candidats pour le site catalytique ont pu être proposés. Le couple acide/base doit être distant d'environ 5,5 Å car GlfH1 agit par rétention. Ensuite, un résidu d'histidine doit être proche pour stabiliser le calcium. De plus, les glycosides hydrolases en TIM barrel ont toutes un sites actif situé sur les boucles à l'extrémité Ct des brins β. Enfin, i'ai aussi utilisé COACh (http://zhanglab.ccmb.med.umich.edu/COACH/) qui permet une recherche automatique de ligands potentiels pour une structure 3-D de protéine donnée. Et comme pour toutes les protéines présentent une structure TIM barrel, le potentiel substrat monosaccharide a bien été placé dans les boucles à l'extrémité Ct des brins β de GlfH1. De là, j'ai pu proposer 3 résidus qui répondaient parfaitement à ce cahier des charges : E170 et E268 pour le couple acide/base et H242 pour la stabilisation du calcium comme montré dans la figure 62. (n.b. : la numérotation employée pour les mutations et la modélisation par homologie ont été réalisées sans le peptide signal de 27 résidus.)

2.3.3 Mutagenèse dirigée

Sur la base des résultats obtenus de modélisation par homologie avec i-Tasser, trois résidus d'acides aminés ont donc été retenus pour être mutés et permettent la caractérisation du site catalytique (E170, H242 et E268). Il a été choisi de muter les acides glutamiques en glutamine et l'histidine en alanine en première intention. En raison du taux de G+C élevé (66 %) chez les mycobactéries, les mutagenèses ont été réalisés en utilisant une PCR en gradient. Une fois les mutants obtenus, j'ai pu isoler les protéines correspondantes selon le même protocole que pour GlfhH1 *Wild Type* (WT). Ensuite, les constantes *michaeliennes* de ces mutants ont été déterminés pour être comparés à GlfH1 WT avec le logiciel Prism v7.0.

Tableau VI K_M et k_{cat} de GlfH1 et des cinq mutants GlfH1-E170Q, GlfH1-E170D, GlfH1-E268Q, GlfH1-E268D, etGlfH1-H242A.

	<i>K_M</i> (mM)	<i>k_{cat}</i> (s ⁻¹)	k_{cat}/K_M (M ⁻¹ .s ⁻¹)
GIfH1-WT	0,6	11	1,8 x 10 ⁴
GlfH1-E170Q	1,1	1	910
GlfH1-E170D	0,4	4 x 10 ⁻³	10
GlfH1-E268Q	6,8	2	290
GlfH1-E268D	0,9	2 x 10 ⁻³	2
GlfH1-H242A	19,2	1	52

Comme montré dans le tableau VI, pour les deux mutants GlfH1-E170Q et GlfH1-E268Q, on observe une diminution du k_{cat} de 10 fois et 5 fois. Ces premiers résultats sont encourageants car les protéines mutées obtenues ont bien conservé une activité catalytique mesurable prouvant qu'il s'agit quand même de protéines actives. Il faut ensuite à prouver que ces deux résidus sont bien les résidus catalytiques. Pour cela, les résidus d'acide glutamique ont été muté en acide aspartique pour éloigner la fonction carboxylique d'environ 1 Å pour obtenir une baisse significative de la constante catalytique sans toutefois la supprimer¹⁹⁶. Les mutants GlfH1-E170D et GlfH1-E268D ont monté une nette diminution de l'activité avec une diminution du k_{cat} de 2750 et 5500 fois respectivement. Ces résultats ont confirmé que les résidus d'acide glutamique en position 170 et 268 sont cruciaux pour l'activité enzymatique de GlfH1 et peuvent être considérés comme les résidus nucléophiles et acides/basiques.

Il reste maintenant à déterminer quel est le résidu d'acide glutamique responsable de l'attaque nucléophile sur le C1 du galactose parmi le résidu E170 et le résidu E268. Celui-ci peut être confirmé en le mutant en résidu d'alanine pour supprimer complètement leur fonction carboxylique, et en rajoutant de l'azoture dans le tampon de réaction. Parce que l'attaque nucléophile du premier résidu d'acide glutamique peut être complémentée par la présence d'azoture¹⁹⁶. Pour cela, les résidus d'acide glutamique ont été muté en résidus d'alanine pour augmenter la distance entre les présumés résidus catalytiques et le substrat, permettant de faciliter l'accès de l'azoture pour la complémentation de l'attaque nucléophile. Les mutants GlfH1-E170A et GlfH1-E268A ont été testé avec et sans azoture (Tableau VII).

Tableau VII K_M et k_{cat} de GlfH1 et des deux mutants GlfH1-E170A et GlfH1-E268A avec et sans la présenced'azoture.

	<i>K_M</i> (mM)	<i>k_{cat}</i> (s ⁻¹)	k_{cat}/K_M (M ⁻¹ .s ⁻¹)
GIfH1-WT	0,6	11	1,8 x 10 ⁴
Sans azoture			
GIfH1-E170A	-	-	-
GlfH1-E268A	0,4	0,01	25
Avec NaN ₃ 2M			
GlfH1-E170A	-	-	-
GlfH1-E268A	0,5	0,1	200

L'activité enzymatique du mutant E170A était trop faible pour être mesurée, les constantes *michaeliennes* ne pouvaient pas être calculées correctement. Ce résultat n'est pas surprenant car la mutation du résidu catalytique d'une glycosidase pourrait induire une diminution des constantes *michaeliennes de l'ordre de* 10^{-6} ¹⁸⁵. Le mutant GlfH1-E268A a montré une diminution du k_{cat} de 1100 fois. De plus, pour le mutant E268A, la présence de 2 M d'azoture de sodium dans le milieu réactionnel permet d'augmenter 10 fois le k_{cat} . Cependant, cet effet n'est pas observé pour le mutant E170A. Ces résultats sont encourageants, la complémentation de l'attaque nucléophile par l'azoture semble indiquer que le résidu E268 est responsable de l'attaque nucléophile sur le C1 du galactose.

J'ai également cherché à savoir si le résidu histidine en position 242 est bien impliqué dans le mécanisme catalytique de GlfH1. Pour cela, le résidu d'histidine 242 a été muté en un résidu

d'alanine. Une diminution du k_{cat} d'un facteur 10 est alors observée (Tableau III) pour le mutant GlfH1-H242A, montrant indiscutablement que le résidu d'histidine en position 242 est également participe au mécanisme catalytique de GlfH1 au-delà des résidus catalytiques attendus. Il reste maintenant à pouvoir prouver l'hypothèse de départ, à savoir que ce résidu d'histidine serait impliqué dans la stabilisation d'un ion calcium comme c'est le cas pour l'endo-1,5- α -L-arabinanase BsArb43B de *Bacillus subtilis*²⁶³. Pour cela, il faudra montrer que l'activité résiduelle de GlfH1-H242A est insensible à la présence de calcium.

Pour finaliser la caractérisation du site catalytique de GlfH1, il reste aussi à confirmer que le résidu E268 est bien responsable de l'attaque nucléophile par une expérience directe. Pour cela, Loïc Chêne de l'équipe du Pr. Stéphane P. Vincent a synthétisé le 2,4 dinitrophényl 2-deoxy-2-fluoro- β -D-galactofuranoside qui permettra d'identifier la position exacte du résidu d'acide aminé qui réalise l'attaque nucléophile¹⁹⁷ donc, dans le cas présent, confirmer le rôle du résidu E268 dans l'attaque nucléophile du C1 du galactofuranose (Fig 63).





Cette confirmation du premier résidu catalytique est actuellement en cours de réalisation. J'ai déjà pu montrer que le 2,4 dinitrophényl 2-deoxy-2-fluoro- β -D-galactofuranoside inactive de façon dose-dépendante GlfH1, ce qui est déjà un résultat intéressant car il confirme que GlfH1

est une ß-glycosidase agissant par rétention¹⁹⁷. Maintenant, il me reste à inactiver complétement un lot de GlfH1 pour pouvoir procéder à l'identification du résidu modifié par LC-MS/MS. La partie spectrométrie de masse sera réalisée par la plateforme P3M (Plateau Plateforme de Protéomique et Peptides Modifiés, resp. Jean-Michel Saliou) de l'institut Pasteur de Lille.

III. Activité de GlfH1

1. Activité de GlfH1 vis-à-vis de l'arabinogalactane synthétique

Si GlfH1 est une exoglycosidase spécifique de la chaîne galactane de l'arabinogalactane, elle aura à hydrolyser des liaisons β 1-5 et β 1-6. Pour savoir si elle est capable de cela, j'ai ensuite étudié l'activité enzymatique de GlfH1 sur un substrat synthétique qui mime la structure de son substrat naturel, c'est-à-dire la chaine galactane de l'arabinogalactane. Les oligosaccharides ont été synthétisés par l'équipe du Dr. Todd Lowary (Lowary Lab, University of Alberta, Canada) et sont constitués du linker disaccharidique α -L-Rhap-(1,3)- α -D-GlcpNAc suivi d'une chaine galactane de 5 à 11 galactofuranoses (C8-Linker-Galfx) respectant l'alternance de liaisons β 1-5 et β 1-6 (voir Fig 64).



Figure 64 : Structure des oligosaccharides synthétiques imitant la structure des chaînes galactanes de l'arabinogalactane utilisés comme substrats de l'exo- β -D-Galf-ase GlfH1. La nomenclature de *Lin-Galf_x* est utilisée avec x comme nombre total de résidus de Galf et *Lin* pour la présence du *linker* disaccharidique a-L-Rhap-(1,3)-a-D-GlcpNAc-C8.

1.1 Analyse HPAEC

Pour cette expérience, des hydrolyses de 0 à 120 h ont été réalisées sur les différents substrats avec ajout régulier de GlfH1 pour compenser son inactivation.

Comme présenté dans les figures 65-69 ci-dessous, l'hydrolyse partielle des oligosaccharides commence à être détectée après deux heures d'incubation avec tous les composés testés. Dans les 24 heures suivant l'incubation, les oligosaccharides initiaux ont presque complètement disparu en donnant des oligosaccharides de degré de polymérisation (DP) d'hexose inférieur. Ceci est illustré par exemple pour *Lin*-Gal f_{12} (Fig 69), qui a finalement généré des oligosaccharides aux petits DP (*Lin*-Gal f_{2-4}) en tant que produits principaux avec du 160 galactose libre après 120 heures d'incubation. Tous les oligosaccharides ont été dégradés, quelle que soit la nature du résidu Gal à l'extrémité non réductrice, que ce soit β -(1,5) ou β -(1,6). Sur les chromatogrammes obtenus par HPAEC des produits d'hydrolyse des différents oligosaccharides *Lin*-Galf_x (Fig 65-69), il arrive fréquemment que le pic soit partiellement dédoublé. Cependant l'analyse par spectrométrie de masse de ces composés n'a toujours monté qu'une seule masse moléculaire. Une hypothèse est que ces composés présentent des conformations distinctes avec une isomérisation lente, peut-être due à l'alternance de liaisons β -(1,5) et β -(1,6) possédant respectivement 2 et 3 degrés de liberté.



Figure 65 : Chromatogramme représentant l'hydrolyse de l'oligosaccharide synthétique *Lin*-Gal f_5 par GlfH1 de t= 0 h à t= 120 h analysé par HPAEC.



Figure 66 : Chromatogramme représentant l'hydrolyse de l'oligosaccharide synthétique *Lin*-Gal f_6 par GlfH1 de t = 0 h à t = 120 h analysé par HPAEC.



Figure 67 : Chromatogramme représentant l'hydrolyse de l'oligosaccharide synthétique *Lin*-Gal f_7 par GlfH1 de t = 0 h à t = 120 h analysé par HPAEC.



Figure 68 : Chromatogramme représentant l'hydrolyse de l'oligosaccharide synthétique *Lin*-Gal f_{11} par GlfH1 de t = 0 h à t = 120 h analysé par HPAEC.



Figure 69 : Chromatogramme représentant l'hydrolyse de l'oligosaccharide synthétique *Lin*-Gal f_{12} par GlfH1 de t = 0 h à t = 120 h analysé par HPAEC.

1.2 Confirmation par MALDI-TOF

La nature des oligosaccharides libérés a ensuite été analysée en spectrométrie de masse après perméthylation. Le spectre de masse a été acquis par MALDI-TOF comme indiqué sur la figure 70 pour l'hydrolyse de l'oligosaccharide *Lin*-Gal f_{12} . Cette analyse confirme que l'oligosaccharide *Lin*-Gal f_{12} a bien été hydrolysé par GlfH1 en générant l'oligosaccharide *Lin*-*Galf*₂ comme produit final après 120 heures d'incubation.





Ces résultats confirment la nature de GlfH1 en tant qu'exo-galactofuranohydrolase qui dégrade le galactane mycobactérien par le clivage séquentiel de résidus β -(1,5) ou β -(1,6) en alternance à partir de l'extrémité non réductrice de la molécule. Il faut aussi souligner que GlfH1 si est capable d'hydrolyser de manière alternative des liaisons β -(1,5) et β -(1,6), il en est de même pour la glycosyltransférase Gl*f*T2 qui a mis en place ces mêmes liaisons pour la biosynthèse de la chaine galactane¹⁴⁸.

Activité de GlfH1 sur son substrat naturel, l'arabinogalactane

GlfH1 étant capable d'utiliser le *p*NP-Gal*f* et les oligosaccharides *Lin*-Gal*f*₅₋₁₂ comme substrats, Il restait à prouver que l'arabinogalactane mycobactérien est bien le substrat naturel de GlfH1. Pour définir l'activité de GlfH1 vis-à-vis de ce substrat (Fig 71), de l'arabinogalactane, de l'AGP et du mAGP purifiés de *M. bovis* BCG ont été incubés avec ou sans GlfH1 et sous agitation pendant 48 heures en utilisant les conditions standards. GlfH1 a été ajouté régulièrement pour maintenir une hydrolyse significative. La réaction enzymatique a été arrêtée par chauffage à 100 °C pendant 5 min. Les monosaccharides libérés par GlfH1 ont été séparés par ultrafiltration en utilisant un filtre avec un seuil de coupure de 1 kDa avant d'être analysés par chromatographie à haute performance d'échange d'anions (HPAEC) et GC/MS.



Figure 71 : Représentation schématique de l'action de GlfH1 sur l'arabinogalactane de *M. tuberculosis.*

2.1 Analyse en HPAEC

L'analyse par chromatographie échange d'anion à haute performance (HPAEC) a été réalisée sur une colonne Carbopac[™] PA-1 (250 mm x 2 mm, Thermo Scientific) avec une pré-colonne de 50 mm (Thermo Scientific). La chromatographie a été ensuite réalisée grâce à un système Dionex ICS-5000 (Thermo Scientific) piloté par le logiciel Chromeleon version 7.1. La détection est réalisée grâce à un détecteur à ampérométrie pulsée (PAD) avec la vague de potentiel carbohydrates (standard quad) du fournisseur. La chromatographie est réalisée en mode isocratique avec [NaOH 100 mM] comme éluant.

Comme montré dans le chromatogramme HPAEC (Fig 72), la libération de résidus de Gal libres après le traitement au GlfH1 de l'arabinogalactane, AGP et mAGP a été clairement observée entre 5 et 7 minutes. Cependant, il apparaît que cette hydrolyse est plus rapide lorsque l'arabinogalactane est liée au PG (AGP). De plus, sa vitesse est moindre lorsque l'arabinogalactane est liée simultanément au PG et aux acides mycoliques (mAGP) (Fig 72). La vitesse d'hydrolyse AGP/arabinogalactane/mAGP a un rapport de 5,6 : 1 : 0,13.

Différentes vitesses d'hydrolyse de GlfH1 ont été observées entre l'arabinogalactane, l'AGP et le mAGP, l'AGP étant le meilleur substrat parmi les trois. Notre hypothèse est que l'arabinogalactane, par sa structure branchée inhabituelle, présente une structure compactée par des liaisons hydrogène²⁷² lorsqu'elle n'est plus liée aux acides PG et mycoliques, rendant ainsi plus difficile l'accès à la chaine galactane par GlfH1. Lorsque l'arabinogalactane est liée au peptidoglycane (AGP), moins de liaisons hydrogènes peuvent se former et l'arabinogalactane devient plus accessible à GlfH1. Au contraire, lorsqu'il est lié simultanément au PG et aux acides mycoliques (mAGP), l'accessibilité de l'arabinogalactane à GlfH1 est limitée par la présence de l'importante couche d'acides mycoliques.





2.2 Confirmations des monosaccharide libéré par GC/MS

Les monosaccharides libérés de l'arabinogalactane, de l'AGP et du mAGP ont été dérivés par per-triméthyl-silylation avant d'être analysés par GC/MS. Les spectres de masse obtenus pour les pics d'élution à 29,85 min et à 31,99 min (Fig 73A et 73C), présentent un m/z de 204. Cet ion 204 correspond à la fragmentation typique du galactose : Me₃SiOCH-CH = OSiMe₃. La nature du monosaccharide libéré sous forme du galactose a été confirmée pour toutes les conditions par analyse GC/MS (Fig 73-75).



Figure 73 : Analyse GC/MS sur le produit réactionnel de GlfH1. L'hydrolyse de l'arabinogalactane par GlfH1 a été réalisée en utilisant les conditions d'hydrolyse standard. A) chromatogramme de galactose libre en standard. B) chromatogramme de monosaccharide hydrolysé par GlfH1 en utilisant l'arabinogalactane comme substrat. C) Spectre de masse à impact électronique pour le signal à 31,99 min du chromatogramme A. D) Spectre de masse à impact électronique pour le signal à 31,98 min du chromatogramme B.



Figure 74 : Analyse GC/MS sur le produit réactionnel de GlfH1. L'hydrolyse de l'AGP par GlfH1 a été réalisée en utilisant les conditions d'hydrolyse standard. A) chromatogramme de galactose libre en standard. B) chromatogramme de monosaccharide hydrolysé par GlfH1 en utilisant AGP comme substrat. C) Spectre de masse à impact électronique pour le signal à 31,99 min du chromatogramme B.



Figure 75 : Analyse GC/MS sur le produit réactionnel de GlfH1. L'hydrolyse du mAGP par GlfH1 a été réalisée en utilisant les conditions d'hydrolyse standard. A) chromatogramme de galactose libre en standard. B) chromatogramme de monosaccharide hydrolysé par GlfH1 en utilisant mAGP comme substrat. C) Spectre de masse à impact électronique pour le signal à 31,99 min du chromatogramme B.

2.3 Analyse par RMN

Pour définir plus en détail l'activité de GlfH1 sur un substrat endogène mycobactérien, l'arabinogalactane soluble a été incubé avec et sans GlfH1 pendant 48 heures en utilisant des conditions standard. Les mélanges réactionnels ont été lyophilisés, échangés deux fois dans de l'oxyde de deutérium et soumis à des analyses par spectroscopie RMN ¹H et ¹³C à haut champ.



Figure 76 : Analyse RMN de l'activité de GlfH1. (A) Région anomérique de l'analyse 1D ¹H-RMN de l'arabinogalactane isolé de *M. bovis* BCG avant (panel du haut) et après (panel du bas) l'hydrolyse de GlfH1. Les spectres 1D ¹H-RMN montrent l'apparition concomitante des signaux α/β -Galp et la diminution de deux signaux (triangles bleus et rouges) identifiés comme -5- β -Galf et -6- β -Galf,. (B) Analyses RMN ¹H-¹H TOCSY et (C) ¹H-¹³C HSQC de l'arabinogalactane isolé de *M. bovis* BCG soumis au traitement par GlfH1.

Les signaux des monosaccharides principaux ont d'abord été identifiés sur l'arabinogalactane intact purifié de *M. bovis* BCG en analysant les spectres RMN 1D-¹H et 2D ¹H-¹H TOCSY et ¹H-¹³C HSQC (Fig 76) en comparaison avec les spectres standards obtenus au laboratoire et avec

la littérature²⁷³. Les spectres obtenus ont ensuite été comparés à ceux obtenus avec l'arabinogalactane de *M. bovis BCG* traité par GlfH1. Comme on le voit sur le spectre 1D¹H-RMN (Fig 76A), les résidus principaux dans l'arabinogalactane non traité ont été identifiés comme 2- α Araf, 5- α Araf et 3,5- α Araf et β Araf terminal pour le domaine arabinane et 5- β Galf et $6-\beta$ Galf pour le domaine galactane en accord avec la littérature. L'analyse RMN de l'arabinogalactane traitée par GlfH1 a montré un spectre très similaire exception faite de deux différences significatives. Tout d'abord, un signal H1 est apparu clairement à δ 4,57 (^{1,2}J_H, H = 7,8Hz) qui était associé à un signal C1 à δ 97,8 sur le spectre ¹H-¹³C HSQC (Fig 76C). Ce signal a été identifiée comme étant du β Galp libre selon les corrélations H/C (H2/C2, H3/C3 et H4/C4 à δ 3.48/73.2, 3.64/74.1 et 3.93/71.7, respectivement (spectre ¹H-¹³C HSQC entier non montrées) et les constantes de couplage observées sur le spectre ¹H-¹H TOCSY ($^{2,3}J_{H}$, H = 9,9Hz, ${}^{3,4}J_{H}$, H = 3,5Hz et ${}^{4,5}J_{H}$, H <1Hz) (Fig 76B). Un signal anomérique supplémentaire a été observé à 5,25 ppm couplé à un carbone à 93,6 ppm (Fig 76C), qui a été attribué à un résidu α Galp, bien que son système de spin n'ait pas pu être entièrement établi en raison de la faible intensité du signal et du chevauchement des signaux. La présence de ces signaux a confirmé la libération de Galp libre après l'hydrolyse de GlfH1. L'observation d'un β Galp en tant que produit principale et d' α Galp en tant que produit mineure et d'aucun Galf libre est en accord avec l'équilibre connu de Gal en solution (30/64/2.5/3.5 pour $\alpha p/\beta p/\alpha f/\beta f$ à 31 °C)²⁶⁷. De plus, l'apparition de signaux de Gal libres était concomitante avec la diminution spécifique de l'intensité des signaux anomères de 5- β Galf et 6- β Galf- à δ 5.02 et 5.23, respectivement sur le spectre 1D ¹H-RMN spectre (Fig 76A). La quantification relative de ces deux signaux par rapport à tous les autres signaux anomères dans les spectres 1D-¹H RMN montre une réduction globale de 25 % des deux signaux dans l'arabinogalactane traité par rapport à l'arabinogalactane non traité. Il est à noter que les intensités de signal de tous les résidus Ara n'ont montré aucune variation entre eux, confirmant en outre que la fraction arabinane de l'arabinogalactane reste intacte par GlfH1.

3. Hydrolyse sélective des liaisons β 1-5 et β 1-6 entre les résidus de Gal*f*

Les études précédentes m'ont permis de démontrer que GlfH1 est capable d'hydrolyser les oligosaccharides quelle que soit la nature de la liaison du résidu de galactose à l'extrémité non réductrice, que ce soit β -(1,5) ou β -(1,6). Il apparaît alors intéressant de déterminer si GlfH1 hydrolyse les liaisons β -(1,5) et β -(1,6) avec la même efficacité. Pour cela, en collaboration avec l'équipe du Pr. Kaoru Takegawa (Kyushu University, Japon), nous avons utilisé deux substrats synthétiques, le 2-aminopyridine-Galf- β -(1,5)-Galf (2PA-Galf_{β 1-5}Galf) et le 2-aminopyridine-Galf- β -(1,6)-Galf (2PA-Galf_{β 1-6}Galf) où les résidus de galactofuranose sont liés par des liaisons β -(1,5) ou β -(1,6). Ces deux substrats synthétiques ont été hydrolysés par GlfH1 puis analysés par chromatographie de partage à polarité de phase inversée.



Figure 77 : Chromatogramme représentant l'hydrolyse des substrats synthétiques 2PA-Gal f_{β_1-5} Galf et 2PA-Gal f_{β_1-6} Galf par 0 à 30 µg de GlfH1 avec un temps d'incubation de 60 min.

Il a d'abord fallu déterminer la quantité nécessaire de GlfH1 pour hydrolyser ces deux substrats. Pour cela, des hydrolyses de 60 min ont d'abord été réalisées sur les deux substrats avec 0 à 30 µg de GlfH1. Comme montré dans les figures 77 et 78 ci-dessous, l'hydrolyse de 2PA-Gal*f*_{β 1-5}Gal*f* peut être détectée en utilisant 0,3 µg d'enzyme, 50 % de substrat a été dégradé avec 3 µg d'enzyme et il a été entièrement dégradé par 30 µg de GlfH1, (Fig 77A et 78). Par contre pour le substrat 2PA-Gal*f*_{β 1-6}Gal*f*, seulement une très petite partie du substrat

a été hydrolysé en utilisant 0,3 μg d'enzyme, et 60 % de substrat est dégradé avec 30 μg de GlfH1 (Fig 77B et 78). Pour la suite d'analyse, la quantité de GlfH1 utilisée a été fixée à 3 μg.



Figure 78 : Représentations graphiques du pourcentage d'hydrolyse des substrats synthétiques par GlfH1 en fonction de la quantité de GlfH1 utilisée. La disparition du substrat 2PA-Gal $f_{\beta1-5}$ Galf est présentée en rouge et la disparition du substrat 2PA-Gal $f_{\beta1-6}$ Galf est présentée en bleu.

Ensuite, des cinétiques d'hydrolyse de 0 à 180 min ont été réalisées sur les différents substrats avec 3 µg de GlfH1. Comme montré dans les figures 79 et 80 ci-dessous, l'hydrolyse du 2PA-Gal $f_{\beta_{1-5}}$ Galf peut être détectée après 15 min d'incubation et 50 % du substrat a été dégradé après 180 min d'incubation (Fig 79A et 80). Cependant pour le substrat 2PA-Gal $f_{\beta_{1-6}}$ Galf, l'hydrolyse n'est qu'à partir de 30 min d'incubation et après 180 min d'incubation, seulement 16 % de la quantité initiale du substrat a été dégradée (Fig 79B et 80).



Figure 79 : Chromatogramme représentant l'hydrolyse des substrats synthétiques 2PA-Gal f_{β_1-5} Galf et 2PA-Gal f_{β_1-6} Galf par 3 µg de GlfH1 de t=0 à t=180 min

Cette analyse montre que la vitesse d'hydrolyse de GlfH1 est plus rapide pour une liaison β 1-5 que dans le cas d'une liaison β -(1,6). Cette différence de cinétique d'hydrolyse entre deux types de liaisons n'a pas souvent été observée, il a seulement été publié très récemment l'existence d'une β -glucosidase de *Lactobacillus antri*²⁷⁴ ayant une vitesse d'hydrolyse différente entre les liaisons β 1-6 et β 1-2.



Figure 80 : Représentations graphiques de la cinétique d'hydrolyse des substrats synthétiques par GlfH1. La disparition du substrat 2PA-Gal $f_{\beta_{1-5}}$ Galf est présentée en rouge et la disparition du substrat 2PA-Gal $f_{\beta_{1-6}}$ Galf est présentée en bleu.

Dans notre cas, les galactofuranoses liés en β 1-6 a un degré de liberté supplémentaire au niveau de la liaison osidique par rapport aux galactofuranoses liés en β 1-5, nous pouvons très bien supposer que cette différence de vitesse d'hydrolyse en soit la conséquence. Lorsque le site catalytique de GlfH1 sera complétement caractérisé (résidus catalytiques et géométrie), il sera possible de procéder à une étude de *docking* de substrats possédant des liaisons β -(1,5) ou β -(1,6) afin de déterminer l'origine moléculaire de cette différence.

IV. Étudier la fonction de GlfH1

Identification de du gène orthologue de *glfH1* chez *M. smegmatis*

Enfin pour aborder les aspects fonctionnels de l'activité de la protéine GlfH1 vis-à-vis de la physiologie des mycobactéries, nous avons cherché à construire un mutant de délétion du orthologue de *glfH1* chez *M. smegmatis*. Cette partie du travail a été réalisée en collaboration avec le Dr. Albertus Viljoen de l'équipe du Dr. Laurent Kremer (FRE3689, Montpellier) qui s'est chargé de la construction des souches mutantes de *M. smegmatis*.





Pour cela, une recherche BLAST limitée à *M. smegmatis* mc² 155 (taxid : 246196) a été effectuée. Le gène *MSMEG_5877* a été identifié comme l'orthologue de GlfH1 chez *M. smegmatis*. En effet, l'alignement global de la séquence MSMEG_5877 avec celle de GlfH1 présente une identité de 73,3 %, une similarité 82,1 % et une absence de *gap*, ce qui est plus que suffisant pour établir un lien d'homologie entre ces deux protéines(Fig 81). Les trois résidus d'acides aminés essentiels à l'activité de GlfH1 (Glu170, His242 et Glu268) sont parfaitement conservés.

Validation du gène orthologue de glfH1 chez M. smegmatis : MSMEG 5877

Pour valider la protéine MSMEG_5877 comme produit du gène orthologue de *Rv3096*, celleci a été exprimée dans *E. coli* et purifiée dans les mêmes conditions que GlfH1. Ses paramètres enzymatiques ont été ensuite déterminés en utilisant les conditions standards. Après avoir comparé les paramètres enzymatiques de MSMEG_5877 et de GlfH1 (Tableau VIII), les deux enzymes ont bien montré des caractéristiques similaires comme le pH optimal, les constantes *michaeliennes*, elles sont toutes les deux inhibées par éthambutol et ont besoin du calcium pour leur fonctionnement optimal.

Le gène *MSMEG_5877* pourrait donc être utilisé pour étudier la fonction physiologique de *GlfH1* en utilisant *M. smegmatis* mc² 155 comme modèle de travail.

MSMEG_5877	GlfH1
4,5	4,5
4	11
1165	18333
Inhibé	Inhibé
Essentiel	Essentiel
	MSMEG_5877 4,5 4 1165 Inhibé Essentiel

Tableau VIIIComparaison des proprités de pH optimum, constantes *michaeliennes*, inhibition par l'éthambutol eteffet du calcium entre MSMEG_5877 et de GlfH1.

Sur cette base, le Dr. Albertus Viljoen de l'équipe du Dr. Laurent Kremer a réalisé plusieurs souches mutantes de *M. smegmatis* correspondant à l'invalidation de *MSMEG_5877* (Δ *GlfH1*) et sa complémentation (Δ *GlfH1* complémentation).

3. Phénotype chez l'amibe

Le Dr. Viljoen a aussi étudié la croissance dans l'amibe de ces différents mutants. Pour cela, les cellules d'*Acanthamoeba polyphaga* ont été infectés par *M. smegmatis* de type sauvage, $\Delta glfH1$ et sa souche complémentée à une MOI de 10 et la croissance de la bactérie a été surveillée 2 heures après l'infection (hpi), 24 hpi et 48 hpi en lysant les cellules puis en plaçant les lysats sur milieu 7H10-OADC pour calculer le cfu/ml. Comme l'indique clairement la figure 82, l'inactivation de *MSMEG_5877* conduits à une diminution important de la croissance chez les amibes, présentant une charge bacillaire inférieure d'environ un demi log à celle de la souche de type sauvage à 48 hpi. De plus, une complémentation génétique de $\Delta glfH1$, conduit à une restauration partielle du phénotype d'infection *Wild Type*.



Figure 82 : La croissance des différents mutants dans l'amibe après 2 h, 24 h et 48 h.

Ces résultats sont prometteurs et proposent un rôle important pour GlfH1 lors de sa croissance chez l'amibe. Le fait que la suppression de *glfH1* (*MSMEG_5877*) indique un phénotype de croissance atténué pendant l'infection des amibes suggère fortement une implication de GlfH1 dans la résistance des mycobactéries aux mécanismes de protection de l'hôte pendant l'infection. Notre hypothèse est que, la délétion du gène *glfH1* pouvant éventuellement perturber l'intégrité de la paroi cellulaire mycobactérienne, entraînant ainsi une augmentation de sa perméabilité. Cette modification structurale permettrait aux molécules antibactériennes / digestives présentes dans l'amibe de pénétrer plus facilement. Cependant, pour confirmer cette hypothèse, une étude plus détaille sur le changement structural de la paroi cellulaire à différente conditions et phases avec la délétion du gène *glfH1*, le rôle de GlfH1 dans la croissance bactérienne et l'infection de *M. tuberculosis* à l'aide de modèles murin reste à faire.

4. Activité de MSMEG_5877 vis-à-vis de l'arabinogalactane de *M. smegmatis*

L'impact de l'expression de MSMEG_5877 sur la structure de l'arabinogalactane de *M. smegmatis* a aussi été abordé. Pour cela, l'arabinogalactane de *M. smegmatis* et des souches mutantes ont été purifiés pour être analysé. Ces arabinogalactanes ont ensuite été analysés par RMN mais aussi par chromatographie en phase gaz (CPG) après dérivation par triméthylsylilation pour déterminer leur composition.



Figure 83 : Impact de la délétion de GlfH1 sur la structure de l'arabinogalactane. Les signaux individuels des anomères RMN d'arabinogalactane ont été identifiés sur 1D ¹H-RMN en utilisant des expériences ¹H-¹³C HSQC (panneau supérieur) et ¹H-¹H COSY (panneau inférieur) utilisant l'arabinogalactane purifié de WT *M. smegmatis*.

Cependant, l'analyse de l'arabinogalactane extraite de WT, de $\Delta glfH1$ et de la souche de $\Delta glfH1$ complémentée a montré que la délétion de glfH1 n'avait pas d'impact significatif sur
la structure de l'arabinogalactane (Fig 83 et 84). Les analyses RMN 1D ¹H et 2D ¹H-¹H TOCSY et ¹H-¹³C HSQC ont permis d'identifier les signaux anomériques des résidus majeurs de l'arabinogalactane isolé de *M. smegmatis* WT : 2- α Araf, 5- α Araf et 3,5- α Araf et β Araf, 5- β Galf et 6- β Galf (Fig 83). Une fois identifiés sur le spectre 1D ¹H RMN, les aires de tous les signaux ont été mesurées, ce qui a permis de quantifier relativement chacun des résidus de l'arabinogalactane. Les quantifications relatives de tous les résidus des arabinogalactanes extraits de WT et de $\Delta glfH1$ ont été comparés (Fig 85) mais ne montrent pas de différence significative entre les trois échantillons. Ces résultats ont été confirmé par l'analyse de CPG sur la composition de l'arabinogalactane des trois échantillons.



Figure 84 : Spectres RMN 1D ¹H (900 MHz) obtenus sur les arabinogalactanes purifiés des souches WT, $\Delta glfH1$ et $\Delta glfH1$ complémentation de *M. smegmatis.*

Le fait qu'il y ait un phénotype d'infection lié à l'activité de GlfH1 sans que l'on observe de variation quantitative de la structure globale de l'arabinogalactane, suggère fortement que l'activité *in vivo* de GlfH1 doit être régulée spatialement et/ou temporellement et/ou ne s'adresse qu'à une fraction de l'arabinogalactane total.



Figure 85 : Quantifications des monosaccharides individuels sur l'arabinogalactane purifiée de souches WT, glfH1 délétée et complémentée en intégrant les signaux de RMN ¹H et de CPG. Les quantifications ont été effectuées sur trois expériences individuelles.

Partie IV : Conclusion

Pendant les vingt dernières années, une intention particulière a été portée aux glycosidases du point de vue thérapeutique : plusieurs inhibiteurs de glycosidases ont été découverts pour une utilisation potentielle dans le traitement du cancer, des infections virales, des maladies neurodégénératives, des maladies auto-immunes et du diabète²¹⁹. Cependant, malgré le fait que plusieurs glycosidases mycobactériennes soient déjà identifiées, leurs fonctions physiologiques restent encore mal connues.

Au cours de mes trois années de thèse, j'ai pu identifier la première galactosidase qui cible l'arabinogalactane des mycobactéries. La caractérisation de GlfH1 a été surprenante à plus d'un titre et montre qu'il s'agit d'une glycosidase aux caractéristiques inhabituelles pour cette fonction enzymatique. Tout d'abord, GlfH1 est incapable d'hydrolyser le $pNP-\beta$ -Dgalactopyranoside suggère que son activité enzymatique est spécifique de la forme furanose du monosaccharide. J'ai également démontré que GlfH1 est un exo-galactofuranosidase qui hydrolyse alternativement des liaisons β -(1,5) et β -(1,6). Aucune des β -galactofuranosidases caractérisés jusqu'à présent chez Aspergillus ou Penicillium n'a montré d'activité similaire sur les liaisons β -(1,5) et β -(1,6). De plus, en raison de la similarité structurelle entre le α -Larabinofuranose et le β -D-galactofuranose, de nombreuses α -L-arabinofuranosidases montrent également une activité de β -D-galactofuranosidase²⁷⁵. Par exemple, les β galactofuranosidases de chez Streptomyces présentent à la fois une activité de α-Larabinofuranosidasique et une activité β -D-galactofuranosidasique²⁶⁴. Contrairement à ces enzymes, GlfH1 ne montre aucune activité vis-à-vis de l'arabinofuranose. De plus, cette enzyme montre une nécessité absolue du calcium pour fonctionner, ce qui est une caractéristique exceptionnelle chez les glycosidases dans la mesure où seule une paire de résidus carboxyliques (Asp ou Glu) est demandée pour le mécanisme catalytique. Le besoin de calcium a été observé pour une endo- α -L-arabinanase de la famille GH43 chez *Bacillus* subtilis, plusieurs exo-mannosidases des familles GH38, GH47, GH92, ainsi qu'une α glucosidase et une α -galactosidase de la famille GH9^{190,263}. Cependant, aucune observation de ce type n'avait été montrée jusqu'à présent chez les β -galactofuranosidases.

Après avoir établi les paramètres clés du fonctionnement de GlfH1, j'ai également réalisé une étude préliminaire visant à identifier un inhibiteur de cette nouvelle enzyme, inhibiteur qui s'est montré être l'éthambutol (EMB). L'éthambutol est un médicament antituberculeux qui perturbe la synthèse de l'arabinogalactane en ciblant les arabinosyltransférases codées par les gènes *embA, embB* et *embC*^{160,163}. Cette molécule est également capable de perturber la fonction de plusieurs protéines effectrices, telles que le cytochrome²⁶⁵. Bien que l'éthambutol présente un K_i en mM, il pourra néanmoins servir de guide pour la découverte d'inhibiteurs effectifs de GlfH1.

L'analyse de la spécificité stéréochimique de GlfH1 montre clairement que GlfH1 agit par rétention. J'ai donc pu affirmer que la distance entre deux résidus catalytiques est d'environ 5,5 Å¹⁸⁶. Ce résultat a été très utile pour la réalisation d'un modèle 3D par homologie de GlfH1, qui a permis de proposer les résidus catalytiques. La mutagenèse dirigée a ensuite permis de prouver que trois résidus d'acides aminés étaient cruciaux pour l'activité de GlfH1. Les résidus E170 et E268 ont été montrés en tant que résidus catalytiques et le résidu H242 semble nécessaire pour stabiliser un ion Ca²⁺. Néanmoins, l'identité du résidu d'acide glutamique (E170 ou E268) qui provoque la première attaque nucléophile reste incertaine mais est en cours d'établissement par une méthode analytique directe ^{186,197}. L'hydrolyse du substrat synthétique qui mime la structure de la chaine galactane de l'arabinogalactane par GlfH1 montre que cette enzyme hydrolyse les oligosaccharides indépendamment de la nature du résidu galactofuranose à l'extrémité non réductrice. J'ai également montré que GlfH1 hydrolyse plus rapidement la liaison β -(1,5) que la liaison β -(1,6). De plus, des différences de vitesse d'hydrolyse par GlfH1 ont été observées en utilisant ses substrats naturels, c'est-à-dire l'arabinogalactane, l'AGP et le mAGP. Parmi les trois substrats, l'AGP est le substrat pour lequel GlfH1 hydrolyse le plus rapidement la chaine galactane. Le rôle de GlfH1 chez la mycobactérie a été abordé en étudiant le produit du gène orthologue (MSMEG_5877) chez *M. smeqmatis*. Et il apparaît que GlfH1 joue un rôle crucial pour la croissance de *M. smeqmatis* dans un modèle d'infection chez l'amibe. Cependant, une étude plus approfondie du rôle de GlfH1 dans la croissance de *M. tuberculosis* pendant l'infection à l'aide d'un modèle murin reste à réaliser.

L'identification de la première glycosidase impliquée dans la dégradation de l'arabinogalactane mycobacterien nous permet de supposer de l'existence d'une famille de glycosidases capables de dégrader et/ou modifier la structure des composés glycanniques associées à la paroi. À cette notion de remodelage structural peut aussi être associée celle de recyclage. En effet, sur la base de l'activité démontrée de GlfH1, il est maintenant possible de proposer une voie de recyclage du galactose libéré de la chaine galactane. En effet, celui-ci

187

peut tout à fait rejoindre la voie de Leloir et servir à reformer de l'UDP-D-galactofuranose nécessaire à la synthèse de la chaine galactane de l'arabinogalactane. Selon la voie KEGG du métabolisme du galactose chez les mycobactéries (https://www.genome.jp/keggbin/show pathway?map00052), plusieurs enzymes sont nécessaires dans la voie de Leloir pour recycler le galactose : la galactose mutarotase, la galactokinase Rv0620/galK (NP_215134.1), galactose-1-phosphate uridylyltransférase Rv0618/galTa (WP_003900189.1) et Rv0619/galTb (CCP43360.1), UDP-galactopyranose mutase Rv3809c/glf (NP_218326.1) et galactosidase GlfH1/*Rv3096* (NP_217612.1). Pour l'instant, il manquait un acteur dans cette voie métabolique, la galactose mutarotase. Et cette enzyme vient justement d'être identifiée chez *M. smegmatis* dans l'équipe du Dr. Yann Guérardel par le Dr. Alexandre Méry (résultat non publiée). Il reste maintenant à résoudre la question du re-import du galactose dans la bactérie. *M. smegmatis* possède des gènes candidats²⁷⁶ à cette fonction mais cela ne semble pas être le cas pour *M. tuberculosis*. Pour supporter cette hypothèse, le recyclage du peptidoglycane chez *M. tuberculosis* et chez *M. bovis BCG* a été montré très récemment. La lipoprotéine LpqI est une exo-β-N-acétylglucosaminidase pouvant cliver des fragments de peptidoglycane *in vitro*²¹³. La mise en évidence de ces deux enzymes mycobactériennes, LpqI et GlfH1, pose les premières bases d'un remodelage et d'un recyclage des constituants majeurs de la paroi cellulaire chez les mycobactéries.

Enfin, l'importance de GlfH1 dans un possible remodelage de la paroi cellulaire peut être confortée par l'étude de son environnement génomique. En effet, son gène est adjacent à *Rv3097*c, également nommé *lipY*, qui code pour une protéine bien caractérisée et exprimant une activité triacylglycérol (TAG) hydrolase. LipY participe à la consommation et au retraitement de ses propres lipides ainsi que des lipides hôtes, contribuant ainsi à la persistance du bacille tuberculeux dans les macrophages spumeux²⁷⁷⁻²⁷⁹. Il reste cependant à déterminer si les voies du TAG et du galactane sont connectées.

L'ensemble de ces travaux présentent l'identification d'une première glycosidase mycobactérienne importante pour la croissance de la bactérie chez l'hôte. Il reste cependant évident que de futures études seront nécessaires pour étudier la contribution potentielle de GlfH1 au cours de la division, de la virulence et de la persistance des cellules mycobactériennes.

188

Partie V : Matériels et Méthodes

I. Considérations générales

Pour toute préparation de tampons, substrats, réactifs et pour toute réaction enzymatique, de l'eau ultrapure avec une résistivité électrique $\geq 18.2 \text{ M}\Omega.\text{cm}$ (ELGA, Veolia Water STI, UK) est utilisée. Toutes les substances chimiques utilisées proviennent de Sigma-Aldrich (France) ou Acros Organics (France) selon disponibilité. Les colonnes HisTrapTM HP et HiTrapTM *Desalting* proviennent de GE Healthcare (USA). Le kit de dosage de protéines Micro BCATM pour la détermination de la concentration en protéines a été obtenu auprès de Thermo Fisher Scientifique (France). La trypsine de qualité spectrométrie de masse (porcine) est fournie par G-BioSciences (USA).

II. Production, purification et analyse de la protéine recombinante obtenue

1. Souches bactériennes, plasmides et milieux de culture

Les souches bactériennes HB101 (NEB), XL1-*blue* (NEB) et BL21 (DE3) (Invitrogen) *d'Escherichia coli* ont été utilisées respectivement pour le clonage et l'expression de protéines. Le vecteur d'expression choisi est le pET32a (Invitrogen).

Les milieux cultures utilisés sont répertoriés dans le tableau IX. Pour la sélection des clones recombinants, l'antibiotique utilisé est l'ampicilline à 50 µg/ml.

Milieu Auto-Induction (AIM)	Milieu LB	Milieu LB solide
Tryptone 10 g/L	Tryptone 10 g/L	Tryptone 10 g/L
Extrait de levure 5 g/L	Extrait de levure 5 g/L	Extrait de levure 5 g/L
(NH ₄) ₂ SO ₄ 3,3 g/L	NaCl 5 g/L	NaCl 5 g/L
KH2PO4 6,8 g/L		Agar 15 g/L
Na ₂ HPO ₄ 7,1 g/L		
Glucose 0,5 g/L		
α-Lactose 2,0 g/L		
MgSO4 0,15 g/L		

Tableau IX Composition des milieux de cultures utilisés pour la culture d'E. coli.

2. Purification des plasmides

Les vecteurs pET32a-Rv3096 et pET32a-MSMEG5877 construits par le Dr. Albertus Viljoen (*Mycobacterial Pathogenesis and Novel Therapeutic Targets*, CNRS-FRE3689, Montpellier) permettent l'expression hétérologue, chez *E. coli*, de la protéine GlfH1 et de la protéine MSMEG_5877 de *Mycobacterium tuberculosis*.

La purification des plasmides a été réalisée à l'aide du kit *Wizard[®] plus SV Minipreps DNA Purification System* (Promega) selon les recommandations du fournisseur. Ce kit permet d'extraire l'ADN plasmidique en petite quantité selon la méthode de lyse alcaline (adaptée de Birboim et Doly, 1979) basée sur la dénaturation différentielle de l'ADN génomique de structure relâchée et de l'ADN plasmidique super-enroulé. Les plasmides purifiés ont été quantifiés par spectrophotométrie à l'aide d'un spectrophotomètre NanoDrop 2000 (Thermo Fisher Scientific).

3. Expression de la protéine recombinante



Figure 86: Représentation schématique du vecteur pET32a (Invitrogen) : sont indiqués le promoteur T7 plaçant l'expression des gènes en aval sous contrôle de l'ARN polymérase du phage T7, la présence de la région opératrice lacO de l'opéron lacZYA répressible par LacI, un site de fixation au ribosome (RBS), puis la présence d'une étiquette thioredoxine (105 résidus d'acides aminés) permettant d'améliorer la solubilité des protéines. L'étiquette contient en outre une séquence de 6 histidines pour permettre la purification de la protéine sur colonne Ni-NTA (Nickel-Nitriloacetic Acid), une étiquette d'affinité S-Tag (15 résidus d'acides aminés) et un site reconnu par une entérokinase (EK) permettant, après purification de la protéine fusionnée, l'élimination de la séquence en amont du site. Un site multiple de clonage (MCS) contenant plusieurs sites de restriction nécessaire au clonage des gènes d'intérêt. Et enfin, le gène lacI permettant la synthèse du répresseur LacI.

Le plasmide pET32a (Invitrogen, Fig 86) permet chez *Escherichia coli* une expression contrôlée (induction lactose, ITPG) et en grande quantité, via l'ARN polymérase du phage T7, de gènes codant pour des protéines d'intérêt. En outre, afin de faciliter sa purification, la protéine

produite reste fusionnée à diverses étiquettes d'affinité (His-tag, S-tag) permettant la détection et l'isolement de la protéine. De même, la présence d'une étiquette Trx (thiorédoxine) particulièrement hydrophile permet d'augmenter la solubilité de la protéine recombinante pour limiter la formation de corps d'inclusion.

Le plasmide purifié à l'issue de l'étape précédente a ensuite été introduit par transformation dans la souche BL21DE3 d'*E. coli*. Pour les transformations réalisées dans cette thèse, la souche d'*E. coli* BL21 DE3 a été rendue compétente par traitement au chlorure de magnésium et congelée à -80 °C selon le protocole de Chung *et al.*, 1989²⁸⁰ avant d'être transformée par un choc thermique de 45 sec. A 42 °C par les plasmides purifiés. Les bactéries transformées ont ensuite été isolées sur milieu LB solide contenant 50 µg/ml d'ampicilline. Cette souche possède dans son génome le gène codant l'ARN polymérase du phage T7 qui permet la transcription des gènes intégrés sur le plasmide pET32a. L'expression de cette polymérase sous contrôle d'un promoteur *lacUV5* ne peut être réalisée qu'après levée de la répression appliquée par la protéine Lacl, par exemple, par ajout d'IPTG (Isopropyl β-D-1-thiogalactopyranoside) lors de la dernière génération de croissance en phase exponentielle. Après une pré-culture d'une nuit d'un clone isolé après transformation en milieu LB à 37°C sous agitation à 160 tours/min, l'expression de GlfH1 et de MSMEG5877 est réalisée non pas en conditions standards (37 °C et induction par l'IPTG) mais à 25 °C sous agitation à 160 tours/min grâce à un milieu auto-inducteur (AIM) et ce, pendant 16 h.

4. Extraction des protéines

Les culots issus des cultures de la souche BL21 d'*E. coli* transformée ont été repris dans 20 ml de tampon TNI [Tris 25 mM pH 7,5, NaCl 300 mM, imidazole 20 mM] contenant le cocktail d'inhibiteurs de protéases (*HaltTM Protease Inhibitor Cocktail, Thermo Scientific®*) et lysés par la Presse de French (SimoAminco) à 1650 psi. Les lysats ont ensuite été centrifugés à 15 000 g (rotor JA-20, Beckman), 4 °C pendant 30 min et les surnageants récupérés pour être soumis à chromatographie.

5. Purification de GlfH1 par chromatographie d'affinité sur métal immobilisé (IMAC)

La protéine GlfH1 a été isolée par chromatographie IMAC à l'aide d'un système Äkta Start (GE Healthcare LifeSciences) géré par le logiciel Unicorn v1.0 et une colonne d'affinité IMAC HisTrap HP 1 ou 5 ml (GE Healthcare LifeSciences). Les trois tampons utilisés pour la chromatographie sont le tampon d'équilibration A [tampon phosphate 20 mM, NaCl 500 mM, imidazole 40 mM, Tween 20 0,05 % (v/v), pH 7,4], le tampon de lavage A⁺ [tampon d'équilibration A, additionné d'ATP 5 mM et de MgCl₂ 10 mM, pH ajusté par NaOH à 8,5] et le tampon d'élution B [tampon phosphate 20 mM, NaCl 500 mM, imidazole 250 mM, Tween 20 0,05% (v/v), pH 7,4]. La purification débute par une première étape de chargement de l'extrait protéique brut sur la colonne réalisée à 1 ml/min. Une deuxième étape d'incubation est ensuite réalisée en utilisant le tampon de lavage A⁺ à température ambiante pendant 1 heure (le débit est interrompu pendant l'incubation dans le tampon) pour éliminer d'éventuelles protéines associées à GlfH1. Elle est suivie d'une troisième étape pour éluer la protéine avec 100% de tampon B à un flux de 1 ml/min. Les fractions contenant la protéine GlfH1 ont été recueillies et diluées 10 fois par le tampon A pour une étape de concentration de nouveau par chromatographie IMAC. Cette étape a été réalisée comme précédemment mais avec une colonne HisTrap HP de 1 ml (GE Healthcare LifeSciences) pour réduire le volume d'élution de GlfH1. Le suivi de l'élution des protéines est mesuré par variation de l'absorbance à 280 nm et le recueil de fractions fixé à 0,5 ml. La mesure de la conductivité est aussi utilisée pour optimiser les étapes d'équilibration.

6. Dessalage de GlfH1 par chromatographie de gel filtration

Les fractions majoritaires contenant GlfH1 ont été rassemblées pour être soumises à une chromatographie de gel filtration afin d'éliminer l'imidazole et autres composés interférents et obtenir la protéine dans le tampon de stockage choisi [Tris 40 mM, DTT 0,1 mM, CaCl₂ 100 mM, NaCl 150 mM, NaN₃ 0,1 %, LDAO 0,2 % (p/v), pH 7,4]. La chromatographie a été réalisée de nouveau à l'aide d'un système Äkta Start sur colonne HiTrap *Desalting* 5 ml (GE Healthcare LifeSciences) avec un volume d'injection de 1 ml, un débit de 5 ml/min et un suivi

d'élution par variation d'absorbance à 280 nm. La protéine GlfH1 est ensuite conservée à 4 °C dans le tampon d'équilibration de la gel filtration.

7. Dosage de protéine par la méthode BCA

La concentration des protéines a été déterminée par la méthode BCA en utilisant le kit *Micro* BCA^{TM} Protein Assay (Thermo Fisher Scientific, USA) en suivant le protocole proposé par le fournisseur. Après deux heures d'incubation à 37 °C, l'absorbance a été mesurée à 562 nm.

8. Electrophorèse en gel de polyacrylamide en présence de dodécyl sulfate de sodium (*SDS-PAGE*)

Ce type de gel permet de séparer les protéines en condition dénaturantes selon leur masse moléculaire apparente. La migration a été réalisée avec un gel de concentration avec T=4 % et un gel de séparation homogène avec T=10 % avec C=3,33 % pour les deux gels. Les échantillons sont solubilisés à raison de 25 % (v/v) final de tampon de Laemmli [4X : Tris-HCl 50 mM, pH 6,8; SDS 2 % (p/v); glycérol 10 % (v/v); bleu de bromophénol 0.0025 % (p/v); ß-mercaptoéthanol 5 % (p/v)]. Après migration, sous voltage constant de 180 V, la présence de protéines sur le gel est révélée par coloration colloïdale au bleu de Coomassie [*Instant Blue*, Expedeon] pendant une nuit. La décoloration est effectuée ensuite dans l'eau ultrapure.

9. Clonage de GlfH1

Le peptide signal de GlfH1 sélectionné par SignalP 4.1 a été éliminé par amplification PCR (*Polymerase Chain Reaction*). Le plasmide pET32a-Rv3096 (7028bp) a été utilisé comme matrice et les oligonucléotides choisis sont répertoriés dans le tableau X. Les sites de restriction utilisés sont *Nco*I et *Hin*dIII.

Tableau X Liste des amorces utilisées pour réaliser le clonage de GlfH1 sans peptide signa	al
---	----

Nom	Séquence nucléotidique de l'amorce
Rv3096_SP_For	5' TTATA CCATGG CTGATGAAGAACCAGGACGTTGGTCGGCCGACCGCGC 3'
Rv3096_SP_Rev	5' GTGGTGGTGCTCGAGTGCGGCCGCAAGCTTAATGC 3'

Les amplifications par PCR ont été réalisées avec la *Q5 Hot Start High Fidelity DNA polymerase* (NEB), une ADN polymérase haute-fidélité minimisant le risque d'erreurs. Elles ont été réalisées selon les recommandations du fournisseur et la composition du mélange réactionnel est détaillée dans le tableau XI. Le tableau XII présente le programme de PCR utilisé.

Les mélanges réactionnels ont subi 25 cycles d'amplification en gradient discontinu de T° (66 °C \pm 6 °C avec 3 °C/sec).

Le produit de PCR obtenu de même que 0,3 µg du vecteur pET32a sont ensuite soumis à digestion par les enzymes de restriction *Nco*I et *Hin*dIII (10 U pour chacune des enzymes) pendant 1 h et à 37°C. A l'issue de cette dernière, les fragments générés sont additionnés de tampon de charge (*Gel Loading Dye Purple* 6X (NEB)) puis soumis à électrophorèse sur gel d'agarose 0,8 % en tampon TAE (Tris-Acétate-EDTA). Après migration, la bande vectrice pET32a (5852 bp) et l'amplicon (1095 bp) correspondant au gène codant la protéine GlfH1 sans peptide signal, tous deux doublements digérés sont visualisés sous UV grâce au bromure d'éthidium contenu dans le gel. Les bandes sont alors découpées du gel et l'ADN extrait à l'aide du kit *QlAquick® Gel Extraction* (QIAGEN) selon les recommandations du fournisseur.

Tableau XIMélange réactionnel (par tube) utilisés pour la PCR.

Réactifs	Volumes (µl)
Plasmide pET32a-Rv3096 10 ng/μl	2,5
5X Q5 reaction buffer	10
dNTPs 10 mM	1
Amorce Sens 100 ng/µl	1,25
Amorce Antisens 100 ng/μl	1,25
5X Q5 High GC Enhancer	10
Q5 Hot Start High-Fidelity DNA Polymerase	0,5
Nuclease-Free Water	23,5

Étape	Température (°C)	Temps	Nombre de cycle
Dénaturation initiale	95	5 min	1
Dénaturation	95	30 sec	
Hybridation	66 \pm 6 avec 3 °C / sec	30 sec	25
Élongation	72	1 min 30 sec	
Élongation	72	3 min	1
Stockage	4	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	

Tableau XIIProgrammes de PCR pour le clonage de GlfH1.

Après avoir ajusté le *ratio* entre le vecteur pET32a et l'ADN correspondant à la protéine GlfH1 sans peptide signal suite à la migration en gel d'agarose, la ligation des 2 éléments a été effectuée par l'ajout au mélange d'1 U d'ADN ligase du phage T4 (Thermo Scientific) suivie d'une incubation 1 h à 22°C. Le produit résultant est ensuite chauffé 10 min à 65 °C pour inactiver l'enzyme avant de servir à la transformation de bactéries compétentes XL1-*blue d'E. coli*. A l'issue de cette dernière, les bactéries ont été étalées sur milieu LB solide contenant 50 µl d'ampicilline. Les colonies obtenues après 18 h de croissance à 37 °C ont alors été utilisées pour ensemencer des cultures de 5 ml en milieu LB contenant 50 µg/ml d'ampicilline. Après une nuit à 37°C, les cellules ont été collectées et les plasmides ont été isolés à l'aide du kit *Wizard® plus SV Minipreps DNA Purification System* (Promega). Les plasmides purifiés ont été quantifiés par spectrophotométrie à l'aide d'un spectrophotomètre NanoDrop 2000 (Thermo Fisher Scientific) et envoyés en séquençage pour confirmer leur séquence d'ADN.

10. Mutagenèse dirigée

Sept mutations acides aminés différentes ont été réalisées indépendamment dans la séquence codante du gène de la protéine GlfH1, résultant en sept substitutions d'acide aminé : E170Q, E170D, E170A, H242A, E268Q, E268D et E268A. Concrètement, le plasmide pET32a-

Rv3096 (7028 bp) a été utilisé comme matrice pour réaliser les PCR. Les oligonucléotides utilisés sont répertoriés dans le tableau XIII.

Mutation de GlfH1	Amorces			
	nom	séquence		
E170Q	E170Q for	5' GGACCTGTGGAATCAACCCGACAATCC 3'		
	E170Q rev	5' GGATTGTCGGGTTGATTCCACAGGTCC 3'		
E170D	E170D for	5' CCTGTGGAATGACCCCGACAATCCC 3'		
	E170D <i>rev</i>	5' GGGATTGTCGGGGTCATTCCACAGG 3'		
E170A	E170A for	5' GACCTGTGGAATGCACCCGACAATCC 3'		
	E170A <i>rev</i>	5' GGATTGTCGGGTGCATTCCACAGGTC 3'		
H242A	H242A for	5' CGTGATCACCTTCGCCAGTTACGCCGCGC 3'		
	H242A rev	5' GCGCGGCGTAACTGGCGAAGGTGATCACG 3'		
E268Q	E268Q for	5' CAATCCTGTGCACCCAGTACCTGGCGCG 3'		
	E268Q rev	5' CGCGCCAGGTACTGGGTGCACAGGATTG 3'		
E268D	E268D for	5' CCTGTGCACCGACTACCTGGCGCG 3'		
	E268D <i>rev</i>	5' CGCGCCAGGTAGTCGGTGCACAGG 3'		
E268A	E268A for	5' CCTGTGCACCGCGTACCTGGCGC 3'		
	E268A rev	5' GCGCCAGGTACGCGGTGCACAGG 3'		

Tableau XIII Liste des amorces utilisées pour réaliser les mutations ponctuelles.

Les PCR ont été réalisées une fois de plus avec la *Q5 Hot Start High Fidelity DNA polymerase* (NEB) selon les recommandations du fournisseur. La composition du mélange réactionnel est détaillée dans le tableau XIV. Le tableau XV présente le programme de PCR utilisé.

Tableau XIVMélanges réactionnels utilisés pour les PCR

Réactifs	Volumes (µl)
Plasmide pET32a-Rv3096 10 ng/µl	2,5
5X Q5 tampon réactionnel	10
dNTPs 10 mM	1
Amorce Sens 100 ng/µl	1,25
Amorce Antisens 100 ng/µl	1,25
5X Q5 High GC Enhancer	10
Q5 Hot Start High-Fidelity DNA Polymerase	0,5
Nuclease-Free Water	23,5

Les mélanges réactionnels ont ensuite subi 25 cycles de PCR (*Polymerase Chain-Reaction*) en gradient (66 °C \pm 10 °C) pour amplifier le plasmide. Parallèlement des témoins négatifs pour

chacune des mutations d'AA envisagées ont été réalisés en omettant la polymérase dans le mélange réactionnel.

Afin de vérifier que les PCR ont bien permis d'amplifier le vecteur plasmidique dans son entier, une quantité aliquote de l'amplicon obtenu dans chacune des conditions a été déposée sur gel d'agarose 0,8 %, après ajout de tampon de charge (*Gel Loading Dye Purple* 6X (NEB)), et soumise à migration électrophorétique. Pour chacune des mutations désirées, l'ADN obtenu à l'issue de la PCR correspondant aux meilleures conditions d'amplification est ensuite incubé avec 10 U d'enzyme de restriction *Dpn*I pendant 1h30 à 37 °C afin de digérer l'ADN matriciel présent dans le mélange réactionnel. Les témoins négatifs évoqués précédemment sont également soumis à l'action enzymatique de *Dpn*I pour confirmer une digestion totale de l'ADN matriciel.

Des bactéries *E. coli* XL1-*blue* chimio-compétentes ont alors été transformées par 2,5 µl du mélange réactionnel résultant pour chacune des mutations envisagées et par une quantité équivalente de mélange réactionnel du témoin négatif correspondant. A l'issue de la transformation, les cellules sont étalées sur milieu LB solide contenant 50 µl d'ampicilline. Après 18 h de croissance à 37 °C, un contrôle des boîtes de Pétri ensemencées avec les témoins négatifs est réalisé. Aucune colonie bactérienne n'est présente, en principe, sur ces dernières. Les colonies obtenues ont pu alors être utilisées pour ensemencer des cultures de 5 ml de LB liquide contenant 50 µg/ml d'ampicilline. Après une nuit à 37 °C, les cellules ont été collectées et les plasmides ont été isolés à l'aide du kit *Wizard® plus SV Minipreps DNA Purification System* (Promega). Les plasmides purifiés ont été quantifiés par spectrophotométrie à l'aide d'un spectrophotomètre NanoDrop et envoyées en séquençage à Eurofins Scientific pour confirmer la présence de la mutation désirée.

Étape	Température (°C)	Temps	Nombre de cycle	
Dénaturation	95	30 sec	1	
initiale				
Dénaturation	95	30 sec		
Hybridation	65 ± 10	1 min	25	
Élongation	72	5 min 30 sec		
Élongation	72	6 min	1	
Stockage	4	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		

Tableau XV	Programmes	de PCR	pour la	a mutagenèse	dirigée d	e GlfH1
------------	------------	--------	---------	--------------	-----------	---------

Après séquençage, la souche BL21(DE3) d'*E. coli* a été transformée par les plasmides recombinants et les bactéries étalées sur milieu LB solide contenant 50 μ l/ml d'ampicilline. Des colonies isolées ont été utilisées pour la production de protéines comme mentionné précédemment.

11. Empreinte peptidique massique (*Peptide Mass Fingerprinting*, PMF)

Après séparation par SDS-PAGE, les bandes de gel contenant les protéines d'intérêt ont été découpées puis digérées à la trypsine pour être analysées et identifiées par spectrométrie de masse. Pour cela, les gels sont préalablement décolorés plusieurs fois pendant 20 min dans $0,5 \text{ ml d'un mélange acétonitrile (ACN)/bicarbonate d'ammonium 50 mM (Bic) 50 : 50 (v/v).$ Les bandes prélevées ont ensuite été déshydratées pendant 10 min dans l'ACN 100 %, puis séchées par dessiccation sous vide (Concentrator 5301, Eppendorf) à température ambiante. Les échantillons ont ensuite été réduits par ajout de 0,5 ml de DTT 20 mM (DTT 20 mM dans Bic 50 mM) pendant 1 h à 56 °C. L'alkylation a été réalisée en rajoutant 0,5 ml d'iodoacétamide (iodoacétamide 100 mM dans Bic 50 mM) pendant 45 min à température ambiante sans lumière et en agitant vigoureusement toutes les 10 min. Après retrait de la solution d'iodoacétamide, les échantillons ont subi trois lavages consécutifs de 10 min par 0,5 ml des solutions Bic 50 mM, ACN/Bic 50 mM 50 : 50 (v/v) et ACN 100 %. Ces lavages ont permis d'éliminer l'iodoacétamide résiduel. Ils ont ensuite été séchés par dessiccation sous vide (Concentrator 5301, Eppendorf) à T°. La digestion des protéines par la trypsine a été réalisée en rajoutant 20 µl de solution de trypsine (Promega) pendant 1 h à température ambiante pour réhydrater les bandes de gel avec la solution de trypsine. La concentration initiale de la trypsine est de $1 \mu g/\mu l$ et elle a été utilisée diluée au $1/50^{e}$ dans le tampon de digestion [NH₄HCO₃ 40 mM/ACN 10 % (v/v)] pour obtenir une concentration de travail de 20 ng/ml. Ensuite 300 µl de tampon de digestion ont été ajoutés aux échantillons et ces derniers ont été incubés pendant une nuit à 37 °C. Après quoi, 150 μ l d'H₂O ultrapure ont été ajoutés et les échantillons incubés pendant 10 min à 37 °C avec agitation vigoureuse. Ensuite, les surnageants ont été récupérés et les bandes du gel ont été extraites deux fois par 50 µl de de la solution [ACN 70 %/ATFA 0,1 % (v/v)] sous agitation pendant 1 h à température ambiante. Les extraits rassemblés ont ensuite été séchés par dessiccation sous vide pendant 6 h. Les peptides ont ensuite été purifiés et concentrés grâce une pointe ZipTip[®] (Millipore) préalablement préparée par deux lavages avec 10 μ l d'ACN 100 % et trois lavages avec 10 μ l d'ATFA 0,1 % (v/v). Les échantillons secs ont été resolubilisés en ajoutant 10 μ l d'ATFA 0,1 % (v/v), puis chargés sur les pointes ZipTip[®] par 4-5 aspiration/refoulement. Après avoir été vidée de tout liquide, la pointe ZipTip[®] a été lavée par 10 μ l d'ATFA 0,1 % (v/v) pour éliminer les contaminants. Les peptides ont enfin été élués avec la matrice acide α -cyano-4-hydroxycinnamique (CHCA) [10 mg/ml dans 2,5 μ l d'ACN 70 %/ ATFA 0,1 % (v/v)] et déposés directement sur une cible MALDI, chauffée à 50 °C jusqu'à cristallisation sur la plaque.

Les peptides purifiés sont analysés par spectrométrie de masse MALDI-TOF (spectromètre AB Sciex ABI 4800) avec 3000 tirs/spot et une gamme de la masse mesurée de 500 à 4000 m/z avec la moyenne fixée à 2150 *m/z*.

À partir d'un spectre de masse contenant les rapports *m/z* de chaque peptide du mélange, les spectres mono-isotopiques ont été sélectionnés par le logiciel de traitement du spectromètre (AB Sciex *Data Explorer*) puis comparés la banque de données NCBInr via MASCOT (<u>http://www.matrixscience.com</u>). En prenant en compte les informations suivants :

- Banque de donnée utilisé pour la recherche : NCBIprot
- Enzyme utilisée pour le clivage : trypsine
- Clivage manqués autorisés : 1
- Taxonomie utilisée pour la recherche : Mycobacterium tuberculosis complexe
- Modification fixée : carbamidométhylation
- Tolérance autorisée entre les masses théoriques et expérimentales : ± 1,2 Da
- Valeurs de masse : MH⁺

III. Analyse bioinformatique

Les séquences de la sous-famille GH5_13 de la banque de données CAZy (<u>http://www.cazy.org</u>) ont été extraites de leur génome de référence de la base de données GenBank (<u>https://www.ncbi.nlm.nih.gov/genome/</u>). Pour toutes les séquences protéiques utilisées, la présence d'un peptide signal a été recherchée avec le serveur SignalP 5.0 ²⁵⁵ et le peptide signal éventuel a été exclu de la séquence. De là, les séquences protéiques ont été numérotées

sans le peptide signal. Les résultats obtenus lors de la consultation des bases de données ont été respectivement obtenus avec la version 230 (février 2019) pour le NCBI (https://www.ncbi.nlm.nih.gov), version 2019_02 (13 février 2019) pour UniProt (https://www.uniprot.org), version du 20 février pour la DGBP de la RCSB (http://www.rcsb.org) et version de février 2019 pour CAZy (http://www.cazy.org). Le logiciel bioinformatique Unipro UGENE v1.32 (février 2019) a été utilisé pour la gestion de des données de séquence, les alignements multiples et l'analyse phylogénique²⁸¹. Les alignements multiples ont été réalisés en utilisant l'algorithme MUSCLE avec les paramètres par défaut pour une précision optimale. L'analyse de la phylogénie a été réalisée par la méthode PHYLIP *Neighbor Joining*²⁸² avec la matrice de substitution PMB Henikoff/Tillier, gamma comme taux de distribution sur tous les sites (coefficient de variation du taux de

substitution entre sites de 0,50) et un *bootstrap* de 100 répétitions avec règle de majorité (règle étendue) comme type de consensus. La méthode PhyML *Maximum Likehood*²⁸³ a également été utilisée pour la confirmation avec la matrice de substitution Dayhoff et un *bootstrap* de 100 répétitions.

IV. Caractérisation de l'activité enzymatique de GlfH1

Typiquement, l'activité enzymatique de GlfH1 a été déterminée en utilisant un substrat synthétique, le $pNP-\beta$ -D-Galactofuranose (pNP-Galf), dans 100 µl de solution tampon (variable) à 37 °C. La réaction est ensuite arrêtée par 50 µl d'une solution de carbonate de sodium 0,5 M. L'absorbance est alors mesurée à 405 nm. Le substrat pNP-Galf a été synthétisé par le Dr. Sydney Villaume de l'équipe du Pr. Stéphane P. Vincent à partir du D-galactose commercial.

V. Modélisation par homologie

La modélisation par homologie de GalfH1 a été réalisée avec l'aide du serveur i-TASSER ²⁸⁴. Ensuite, le modèle de premier plan a été affiné en utilisant le serveur de dynamique moléculaire FG-MD²⁸⁴ afin d'améliorer la géométrie locale de la structure, de supprimer les conflits stériques et d'améliorer l'angle de torsion et les réseaux de liaison à l'hydrogène.

VI. Préparation de l'arabinogalactane de *Mycobacterium bovis BCG*

1. Culture de M. bovis BCG

Mycobacterium bovis BCG (souche Pasteur) a été cultivée en milieu Sauton [Pour 1L de culture : KH_2PO_4 0,5 g ; 0,5 g de MgSO_4 ; 4 g de L-asparagine ; 2 g d'acide citrique ; 50 mg de citrate d'ammonium ferrique ; 60 mL de glycérol ; 100 uL de ZnSO₄ à 1 % p/v, ajuster le pH à 7. Stériliser à 120 °C pendant 10 min. Ajouter 2,5 mL de Tween 80 20 % (v/v), qsp à 1L avec de l'eau ultrapure, stérilisé par une membrane avec un seuil de coupure de 0,22 µm] à 37 °C sans agitation. Les mycobactéries ont été récupérés en phase stationnaire (A₆₀₀ \approx 3), pour cela, la culture est centrifugée à 4000 g pendant 10 min à 4 °C, puis lavé par du tampon PBS 1X. Le culot récupéré a conservé à -20 °C pour la purification de l'AG.

2. Purification de l'arabinogalactane de Mycobacterium bovis

BCG

La purification de l'arabinogalactane à partir de *M. bovis BCG* a été réalisée en se basant sur la méthode publiée par Gurdyal S. Besra ²⁸⁵.

2.1 Préparation du complexe mAGP

A partir d'environ 5-8 g de culot humide, 20 ml de tampon de lyse [2 mM DTT, 2 mM MgCl₂ dans PBS 1X] avec inhibiteur de protéase (*HaltTM Protease Inhibitor Cocktail, Termo Scientific®*) et lysé par la Presse de French (SimoAminco) à 1650 psi. L'opération est répétée une seconde fois. Le lysat a ensuite été centrifugé à 14 000 tr.min⁻¹ (rotor JA-20, Beckman) à +4 °C pendant 30 min et le culot C1 récupéré. Le culot C1 est ensuite repris dans 20 ml du tampon d'extraction [2 % (v/v) Triton X-100 dans PBS 1X], incubé à 4 °C pendant une nuit sous agitation douce. L'échantillon est centrifugé à 10 000 tr.min⁻¹ (rotor JA-20, Beckman) à +4 °C pendant 30 min. Le surnageant S2 contient les LAMs, LMs et PIMs. Le culot C2 est récupéré, il a ensuite été extrait trois fois pendant 1 h à 95 °C par un tampon d'extraction [2 % (v/v) SDS dans PBS 1X],

puis centrifugé à 10 000 tr.min⁻¹ (rotor JA-20, Beckman) à 20 °C pendant 20 min. Le matériel insoluble récupéré a ensuite subi ensuite trois lavages consécutifs par 20 ml de :

- eau
- 80 % acétone / 20 % eau (v/v)
- acétone

Des centrifugations ont été réalisées entre chaque lavage (10 000 tr.min⁻¹ avec le rotor JA-20, Beckman, 4 °C, 15 min) pour éliminer le surnageant. Après la lyophilisation du culot, le mAGP purifié est récupéré.

2.2 Purification de l'AG

Le mAGP est ensuite soumis à une saponification par une solution de KOH 2 % (p/v) dans le toluène/méthanol 1: 1 à 70 °C sous reflux pendant 48 h et sous agitation. Après une centrifugation à 10 000 tr.min⁻¹ (rotor JA-20, Beckman) à 20 °C pendant 30 min, le culot récupéré et il correspond à l'AGP. Puis, il a été repris dans une solution de NaOH 2 M pour effectuer une lyse alcaline à 80 °C pendant 16 h sous agitation. Centrifugé à nouveau en utilisant les mêmes conditions qu'avant. Le culot correspond a l'AG, il a été dialysé pendant 72 h en utilisant une membrane avec un seuil de 1 kDa. L'échantillon est lyophilisé pour récupérer l'AG purifié. Pour 5 g de culot humide de mycobactéries, un moyen d'environ 60 mg de l'AG peuvent être obtenu.

VII. Conditions hydrolyse par GlfH1

1. Mode d'action de GlfH1

La stéréochimie de l'activité de l'enzyme GlfH1 (rétention ou inversion) a été déterminée par chromatographie en phase gazeuse. 100 μ g de *p*NPGal*f* ont été hydrolysés par GlfH1 en présence de méthanol à 20 % (v/v) à 37 °C pendant 24 h, en utilisant le tampon citrate 20 mM pH 4,5, puis lyophilisés avant d'être dérivés par per-triméthyl-silylation (TMS).

2. Hydrolyse d'oligosaccharides synthétiques par GlfH1

L'hydrolyse des oligosaccharides synthétiques *C8-Linker-Galf_x* a été réalisée à 37 °C pendant une durée comprise entre 0 et 120 h avec 15 µg de substrat, 10 µg de GlfH1 dans le tampon citrate de sodium 20 mM avec 20 mM CaCl₂ à pH 4,5 sur un volume total de 100 µl, 10 µg de GlfH1 ayant été rajoutés toutes les 2 h pendant 8 h, puis toutes les 24 heures afin d'optimiser l'hydrolyse. Après incubation, les échantillons ont été chauffés à 100°C pendant 5 min pour dénaturer totalement l'enzyme puis centrifugés à 10 000 g pendant 2 min. Les surnageants ont été recueillis pour l'analyse HPAEC-PAD et MALDI-TOF.

3. Hydrolyse des polysaccharides naturels par GlfH1

Les polysaccharides naturels (environ 1 mg de chaque) ont été hydrolysés par GlfH1 dans le tampon citrate 20 mM avec 20 mM CaCl₂ à pH 4,5 à 37 °C pendant 48 heures sous agitation, la protéine GlfH1 a été ajouté toutes les 24 heures afin d'optimiser l'hydrolyse. Après incubation, les échantillons ont été chauffés à 100 °C pendant 5 min et centrifugés à 10 000 g pendant 2 min. Les monosaccharides libérés par le GlfH1 ont été séparés par ultrafiltration avec un filtre de 1 kDa, des fragments <1 kDa ont été récupérés pour des analyses HPAEC-PAD et GC / MS.

VIII. Méthodes analytiques

1. Analyse HPAEC-PAD

Pour analyser les monosaccharides ou les oligosaccharides libérés par GlfH1, une chromatographie échangeuse d'anions à haute performance a été réalisée à l'aide d'un système Dionex ICS5000 HPAEC (contrôlé par le logiciel Chromeleon version 7.1.) avec détection ampérométrique à impulsions (PAD) (Thermo Scientific). Les séparations ont été effectuées sur une colonne CarboPac PA1 (2 mm × 250 mm, Thermo Scientific) équipée d'une colonne de garde (2 mm × 50 mm, Thermo Scientific) pour les monosaccharides ou sur une colonne de x 250 mm, Thermo Scientific) équipée d'une colonne de garde (3 mm × 50 mm, Thermo Scientific) pour les monosaccharides. Les séparations de monosaccharides ont été réalisées en mode isocratique dans soude 100 mM à 0,25 ml.min⁻¹.

Les séparztions d'oligosaccharides ont été réalisées aussi dans la soude 100 mM et avec le gradient d'acétate de sodium 1 M présenté dans le tableau XVI. La détection a été réalisée par ampérométrie pulsée avec la vague de potentiel *Quadruple Standard Quad (carbohydrates)* fournie par le constructeur.

Temps (min)	NaOH 100 mM	Acétate de sodium 1 M	Débit (ml.min ⁻¹)
0 - 58	100 %	0 %	0,25
58 - 88	60 %	40 %	0,25
88 - 95	20 %	80 %	0,25
95 - 115	100 %	0 %	0,25
115 - 117	100 %	0 %	0,05

Tableau XVIProgramme d'élution sur colonne CarboPac PA200 pour la séparation d'oligosaccharides par HPAEC-
PAD.

2. Analyse par chromatographie en phase gazeuse (CPG)

2.1 Dérivation par triméthylsilylation (TMS)

Les monosaccharides hydrolysés par GlfH1 après lyophilisation ont été dérivés par triméthylsilylation (TMS). Pour la silylation, 20 µl de pyridine et 20 µl de Bis-Silyl-TriFluoroAcétamide (BSTFA) ont été ajoutés à l'échantillon qui est ensuite incubé à T° ambiante pendant 2 h, séché sous azote et finalement repris dans 100 µl d'heptane pour injecter en GC.

2.2 Analyses en GC-FID et GC/MS

Les monosaccharides dérivés par la méthode TMS ont d'abord été analysés par chromatographie en phase gazeuse couplée à un détecteur à ionisation de flamme (GC-FID) (CPG Thermo Scientific Trace GC Ultra) avec colonne Uptibond 1 premium (Interchim, longeur 30 m, diamètre interne 0,25 mm, épaisseur de la phase stationnaire 0,25 µm). Pour les analyses en GC / MS, un chromatographe en phase gazeuse (Thermo Scientific Trace GC Ultra) et une colonne capillaire SolGel-1ms[™] (0,25 mm 30 m, film 0,25 µm, SGE) ont été utilisés en couplage avec un spectrométre de masse (TSQ Quantum GC, Thermo Scientific).

Le programme de température utilisé pour les monosaccharides dérivés au TMS était de 120 °C à 240 °C avec une pente de 2 ° C / min ; maintien à 240 °C pendant 10 min.

3. Analyse spectromètre de masse MALDI-TOF

Les oligosaccharides hydrolysés ont été perméthylés selon la méthode de Ciucanu et Kerek 1984²⁸⁶, pour l'analyse par spectrométrie de masse.

À partir des substrats perméthylés préalablement dissous dans 50 µl d'acétonitrile, 1 µl d'échantillon a été prélevé pour être mélangé directement sur la plaque de MALDI avec 1 µl de solution de matrice superDHB mg/ml dans l'acétonitrile 50 % v/v, NaOH 1 mM). Les spectres MALDI-TOF ont été réalisés en utilisant un spectromètre de masse *4800 Proteomics Analyzer* (*Applied Biosystems*, Framingham, MA, USA) en mode positif avec *reflectron* et extraction retardée en utilisant un mode d'accélération de 20 kV, un retard d'impulsion de 200 ns et une tension de grille de 66 %. Trois mille tirs laser ont été effectués et accumulés pour chaque spectre.

4. Préparations des échantillons pour la Résonance magnétique nucléaire (RMN)

Les échantillons ont été lyophilisés, échangés deux fois dans de D₂O (99,96 % ²H, Eurisotop[®]) et soumis à des analyses par spectroscopie RMN ¹H et ¹³C à haut champ. Tous les échantillons ont été placés dans des tubes de 5 mm adaptés au D₂O. De l'acétone a été ajoutée en tant qu'étalon interne, à partir d'une solution de 2,5 μ l d'acétone dans 10 ml de D₂O. Les échantillons ont été ensuite analysés par la RMN ¹H et ¹³C à champ élevé.

Les analyses RMN sont acquises sur plusieurs spectromètres différents selon les expériences :

- un spectromètre Bruker Avance II 400 MHz (Bruker, Billerica, Massachusetts, USA) équipé d'une sonde BBO de 5 mm résonant à 400 MHz pour le ¹H et 100,6 MHz pour le ¹³C;
- un spectromètre Bruker Avance III 800 MHz équipé d'une sonde BBI résonant à 800
 MHz pour le ¹H et à 201,2 MHz pour le ¹³C;
- un spectromètre Bruker Avance III 900 MHz avec une sonde BBO résonant à 900 MHz
 pour le ¹H et à 226,3 MHz pour le ¹³C.

Annexes





Annexe : Protocole utilié pour la purification de l'arabinogalactane mycobactérien.

Bibliographie

- 1. World Health Organization. *Global Tuberculosis Report 2018*. (2018).
- 2. Pai, M. et al. Tuberculosis. Nat Rev Dis Primers 2, 845–23 (2016).
- Sotgiu, G. *et al.* Do we need a new Fleming e poque: The nightmare of drugresistant tuberculosis. *International Journal of Mycobacteriology* 2, 123–125 (2013).
- 4. Ludwig, W. *et al.* Road map of the phylum Actinobacteria. *Bergey's Manual® of Systematic Bacteriology* 57, 1–28 (2012).
- Goodfellow, M. & Williams, S. Ecology of actinomycetes. *Annu. Rev. Microbiol.* 37, 189–216 (1983).
- Barka, E. A. *et al.* Taxonomy, Physiology, and Natural Products of Actinobacteria.
 Microbiol. Mol. Biol. Rev. 80, 1–43 (2015).
- Ahmed, A. *et al.* Comparative Genomic Analyses of 17 Clinical Isolates of Gardnerella vaginalis Provide Evidence of Multiple Genetically Isolated Clades Consistent with Subspeciation into Genovars. *Journal of Bacteriology* 194, 3922– 3937 (2012).
- Zimmermann, W. Degradation of lignin by bacteria. *Journal of Biotechnology* 13, 119–130 (1990).
- 9. Bérdy, J. Bioactive microbial metabolites. *J. Antibiot.* 58, 1–26 (2005).
- 10. Ventura, M. *et al.* Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. *Microbiol. Mol. Biol. Rev.* 71, 495–548 (2007).
- 11. Girard, G. *et al.* A novel taxonomic marker that discriminates between morphologically complex actinomycetes. *Open Biology* 3, 130073–130073 (2013).
- Gao, B. & Gupta, R. S. Phylogenetic Framework and Molecular Signatures for the Main Clades of the Phylum Actinobacteria. *Microbiol. Mol. Biol. Rev.* 76, 66–112 (2012).
- Euzéby, J. P. List of Bacterial Names with Standing in Nomenclature: a folder available on the Internet. *International Journal of Systematic Bacteriology* 47, 590–592 (1997).
- Burkovski, A. The role of corynomycolic acids in Corynebacterium-host interaction.
 Antonie Van Leeuwenhoek 1–9 (2018).

- 15. Aubert, D. A formal analysis of phylogenetic terminology: towards a reconsideration of the current paradigm in systematics. *Phytoneuron* (2015).
- 16. Marrakchi, H., Lanéelle, M.-A. & Daffé, M. Mycolic Acids: Structures, Biosynthesis, and Beyond. *Chemistry & Biology* 1–19 (2013).
- Parte, A. C. LPSN—list of prokaryotic names with standing in nomenclature.
 Nucleic Acids Research 42, D613–D616 (2013).
- Gupta, R. S., 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.* 9, 2095–41 (2018).
- 19. Velayati, A. A. & Farnia, P. *Atlas of Mycobacterium Tuberculosis*. (Academic Press, 2016).
- 20. Pieters, J. & Mckinney, J. D. *Pathogenesis of Mycobacterium tuberculosis and its Interaction with the Host Organism*. (Springer Science & Business Media, 2013).
- 21. Tientcheu, L. D. *et al.* Immunological consequences of strain variation within the Mycobacterium tuberculosiscomplex. *Eur. J. Immunol.* 47, 432–445 (2017).
- Garnier, T. *et al.* The complete genome sequence of Mycobacterium bovis. *PNAS* 100, 7877–7882 (2003).
- 23. Luca, S. & Mihaescu, T. History of BCG Vaccine. *Maedica (Buchar)* 8, 53–58 (2013).
- 24. Forrellad, M. A. *et al.* Virulence factors of the Mycobacterium tuberculosiscomplex. *Virulence* 4, 3–66 (2014).
- Porvaznik, I., Solovič, I. & Mokrý, J. Non-Tuberculous Mycobacteria: Classification,
 Diagnostics, and Therapy. *Adv. Exp. Med. Biol.* 944, 19–25 (2017).
- 26. Society, A. T. Diagnosis and treatment of disease caused by nontuberculous mycobacteria. This official statement of the American Thoracic Society was approved by the Board of Directors, March 1997. Medical Section of the American Lung Association. *American journal of respiratory and critical care medicine* 156, S1–25 (1997).
- Wolinsky, E. Nontuberculous mycobacteria and associated diseases. *Am. Rev. Respir. Dis.* 119, 107–159 (1979).
- van Ingen, J., Boeree, M. J., van Soolingen, D. & Mouton, J. W. Resistance mechanisms and drug susceptibility testing of nontuberculous mycobacteria. *Drug Resistance Updates* 15, 149–161 (2012).
- 29. Koh, W.-J., Kwon, O. J. & Lee, K. S. Nontuberculous mycobacterial pulmonary diseases in immunocompetent patients. *Korean J Radiol* 3, 145–157 (2002).
- King, G. M. Uptake of Carbon Monoxide and Hydrogen at Environmentally Relevant Concentrations by Mycobacteria. *Applied and Environmental Microbiology* 69, 7266–7272 (2003).
- Li, Q., Ge, F., Tan, Y., Zhang, G. & Li, W. Genome-Wide Transcriptome Profiling of Mycobacterium smegmatis MC2 155 Cultivated in Minimal Media Supplemented with Cholesterol, Androstenedione or Glycerol. *IJMS* 17, 689–17 (2016).
- 32. Reyrat, J. M. & Kahn, D. Mycobacterium smegmatis: an absurd model for tuberculosis? *Trends Microbiol.* 9, 472–474 (2001).
- Nunes-Alves, C. *et al.* In search of a new paradigm for protective immunity to TB.
 Nat Rev Micro 12, 289–299 (2014).
- 34. Cosma, C. L., Sherman, D. R. & Ramakrishnan, L. The Secret Lives of the Pathogenic
 Mycobacteria. *Annu. Rev. Microbiol.* 57, 641–676 (2003).
- 35. Russell, D. G. Who puts the tubercle in tuberculosis? *Nat Rev Micro* 5, 39–47 (2006).
- 36. Dheda, K., SCHWANDER, S. K., ZHU, B., van ZYL-SMIT, R. N. & Zhang, Y. The immunology of tuberculosis: From bench to bedside. *Respirology* 15, 433–450 (2010).
- Tailleux, L. *et al.* DC-SIGN Induction in Alveolar Macrophages Defines Privileged
 Target Host Cells for Mycobacteria in Patients with Tuberculosis. *PLoS Med* 2, e381–11 (2005).
- Kang, P. B. *et al.* The human macrophage mannose receptor directs Mycobacterium tuberculosislipoarabinomannan-mediated phagosome biogenesis. *J Exp Med* 202, 987–999 (2005).
- Tailleux, L. *et al.* DC-SIGN Is the Major Mycobacterium tuberculosisReceptor on
 Human Dendritic Cells. *J Exp Med* 197, 121–127 (2003).

- 40. Geijtenbeek, T. B. H. *et al.* Mycobacteria Target DC-SIGN to Suppress Dendritic Cell Function. *J Exp Med* 197, 7–17 (2003).
- 41. Kaufmann, S. H. How can immunology contribute to the control of tuberculosis? *Nature Reviews Immunology* 1, 20–30 (2001).
- 42. Kaufmann, S. H. E. Protection against tuberculosis: cytokines, T cells, and macrophages. *Ann. Rheum. Dis.* 61 Suppl 2, ii54–8 (2002).
- 43. Schwander, S. K. *et al.* T lymphocytic and immature macrophage alveolitis in active pulmonary tuberculosis. *J. Infect. Dis.* 173, 1267–1272 (1996).
- Schwander, S. K. *et al.* Enhanced responses to Mycobacterium tuberculosis antigens by human alveolar lymphocytes during active pulmonary tuberculosis. *J. Infect. Dis.* 178, 1434–1445 (1998).
- 45. Herrera, M. T. *et al.* Compartmentalized bronchoalveolar IFN-γ and IL-12 response in human pulmonary tuberculosis. *Tuberculosis* 89, 38–47 (2009).
- 46. Pawlowski, A., Jansson, M., Sköld, M., Rottenberg, M. E. & Källenius, G. Tuberculosis and HIV Co-Infection. *PLoS Pathog* 8, e1002464–7 (2012).
- 47. Keane, J. *et al.* Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N Engl J Med* 345, 1098–1104 (2001).
- 48. Newport, M. J. *et al.* A mutation in the interferon-gamma-receptor gene and susceptibility to mycobacterial infection. *N Engl J Med* 335, 1941–1949 (1996).
- 49. Pai, M., Zwerling, A. & Menzies, D. Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. *Ann. Intern. Med.* 149, 177–184 (2008).
- 50. Pai, M. *et al.* Gamma Interferon Release Assays for Detection of Mycobacterium tuberculosis Infection. *Clinical Microbiology Reviews* 27, 3–20 (2014).
- 51. Pai, M. & Sotgiu, G. Diagnostics for latent TB infection: incremental, not transformative progress. *Eur. Respir. J.* 47, 704–706 (2016).
- 52. Kik, S. V., Denkinger, C. M., Chedore, P. & Pai, M. Replacing smear microscopy for the diagnosis of tuberculosis: what is the market potential? *Eur. Respir. J.* 43, 1793–1796 (2014).
- 53. Saeed, M. *et al.* GeneXpert technology. A breakthrough for the diagnosis of tuberculous pericarditis and pleuritis in less than 2 hours. *SMJ* 38, 699–705 (2017).

- 54. World Health Organization. The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs: policy guidance. (2016).
- 55. World Health Organization. Molecular line probe assays for rapid screening of patient at risk of multidrug-resistant tuberculosis (MDR-TB). (2008).
- 56. David, H. L. Probability distribution of drug-resistant mutants in unselected populations of Mycobacterium tuberculosis. *Appl Microbiol* 20, 810–814 (1970).
- 57. Joshi, J. Tuberculosis chemotherapy in the 21 stcentury: Back to the basics. *Lung India* 28, 193–9 (2011).
- 58. Nath, H. & Ryoo, S. in *Tuberculosis Current Issues in Diagnosis and Management* 1–19
- 59. Banerjee, A. *et al.* inhA, a gene encoding a target for isoniazid and ethionamide in
 Mycobacterium tuberculosis. *Science* 263, 227–230 (1994).
- 60. Ramaswamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber. Lung Dis.* 79, 3–29 (1998).
- 61. Telenti, A., Imboden, P., Marchesi, F., Schmidheini, T. & Bodmer, T. Direct, automated detection of rifampin-resistant Mycobacterium tuberculosis by polymerase chain reaction and single-strand conformation polymorphism analysis. *Antimicrobial Agents and Chemotherapy* 37, 2054–2058 (1993).
- 62. Zhang, Y. & Mitchison, D. The curious characteristics of pyrazinamide: a review.*Int. J. Tuberc. Lung Dis.* 7, 6–21 (2003).
- 63. Tubercle, D. M. *The action of antituberculosis drugs in short-course chemotherapy*. *Elsevier* (1985).
- 64. Konno, K., Feldmann, F. M. & McDermott, W. Pyrazinamide susceptibility and amidase activity of tubercle bacilli. *Am. Rev. Respir. Dis.* 95, 461–469 (1967).
- 65. Zhang, Y. & Yew, W.-W. Mechanisms of drug resistance in Mycobacterium tuberculosis: update 2015. *Int. J. Tuberc. Lung Dis.* 19, 1276–1289 (2015).
- 66. Zhang, Y., Scorpio, A., Nikaido, H. & Sun, Z. Role of acid pH and deficient efflux of pyrazinoic acid in unique susceptibility of Mycobacterium tuberculosis to pyrazinamide. *Journal of Bacteriology* (1999).

- 67. Cheng, S. J., Thibert, L., Sanchez, T., Heifets, L. & Zhang, Y. pncA mutations as a major mechanism of pyrazinamide resistance in Mycobacterium tuberculosis: spread of a monoresistant strain in Quebec, Canada. *Antimicrobial Agents and Chemotherapy* 44, 528–532 (2000).
- 68. Lemaitre, N., Sougakoff, W., Truffot-Pernot, C. & Jarlier, V. Characterization of new mutations in pyrazinamide-resistant strains of Mycobacterium tuberculosis and identification of conserved regions important for the catalytic activity of the pyrazinamidase PncA. *Antimicrobial Agents and Chemotherapy* 43, 1761–1763 (1999).
- 69. Takayama, K. & Kilburn, J. Inhibition of synthesis of arabinogalactan by ethambutol in Mycobacterium smegmatis. *Antimicrobial Agents and Chemotherapy* (1989).
- 70. Wolucka, B. A., McNeil, M. R., de Hoffmann, E., Chojnacki, T. & Brennan, P. J. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. J. Biol. Chem 269, 23328–23335 (1994).
- 71. Telenti, A. *et al.* The emb operon, a gene cluster of Mycobacterium tuberculosis involved in resistance to ethambutol. *Nat Med* 3, 567–570 (1997).
- He, L. *et al.* ubiA (Rv3806c) encoding DPPR synthase involved in cell wall synthesis is associated with ethambutol resistance in Mycobacterium tuberculosis.
 Tuberculosis 1–6 (2015).
- 73. Inverstigation, B. M. R. C. Streptomycin treatment of pulmonary tuberculosis.
 British Medical Journal 1–18 (1948).
- 74. Finken, M., Kirschner, P., Meier, A., Wrede, A. & Böttger, E. C. Molecular basis of streptomycin resistance in Mycobacterium tuberculosis: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Molecular Microbiology* 9, 1239–1246 (1993).
- 75. Cooksey, R. C., Morlock, G. P., McQueen, A., Glickman, S. E. & Crawford, J. T. Characterization of streptomycin resistance mechanisms among Mycobacterium tuberculosis isolates from patients in New York City. *Antimicrobial Agents and Chemotherapy* 40, 1186–1188 (1996).

- 76. Okamoto, S. *et al.* Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. *Molecular Microbiology* 63, 1096–1106 (2007).
- 77. Nachega, J. B. & Chaisson, R. E. Tuberculosis drug resistance: a global threat. *Clin. Infect. Dis.* 36, S24–30 (2003).
- Frieden, T. R. *et al.* The emergence of drug-resistant tuberculosis in New York City.
 N Engl J Med 328, 521–526 (1993).
- 79. Ben Amor, Y., Day, M. S. & Schluger, N. W. Preventing the next generation of extensively drug-resistant tuberculosis. *Int. J. Tuberc. Lung Dis.* 14, 525–527 (2010).
- Villarino, M. E., Geiter, L. J. & Simone, P. M. The multidrug-resistant tuberculosis challenge to public health efforts to control tuberculosis. *Public Health Rep* 107, 616–625 (1992).
- 81. Mishra, A. K., Driessen, N. N., Appelmelk, B. J. & Besra, G. S. Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in Mycobacterium tuberculosisphysiology and host–pathogen interaction. *FEMS Microbiol Rev* 35, 1126–1157 (2011).
- 82. Daffé, M. The cell envelope of tubercle bacilli. *Tuberculosis* 95, S155–S158 (2015).
- Alderwick, L. J., Harrison, J., Lloyd, G. S. & Birch, H. L. The Mycobacterial Cell Wall—Peptidoglycan and Arabinogalactan. *Cold Spring Harb Perspect Med* 5, a021113–16 (2015).
- 84. Jarlier, V. & Nikaido, H. Permeability barrier to hydrophilic solutes in Mycobacterium chelonei. *Journal of Bacteriology* 172, 1418–1423 (1990).
- 85. Kaur, D., Guerin, M. E., Skovierova, H., Brennan, P. J. & Jackson, M. *Chapter 2 Biogenesis of the Cell Wall and Other Glycoconjugates of Mycobacterium tuberculosis*. *Advances in Applied Microbiology* 69, 23–78 (Elsevier Inc., 2009).
- Puffal, J., García-Heredia, A., Rahlwes, K. C., Siegrist, M. S. & Morita, Y. S. Spatial control of cell envelope biosynthesis in mycobacteria. *Pathog Dis* 76, 938 (2018).
- 87. Daffé, M. The cell envelope of tubercle bacilli. *Tuberculosis* 1–4 (2015).
- 88. Barry, C. E., Crick, D. C. & McNeil, M. R. Targeting the formation of the cell wall core of M. tuberculosis. *Infect Disord Drug Targets* 7, 182–202 (2007).

- 89. Schleifer, K. H. & Kandler, O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* 36, 407–477 (1972).
- 90. Brennan, P. J. & Nikaido, H. The envelope of mycobacteria. *Annu. Rev. Biochem.*64, 29–63 (1995).
- 91. Mahapatra, S., Scherman, H., Brennan, P. J. & Crick, D. C. N Glycolylation of the Nucleotide Precursors of Peptidoglycan Biosynthesis of Mycobacterium spp. Is Altered by Drug Treatment. *Journal of Bacteriology* 187, 2341–2347 (2005).
- 92. Raymond, J. B., Mahapatra, S., Crick, D. C. & Pavelka, M. S., Jr. Identification of the namHGene, Encoding the Hydroxylase Responsible for the N-Glycolylation of the Mycobacterial Peptidoglycan. J. Biol. Chem. 280, 326–333 (2004).
- 93. Chatterjee, D., Bozic, C. M., McNeil, M. & Brennan, P. J. Structural features of the arabinan component of the lipoarabinomannan of Mycobacterium tuberculosis.
 J. Biol. Chem. 266, 9652–9660 (1991).
- 94. Lavollay, M. *et al.* The Peptidoglycan of Stationary-Phase Mycobacterium tuberculosis Predominantly Contains Cross-Links Generated by L,D-Transpeptidation. *Journal of Bacteriology* 190, 4360–4366 (2008).
- 95. McNeil, M., Daffé, M. & Brennan, P. J. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J. Biol. Chem.* 265, 18200–18206 (1990).
- 96. Daffé, M., Brennan, P. J. & 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.* 265, 6734–6743 (1990).
- 97. Misaki, A. & Yukawa, S. Studies on cell walls of Mycobacteria. II. Constitution of polysaccharides from BCG cell walls. *J. Biochem.* 59, 511–520 (1966).
- 98. Azuma, I., Kimura, H., Niinaka, T., Aoki, T. & Yamamura, Y. Chemical and Immunological Studies on Mycobacterial Polysaccharides I. Purification and Properties of Polysaccharides from Human Tubercle Bacilli. *Journal of Bacteriology* 95, 263–271 (1968).

- 99. Kanetsuna, F. Chemical analyses of mycobacterial cell walls. *Biochim. Biophys. Acta* 158, 130–143 (1968).
- 100. Kanetsuna, F., Imaeda, T. & Cunto, G. On the linkage between mycolic acid and arabinogalactan in phenol-treated myobacterial cell walls. *Biochim. Biophys. Acta* 173, 341–344 (1969).
- 101. Besra, G. S. *et al.* A new interpretation of the structure of the mycolylarabinogalactan complex of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and ¹H nuclear magnetic resonance spectroscopy. *Biochemistry* 34, 4257–4266 (1995).
- 102. Bhamidi, S. *et al.* The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of Mycobacterium tuberculosis mycolyl arabinogalactan. *J. Biol. Chem.* 283, 12992–13000 (2008).
- 103. Jankute, M., Cox, J. A. G., Harrison, J. & Besra, G. S. Assembly of the Mycobacterial Cell Wall. *Annu. Rev. Microbiol.* 69, 405–423 (2015).
- 104. Bhamidi, S. *et al.* Detailed structural and quantitative analysis reveals the spatial organization of the cell walls of in vivo grown Mycobacterium leprae and in vitro grown Mycobacterium tuberculosis. *J. Biol. Chem.* 286, 23168–23177 (2011).
- 105. Jackson, M. The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med* 4, (2014).
- 106. Asselineau, C. & Asselineau, J. Trehalose-containing glycolipids. *Prog Chem Fats Other Lipids* 16, 59–99 (1978).
- 107. Goren, M. in *Handbook of Lipid Research* 6, (Bergey's Manual[®] of Systematic Bacteriology, 1990).
- 108. Liu, J., Barry, C. E., III, Besra, G. S. & Nikaido, H. Mycolic Acid Structure Determines the Fluidity of the Mycobacterial Cell Wall. *J. Biol. Chem.* 271, 29545–29551 (1996).
- Takayama, K., Wang, C. & Besra, G. S. Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis. *Clinical Microbiology Reviews* 18, 81–101 (2005).

- 110. Watanabe, M., Aoyagi, Y., Ridell, M. & Minnikin, D. E. Separation and characterization of individual mycolic acids in representative mycobacteria. *Microbiology* 147, 1825–1837 (2001).
- 111. Watanabe, M. *et al.* Location of functional groups in mycobacterial meromycolate chains; the recognition of new structural principles in mycolic acids. *Microbiology* 148, 1881–1902 (2002).
- 112. Qureshi, N., Takayama, K., Jordi, H. C. & Schnoes, H. K. Characterization of the purified components of a new homologous series of alpha-mycolic acids from Mycobacterium tuberculosis H37Ra. *J. Biol. Chem.* 253, 5411–5417 (1978).
- 113. Dubnau, E. *et al.* Oxygenated mycolic acids are necessary for virulence of Mycobacterium tuberculosis in mice. *Molecular Microbiology* 36, 630–637 (2000).
- 114. Glickman, M. S., Cox, J. S. & Jacobs, W. R. A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis. Mol. Cell* 5, 717–727 (2000).
- 115. George, K. M., Yuan, Y., Sherman, D. R. & Barry, C. E. The biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. Identification and functional analysis of CMAS-2. *J. Biol. Chem.* 270, 27292–27298 (1995).
- 116. Yuan, Y., Lee, R. E., Besra, G. S., Belisle, J. T. & Barry, C. E. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. *PNAS* 92, 6630–6634 (1995).
- 117. Zhang, Z., Bulloch, E. M. M., Bunker, R. D., Baker, E. N. & Squire, C. J. Structure and function of GImU from *Mycobacterium tuberculosis*. *Acta Cryst (2009)*. *D65*, 275-283.
- Jagtap, P. K. A. *et al.* Substrate-bound Crystal Structures Reveal Features Unique to *Mycobacterium tuberculosis* N-Acetyl-glucosamine 1-Phosphate Uridyltransferase and a Catalytic Mechanism for Acetyl Transfer. *J. Biol. Chem.* 287, 39524–39537 (2012).
- 119. Li, S. *et al.* Identification of *M. tuberculosis* Rv3441c and *M. smegmatis* MSMEG_1556 and Essentiality of M. smegmatis MSMEG_1556. *PLoS ONE* 7, e42769–9 (2012).

- 120. De Smet, K. A., Kempsell, K. E., Gallagher, A., Duncan, K. & Young, D. B. Alteration of a single amino acid residue reverses fosfomycin resistance of recombinant MurA from *Mycobacterium tuberculosis*. *Microbiology* 145 (Pt 11), 3177–3184 (1999).
- 121. Xu, L. *et al.* Characterization of mycobacterial UDP-N-acetylglucosamine enolpyruvyle transferase (MurA). *Research in Microbiologoy* 165, 91–101 (2014).
- 122. Benson, T. E., Marquardt, J. L., Marquardt, A. C., Etzkorn, F. A. & Walsh, C. T. Overexpression, purification, and mechanistic study of UDP-N-acetylenolpyruvylglucosamine reductase. *Biochemistry* 32, 2024–2030 (1993).
- 123. Basavannacharya, C., Robertson, G., Munshi, T., Keep, N. H. & Bhakta, S. ATPdependent MurE ligase in *Mycobacterium tuberculosis*: Biochemical and structural characterisation. *Tuberculosis* 90, 16–24 (2010).
- 124. Basavannacharya, C. *et al.* Essential residues for the enzyme activity of ATPdependent MurE ligase from *Mycobacterium tuberculosis*. *Protein Cell* 1, 1011– 1022 (2010).
- 125. Munshi, T. *et al.* Characterisation of ATP-dependent Mur ligases involved in the biogenesis of cell wall peptidoglycan in *Mycobacterium tuberculosis. PLoS ONE* 8, e60143 (2013).
- 126. Eniyan, K., Kumar, A., Rayasam, G. V., Perdih, A. & Bajpai, U. Development of a one-pot assay for screening and identification of Mur pathway inhibitors in *Mycobacterium tuberculosis. Sci. Rep.* 1–12 (2016).
- Maitra, A. *et al.* Cell wall peptidoglycan in Mycobacterium tuberculosis: An Achilles' heel for the TB-causing pathogen. *FEMS Microbiol Rev* 293, 9770–28 (2019).
- 128. Münch, D. *et al.* Identification and in vitro Analysis of the GatD/MurT Enzyme-Complex Catalyzing Lipid II Amidation in *Staphylococcus aureus*. *PLoS Pathog* 8, e1002509–11 (2012).
- 129. Ruiz, N. Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *PNAS* 105, 15553–15557 (2008).
- 130. Mohammadi, T. *et al.* Identification of FtsW as a transporter of lipid- linked cell wall precursors across the membrane. *The EMBO Journal* 30, 1425–1432 (2011).

- 131. Sham, L. T. *et al.* MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. *Science* 345, 220–222 (2018).
- Hett, E. C., Chao, M. C. & Rubin, E. J. Interaction and Modulation of Two Antagonistic Cell Wall Enzymes of Mycobacteria. *PLoS Pathog* 6, e1001020–15 (2010).
- 133. Schoonmaker, M. K., Bishai, W. R. & Lamichhane, G. Nonclassical Transpeptidases of *Mycobacterium tuberculosis* Alter Cell Size, Morphology, the Cytosolic Matrix, Protein Localization, Virulence, and Resistance to beta-Lactams. *Journal of Bacteriology* 196, 1394–1402 (2014).
- 134. Sacco, E. *et al.* Activation of the L,D-transpeptidation peptidoglycan cross-linking pathway by a metallo-D,D-carboxypeptidase in *Enterococcus faecium*. *Molecular Microbiology* 75, 874–885 (2010).
- 135. Mikusova, K., Mikus, M., Besra, G. S., Hancock, I. & Brennan, P. J. Biosynthesis of the linkage region of the mycobacterial cell wall. *J. Biol. Chem.* 271, 7820–7828 (1996).
- 136. Jin, Y., Xin, Y., Zhang, W. & Ma, Y. *Mycobacterium tuberculosis* Rv1302 and *Mycobacterium smegmatis* MSMEG__4947 have WecA function and MSMEG__4947 is required for the growth of *M. smegmatis*. *FEMS Microbiol. Lett*. 310, 54–61 (2010).
- 137. Ishizaki, Y. *et al.* Inhibition of the first step in synthesis of the mycobacterial cell wall core, catalyzed by the GlcNAc-1-phosphate transferase WecA, by the novel caprazamycin derivative CPZEN-45. *J. Biol. Chem.* 288, 30309–30319 (2013).
- 138. Mills, J. A. *et al.* Inactivation of the mycobacterial rhamnosyltransferase, which is needed for the formation of the arabinogalactan-peptidoglycan linker, leads to irreversible loss of viability. *J. Biol. Chem.* 279, 43540–43546 (2004).
- 139. Ma, Y. *et al.* Determination of the pathway for rhamnose biosynthesis in mycobacteria: cloning, sequencing and expression of the *Mycobacterium tuberculosis* gene encoding alpha-D-glucose-1-phosphate thymidylyltransferase. *Microbiology* 143 (Pt 3), 937–945 (1997).
- 140. Hoang, T. T., Ma, Y., Stern, R. J., McNeil, M. R. & Schweizer, H. P. Construction and use of low-copy number T7 expression vectors for purification of problem

proteins: purification of *Mycobacterium tuberculosis* RmID and *Pseudomonas aeruginosa* LasI and RhII proteins, and functional analysis of purified RhII. *Gene* 237, 361–371 (1999).

- 141. Stern, R. J. *et al.* Conversion of dTDP-4-keto-6-deoxyglucose to free dTDP-4-ketorhamnose by the rmIC gene products of *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology* 145 (Pt 3), 663–671 (1999).
- 142. Ma, Y. *et al.* Drug Targeting *Mycobacterium tuberculosis* Cell Wall Synthesis: Genetics of dTDP-Rhamnose Synthetic Enzymes and Development of a Microtiter Plate-Based Screen for Inhibitors of Conversion of dTDP-glucose to dTDP-Rhamnose. *Antimicrobial Agents and Chemotherapy* 45, 1407–1416 (2001).
- 143. Li, W., Xin, Y., McNeil, M. R. & Ma, Y. rmlB and rmlC genes are essential for growth of mycobacteria. *Biochemical and Biophysical Research Communications* 342, 170–178 (2006).
- 144. Lemaire, H. G., Müller-Hill, B.1986. Nucleotide sequences of the *gal* E gene and the *gal* T gene of *E.coli*. *Nucleic Acids Research* 14, 7705–7711 (1986).
- 145. Weston, A. *et al.* Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. *Tuber. Lung Dis.* 78, 123–131 (1997).
- 146. Mikusova, K. *et al.* Identification of a Novel Galactosyl Transferase Involved in Biosynthesis of the Mycobacterial Cell Wall. *Journal of Bacteriology* 188, 6592– 6598 (2006).
- 147. Alderwick, L. J. *et al.* Expression, purification and characterisation of soluble GlfT and the identification of a novel galactofuranosyltransferase Rv3782 involved in priming GlfT-mediated galactan polymerisation in *Mycobacterium tuberculosis*. *Protein Expression and Purification* 58, 332–341 (2008).
- 148. Kremer, L. *et al.* Galactan Biosynthesis in Mycobacterium tuberculosis. *J. Biol. Chem.* 276, 26430–26440 (2001).
- 149. May, J. F., Splain, R. A., Brotschi, C. & Kiessling, L. L. A tethering mechanism for length control in a processive carbohydrate polymerization. *PNAS* 106, 11851–11856 (2009).
- 150. Alderwick, L. J. *et al.* Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an

arabinan-deficient mutant with a cell wall galactan core. *J. Biol. Chem.* 280, 32362–32371 (2005).

- 151. Jiang, T. *et al.* The effect of MSMEG_6402 gene disruption on the cell wall structure of *Mycobacterium smegmatis*. *Microbial Pathogenesis* 51, 156–160 (2011).
- 152. Mikusova, K. *et al.* Decaprenylphosphoryl Arabinofuranose, the Donor of the D-Arabinofuranosyl Residues of Mycobacterial Arabinan, Is Formed via a Two-Step Epimerization of Decaprenylphosphoryl Ribose. *Journal of Bacteriology* 187, 8020–8025 (2005).
- 153. Meniche, X. *et al.* Partial redundancy in the synthesis of the D-arabinose incorporated in the cell wall arabinan of *Corynebacterineae*. *Microbiology* 154, 2315–2326 (2008).
- 154. Crellin, P. K., Brammananth, R. & Coppel, R. L. Decaprenylphosphoryl-β-D-Ribose
 2['] -Epimerase, the Target of Benzothiazinones and Dinitrobenzamides, Is an Essential Enzyme in *Mycobacterium smegmatis*. *PLoS ONE* 6, e16869–8 (2011).
- 155. Makarov, V. *et al.* Benzothiazinones Kill Mycobacterium tuberculosis by Blocking Arabinan Synthesis. *Science* 324, 801–804 (2009).
- 156. Alderwick, L. J., Seidel, M., Sahm, H., Besra, G. S. & Eggeling, L. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis. J. Biol. Chem.* 281, 15653–15661 (2006).
- 157. Shi, L. *et al.* Transfer of the First Arabinofuranose Residue to Galactan Is Essential for *Mycobacterium smegmatis* Viability. *Journal of Bacteriology* 190, 5248–5255 (2008).
- 158. Escuyer, V. E. *et al.* The Role of the embAand embBGene Products in the Biosynthesis of the Terminal Hexaarabinofuranosyl Motif of Mycobacterium smegmatisArabinogalactan. *J. Biol. Chem.* 276, 48854–48862 (2001).
- 159. Jankute, M., Grover, S., Rana, A. K. & Besra, G. S. Arabinogalactan and lipoarabinomannan biosynthesis: structure, biogenesis and their potential as drug targets. *Future Microbiology* 7, 129–147 (2012).

- 160. Belanger, A. E. *et al.* The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. *PNAS* 93, 11919–11924 (1996).
- 161. Birch, H. L. *et al.* A truncated lipoglycan from mycobacteria with altered immunological properties. *PNAS* 107, 2634–2639 (2010).
- Birch, H. L. *et al.* Biosynthesis of mycobacterial arabinogalactan: identification of a novel (13) arabinofuranosyltransferase. *Molecular Microbiology* 280, 32362–16 (2008).
- 163. Goude, R., Amin, A. G., Chatterjee, D. & Parish, T. The Arabinosyltransferase EmbC Is Inhibited by Ethambutol in Mycobacterium tuberculosis. *Antimicrobial Agents and Chemotherapy* 53, 4138–4146 (2009).
- 164. Škovierová, H. *et al.* AftD, a novel essential arabinofuranosyltransferase from mycobacteria. *Glycobiology* 19, 1235–1247 (2009).
- Belisle, J. T. Role of the Major Antigen of Mycobacterium tuberculosis in Cell Wall
 Biogenesis. *Science* 276, 1420–1422 (1997).
- 166. Seidel, M. *et al.* 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.* 282, 14729–14740 (2007).
- 167. Lee, A. *et al.* Sequencing of Oligoarabinosyl Units Released from Mycobacterial Arabinogalactan by Endogenous Arabinanase: Identification of Distinctive and Novel Structural Motifs †. *Biochemistry* 45, 15817–15828 (2006).
- 168. Draper, P., Khoo, K. H., Chatterjee, D., Dell, A. & Morris, H. R. Galactosamine in walls of slow-growing mycobacteria. *Biochem. J.* 327 (Pt 2), 519–525 (1997).
- 169. Škovierová, H. *et al.* Biosynthetic Origin of the Galactosamine Substituent of Arabinogalactan in Mycobacterium tuberculosis. *J. Biol. Chem.* 285, 41348–41355 (2010).
- 170. Jackson, M. *et al.* Inactivation of the antigen 85C gene profoundly affects the mycolate content and alters the permeability of the *Mycobacterium tuberculosis* cell envelope. *Molecular Microbiology* 31, 1573–1587 (1999).

- 171. Tong, L. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *CMLS, Cell. Mol.Life Sci* 62, 1784–1803 (2005).
- 172. Gago, G., Diacovich, L., Arabolaza, A., Tsai, S.-C. & Gramajo, H. Fatty acid biosynthesis in actinomycetes. *FEMS Microbiol Rev* 35, 475–497 (2011).
- 173. Kremer, L. *et al.* Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of Mycobacterium tuberculosis fatty acid synthase II. *J. Biol. Chem.* 276, 27967– 27974 (2001).
- Glickman, M. S. The *mmaA2* Gene of *Mycobacterium tuberculosis* Encodes the Distal Cyclopropane Synthase of the α-Mycolic Acid. *J. Biol. Chem.* 278, 7844–7849 (2003).
- 175. Galandrin, S. *et al.* Assay Development for Identifying Inhibitors of the Mycobacterial FadD32 Activity. *J Biomol Screen* 18, 576–587 (2013).
- 176. Thanna, S. & Sucheck, S. J. Targeting the trehalose utilization pathways of *Mycobacterium tuberculosis. Med. Chem. Commun.* 7, 69–85 (2016).
- 177. Hiratake, J. & Sakata, K. Glycosylamidines as potent selective and easily accessible glycosidase inhibitors and their application to affinity chromatography. *Meth. Enzymol.* 363, 421–444 (2003).
- 178. Davies, G. J., Gloster, T. M. & Henrissat, B. Recent structural insights into the expanding world of carbohydrate-active enzymes. *Current Opinion in Structural Biology* 15, 637–645 (2005).
- 179. Wolfenden, R., Lu, X. & Young, G. Spontaneous hydrolysis of glycosides. *J. Am. Chem. Soc.* 6814–6815 (1998).
- Wolfenden, R. Benchmark Reaction Rates, the Stability of Biological Molecules in Water, and the Evolution of Catalytic Power in Enzymes. *Annu. Rev. Biochem.* 80, 645–667 (2011).
- 181. Lairson, L. L. & Withers, S. G. Mechanistic analogies amongst carbohydrate modifying enzymes. *Chem. Commun.* 2243–6 (2004).
- 182. Laine, R. A. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10(12) structures for a reducing hexasaccharide: the

Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4, 759–767 (1994).

- 183. Kötzler, M. P., Hancock, S. M. & Withers, S. Glycosidases: Functions, families and folds. *eLS* (2014).
- 184. Vocadlo, D. J. & Davies, G. J. Mechanistic insights into glycosidase chemistry. *Curr Opin Chem Biol* 12, 539–555 (2008).
- 185. Ly, H. D. & Withers, S. G. Mutagenesis of glycosidases. Annu. Rev. Biochem. 68, 487–522 (1999).
- 186. Zechel, D. L. & Withers, S. Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Accounts of Chemical Research* 33, 11–18 (2000).
- 187. Zhu, Y. *et al.* Mechanistic insights into a Ca2+-dependent family of αmannosidases in a human gut symbiont. *Nat Chem Biol* 6, 125–132 (2009).
- Juers, D. H. *et al.* A Structural View of the Action of Escherichia coli(lacZ) β Galactosidase †,‡. *Biochemistry* 40, 14781–14794 (2001).
- 189. Richard, J. P., Huber, R. E., Lin, S., Heo, C. & Amyes, T. L. Structure-reactivity relationships for beta-galactosidase (Escherichia coli, lac Z). 3. Evidence that Glu-461 participates in Brønsted acid-base catalysis of beta-D-galactopyranosyl group transfer. *Biochemistry* 35, 12377–12386 (1996).
- 190. Okuyama, M. *et al.* Catalytic role of the calcium ion in GH97 inverting glycoside hydrolase. *FEBS Letters* 588, 3213–3217 (2014).
- 191. Gloster, T. M., Turkenburg, J. P., Potts, J. R., Henrissat, B. & Davies, G. J. Divergence of Catalytic Mechanism within a Glycosidase Family Provides Insight into Evolution of Carbohydrate Metabolism by Human Gut Flora. *Chemistry & Biology* 15, 1058–1067 (2008).
- 192. Livingstone, G., Franks, F. & Aspinall, L. The effects of aqueous solvent structure on the mutarotation kinetics of glucose. *Journal of Solution Chemistry* (1977).
- 193. Pigman, W. & Isbell, H. S. in 23, 11–57 (Elsevier, 1968).
- 194. Withers, S. G. & Aebersold, R. Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci.* 4, 361–372 (1995).
- 195. Dias, F. M. V. *et al.* Insights into the Molecular Determinants of Substrate Specificity in Glycoside Hydrolase Family 5 Revealed by the Crystal Structure and

Kinetics of Cellvibrio mixtus Mannosidase 5A. J. Biol. Chem. 279, 25517–25526 (2004).

- 196. Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J. & Withers, S. G. Changing enzymic reaction mechanisms by mutagenesis: conversion of a retaining glucosidase to an inverting enzyme. *J. Am. Chem. Soc.* 116, 11594–11595 (1994).
- 197. Vocadlo, D. J. & Withers, S. G. Identification of active site residues in glycosidases by use of tandem mass spectrometry. *Methods Mol. Biol.* 146, 203–222 (2000).
- 198. Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280 (Pt 2), 309–316 (1991).
- 199. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* 42, D490–D495 (2013).
- 200. Helbert, W. *et al.* Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. *PNAS* 116, 6063–6068 (2019).
- 201. Gerlt, J. Evolution of function in (β/α) 8-barrel enzymes. *Curr Opin Chem Biol* 7, 252–264 (2003).
- 202. Wierenga, R. The TIM barrel fold: a versatile framework for efficient enzymes. *FEBS Letters* (2011).
- 203. Pons, T., Gómez, R., Chinea, G. & Valencia, A. Beta-propellers: associated functions and their role in human diseases. *Curr. Med. Chem.* 10, 505–524 (2003).
- 204. Kreisberg, J. F., Betts, S. D. & King, J. β-Helix core packing within the triplestranded oligomerization domain of the P22 tailspike. *Protein Sci.* 9, 2338–2343 (2000).
- 205. Davies, G. J. & Williams, S. J. Carbohydrate-active enzymes: sequences, shapes, contortions and cells. *Biochemical Society Transactions* 44, 79–87 (2016).
- 206. van Wyk, N., Drancourt, M., Henrissat, B. & Kremer, L. Current perspectives on the families of glycoside hydrolases of Mycobacterium tuberculosis: their importance and prospects for assigning function to unknowns. *Glycobiology* 27, 112–122 (2017).
- 207. Miah, F. *et al.* Flux through Trehalose Synthase Flows from Trehalose to the Alpha Anomer of Maltose in Mycobacteria. *Chemistry & Biology* 20, 487–493 (2013).

- 208. Kalscheuer, R. *et al.* Self-poisoning of Mycobacterium tuberculosis by targeting GlgE in an α -glucan pathway. *Nat Chem Biol* 6, 376–384 (2010).
- 209. Garg, S. K., Alam, M. S., Kishan, K. V. R. & Agrawal, P. Expression and characterization of α -(1,4)-glucan branching enzyme Rv1326c of Mycobacterium tuberculosis H37Rv. *Protein Expression and Purification* 51, 198–208 (2007).
- 210. Carroll, J. D., Pastuszak, I., Edavana, V. K., Pan, Y. T. & Elbein, A. D. A novel trehalase from Mycobacterium smegmatis purification, properties, requirements. *FEBS Journal* 274, 1701–1714 (2007).
- 211. Varrot, A. *et al.* Mycobacterium tuberculosisStrains Possess Functional Cellulases.
 J. Biol. Chem. 280, 20181–20184 (2005).
- 212. Koeck, D. E., Pechtl, A., Zverlov, V. V. & Schwarz, W. H. Genomics of cellulolytic bacteria. *Current Opinion in Biotechnology* 29, 171–183 (2014).
- 213. Moynihan, P. J. *et al.* The hydrolase LpqI primes mycobacterial peptidoglycan recycling. *Nature Communications* 10, 2647 (2019).
- 214. Hett, E. C. *et al.* A partner for the resuscitation-promoting factors of *Mycobacterium tuberculosis*. *Molecular Microbiology* 66, 658–668 (2007).
- 215. Hett, E. C., Chao, M. C., Deng, L. L. & Rubin, E. J. A Mycobacterial Enzyme Essential for Cell Division Synergizes with Resuscitation-Promoting Factor. *PLoS Pathog* 4, e1000001–9 (2008).
- 216. Nikitushkin, V. D. *et al.* A product of RpfB and RipA joint enzymatic action promotes the resuscitation of dormant mycobacteria. *FEBS Journal* 282, 2500–2511 (2015).
- 217. Berg, S., Kaur, D., Jackson, M. & Brennan, P. J. The glycosyltransferases of *Mycobacterium tuberculosis* roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology* 17, 35–56R (2007).
- Rivera-Marrero, C. A., Ritzenthaler, J. D., Roman, J. & Moremen, K. W. Molecular cloning and expression of an α-mannosidase gene in Mycobacterium tuberculosis.
 Microbial Pathogenesis 30, 9–18 (2001).
- 219. Wadood, A. *et al.* Selective glycosidase inhibitors: A patent review (2012-present).
 Int. J. Biol. Macromol. 111, 82–91 (2018).

- 220. Kieser, K. J. & Rubin, E. J. How sisters grow apart: mycobacterial growth and division. *Nat Rev Micro* 12, 550–562 (2014).
- 221. Cordillot, M. *et al.* In VitroCross-Linking of *Mycobacterium tuberculosis* Peptidoglycan by L,D-Transpeptidases and Inactivation of These Enzymes by Carbapenems. *Antimicrobial Agents and Chemotherapy* 57, 5940–5945 (2013).
- 222. Saxena, A., Srivastava, V., Srivastava, R. & Srivastava, B. S. Identification of genes of Mycobacterium tuberculosis upregulated during anaerobic persistence by fluorescence and kanamycin resistance selection. *Tuberculosis* 88, 518–525 (2008).
- 223. Rachman, H. *et al.* Unique Transcriptome Signature of Mycobacterium tuberculosis in Pulmonary Tuberculosis. *Infect. Immun.* 74, 1233–1242 (2006).
- 224. Vandal, O. H. *et al.* Acid-Susceptible Mutants of Mycobacterium tuberculosis Share Hypersusceptibility to Cell Wall and Oxidative Stress and to the Host Environment. *Journal of Bacteriology* 191, 625–631 (2009).
- 225. Patru, M. M. & Pavelka, M. S. A Role for the Class A Penicillin-Binding Protein PonA2 in the Survival of *Mycobacterium smegmatis* under Conditions of Nonreplication. *Journal of Bacteriology* 192, 3043–3054 (2010).
- Gray, K. M., Keer, J., Williams, H. D. & Smeulders, M. J. Mutants of Mycobacterium smegmatis impaired in stationary-phase survival. *Microbiology* 146, 2209–2217 (2000).
- 227. Kumar, P. *et al.* Meropenem inhibits D,D-carboxypeptidase activity in Mycobacterium tuberculosis. *Molecular Microbiology* 86, 367–381 (2012).
- 228. Bourai, N., Jacobs, W. R., Jr. & Narayanan, S. Deletion and overexpression studies on DacB2, a putative low molecular mass penicillin binding protein from *Mycobacterium tuberculosis* H₃₇Rv. *Microbial Pathogenesis* 52, 109–116 (2012).
- 229. Baranowski, C. *et al.* Maturing Mycobacterium smegmatis peptidoglycan requires non-canonical crosslinks to maintain shape. *Elife* 7, 100 (2018).
- 230. Votyakova, T. V., Kaprelyants, A. S. & Kell, D. B. Influence of Viable Cells on the Resuscitation of Dormant Cells in Micrococcus luteus Cultures Held in an Extended Stationary Phase: the Population Effect. *Applied and Environmental Microbiology* 60, 3284–3291 (1994).

- 231. Mukamolova, G. V., Yanopolskaya, N. D., Kell, D. B. & Kaprelyants, A. S. On resuscitation from the dormant state of Micrococcus luteus. *Antonie Van Leeuwenhoek* 73, 237–243 (1998).
- 232. Mukamolova, G. V., Kormer, S. S., Kell, D. B. & Kaprelyants, A. S. Stimulation of the multiplication of *Micrococcus luteus* by an autocrine growth factor. *Arch. Microbiol.* 172, 9–14 (1999).
- 233. Mukamolova, G. V. *et al.* The *rpf* gene of *Micrococcus luteus* encodes an essential secreted growth factor. *Molecular Microbiology* 46, 611–621 (2002).
- 234. Cole, S. T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544 (1998).
- 235. Mukamolova, G. V., Kaprelyants, A. S., Young, D. I., Young, M. & Kell, D. B. A bacterial cytokine. *PNAS* 95, 8916–8921 (1998).
- Sutcliffe, I. C. & Harrington, D. J. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. *FEMS Microbiol Rev* 28, 645–659 (2004).
- Böth, D., Schneider, G. & Schnell, R. Peptidoglycan Remodeling in *Mycobacterium tuberculosis*: Comparison of Structures and Catalytic Activities of RipA and RipB.
 Journal of Molecular Biology 413, 247–260 (2011).
- 238. Kumar, A. *et al.* The structure of Rv3717 reveals a novel amidase from *Mycobacterium tuberculosis. Acta Crystallogr. D Biol. Crystallogr.* 69, 2543–2554 (2013).
- Prigozhin, D. M., Mavrici, D., Huizar, J. P., Vansell, H. J. & Alber, T. Structural and biochemical analyses of *Mycobacterium tuberculosis N*-acetylmuramyl-L-alanine amidase Rv3717 point to a role in peptidoglycan fragment recycling. *J. Biol. Chem.* 288, 31549–31555 (2013).
- 240. Goodell, E. W. Recycling of murein by *Escherichia coli*. *Journal of Bacteriology* 163, 305–310 (1985).
- Ahangar, M. S. *et al.* Structural and functional determination of homologs of the *Mycobacterium tuberculosis N*-acetylglucosamine-6-phosphate deacetylase (NagA). *J. Biol. Chem.* 293, 9770–9783 (2018).

- 242. Seiler, P. *et al.* Cell-wall alterations as an attribute of Mycobacterium tuberculosis in latent infection. *J. Infect. Dis.* 188, 1326–1331 (2003).
- 243. Bhamidi, S. *et al.* A bioanalytical method to determine the cell wall composition of Mycobacterium tuberculosis grown in vivo. *Analytical Biochemistry* 421, 240– 249 (2012).
- 244. Barkan, D., Hedhli, D., Yan, H.-G., Huygen, K. & Glickman, M. S. Mycobacterium tuberculosis Lacking All Mycolic Acid Cyclopropanation Is Viable but Highly Attenuated and Hyperinflammatory in Mice. *Infect. Immun.* 80, 1958–1968 (2012).
- 245. Jain, M. *et al.* Lipidomics reveals control of Mycobacterium tuberculosis virulence lipids via metabolic coupling. *PNAS* 104, 5133–5138 (2007).
- Vander Beken, S. *et al.* Molecular structure of the Mycobacterium tuberculosis virulence factor, mycolic acid, determines the elicited inflammatory pattern. *Eur. J. Immunol.* 41, 450–460 (2010).
- 247. Bhatt, A. *et al.* Deletion of kasB in Mycobacterium tuberculosis causes loss of acidfastness and subclinical latent tuberculosis in immunocompetent mice. *PNAS* 104, 5157–5162 (2007).
- Xin, Y., Huang, Y. & McNeil, M. The presence of an endogenous endo-D-arabinase in *Mycobacterium smegmatis* and characterization of its oligoarabinoside product.
 Biochimica et Biophysica Acta (BBA) General Subjects (1999).
- 249. Rombouts, Y. *et al.* Exposure of Mycobacteria to Cell Wall-inhibitory Drugs Decreases Production of Arabinoglycerolipid Related to Mycolyl-arabinogalactanpeptidoglycan Metabolism. *J. Biol. Chem.* 287, 11060–11069 (2012).
- 250. Wheat, W. H. *et al.* The presence of a galactosamine substituent on the arabinogalactan of Mycobacterium tuberculosis abrogates full maturation of human peripheral blood monocyte-derived dendritic cells and increases secretion of IL-10. *Tuberculosis* 95, 476–489 (2015).
- 251. Saha, B. C. Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. *Biotechnology Advances* 18, 403–423 (2000).
- 252. Margolles-Clark, E. & Environ, M. T. A. Cloning of genes encoding alpha-Larabinofuranosidase and beta-xylosidase from Trichoderma reesei by expression in Saccharomyces cerevisiae. *Am Soc Microbiol* (1996).

- 253. Ferrer, M. *et al.* Functional Metagenomics Unveils a Multifunctional Glycosyl Hydrolase from the Family 43 Catalysing the Breakdown of Plant Polymers in the Calf Rumen. *PLoS ONE* 7, e38134–11 (2012).
- 254. Naumoff, D. G. Conserved sequence motifs in levansucrases and bifunctional beta-xylosidases and alpha-L-arabinases. *FEBS Letters* 448, 177–179 (1999).
- 255. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423 (2019).
- 256. Burgess, R. R. & Deutscher, M. P. *Guide to Protein Purification*. (Academic Press, 2009).
- 257. Structural Genomics Consortium *et al.* Protein production and purification. *Nat Meth* 5, 135–146 (2008).
- 258. Calloni, G. *et al.* DnaK Functions as a Central Hub in the E. coli Chaperone Network. *Cell Reports* 1, 251–264 (2012).
- 259. Williams, G. J., Breazeale, S. D., Raetz, C. R. H. & Naismith, J. H. Structure and Function of Both Domains of ArnA, a Dual Function Decarboxylase and a Formyltransferase, Involved in 4-Amino-4-deoxy-L-arabinose Biosynthesis. *J. Biol. Chem.* 280, 23000–23008 (2005).
- 260. Breazeale, S. D., Ribeiro, A. A., McClerren, A. L. & Raetz, C. R. H. A Formyltransferase Required for Polymyxin Resistance in Escherichia coliand the Modification of Lipid A with 4-Amino-4-deoxy-l-arabinose. *J. Biol. Chem.* 280, 14154–14167 (2005).
- 261. Marcotte, E. M. Detecting Protein Function and Protein-Protein Interactions from Genome Sequences. *Science* 285, 751–753 (1999).
- Wallis, G., Hemming, F. W. & Peberdy, J. An extracellular β-galactofuranosidase from Aspergillus niger and its use as a tool for glycoconjugate analysis. *Biochimica et Biophysica Acta (BBA) General Subjects* 1525, 19–28 (2001).
- 263. de Sanctis, D., Inácio, J. M., Lindley, P. F., de Sá-Nogueira, I. & Bento, I. New evidence for the role of calcium in the glycosidase reaction of GH43 arabinanases.
 FEBS Journal 277, 4562–4574 (2010).

- 264. Matsunaga, E. *et al.* Identification and Characterization of a Novel Galactofuranose-Specific β-D-Galactofuranosidase from Streptomyces Species.
 PLoS ONE 10, e0137230 (2015).
- 265. Lee, S. Y. *et al.* Inhibition of cytochrome P450 by ethambutol in human liver microsomes. *Toxicology Letters* 229, 33–40 (2014).
- 266. Marangoni, A. G. *Enzyme Kinetics*. (John Wiley & Sons, 2003).
- 267. Angyal, S. The composition of reducing sugars in solution. *Adv Carbohydr Chem Biochem* 42, (1984).
- 268. Kamerling, J. P., Gerwig, G. J., Vliegenthart, J. F. & Clamp, J. R. Characterization by gas-liquid chromatography-mass spectrometry and proton-magnetic-resonance spectroscopy of pertrimethylsilyl methyl glycosides obtained in the methanolysis of glycoproteins and glycopeptides. *Biochem. J.* 151, 491–495 (1975).
- 269. Zhou, P. *et al.* Structural insights into the substrate specificity and transglycosylation activity of a fungal glycoside hydrolase family 5 [beta]-mannosidase. *Acta Cryst* D70, 2970–2982 (2014).
- Goncalves, A. M. D. *et al.* Endo-[beta]-d-1,4-mannanase from Chrysonilia sitophila displays a novel loop arrangement for substrate selectivity. *Acta Cryst* D68, 1468–1478 (2012).
- Hilge, M. *et al.* High-resolution native and complex structures of thermostable beta-mannanase from Thermomonospora fusca substrate specificity in glycosyl hydrolase family 5. *Structure* 6, 1433–1444 (1998).
- 272. Hong, X. & Hopfinger, A. J. Construction, Molecular Modeling, and Simulation of Mycobacterium tuberculosis Cell Walls. *Biomacromolecules* 5, 1052–1065 (2004).
- 273. Lee, R. E. B., Li, W., Chatterjee, D. & Lee, R. E. Rapid structural characterization of the arabinogalactan and lipoarabinomannan in live mycobacterial cells using 2D and 3D HR-MAS NMR: structural changes in the arabinan due to ethambutol treatment and gene mutation are observed. *Glycobiology* 15, 139–151 (2005).
- 274. Kim, Y. S. & Ma, J. Y. Insight into the Hydrolytic Selectivity of β-Glucosidase to Enhance the Contents of Desired Active Phytochemicals in Medicinal Plants.
 BioMed Research International 2018, 1–10 (2018).

- Tefsen, B. *et al.* Fungal α-arabinofuranosidases of glycosyl hydrolase families 51
 and 54 show a dual arabinofuranosyl- and galactofuranosyl-hydrolyzing activity.
 Biological Chemistry 393, 1–9 (2012).
- 276. Titgemeyer, F. *et al.* A Genomic View of Sugar Transport in Mycobacterium smegmatis and Mycobacterium tuberculosis. *Journal of Bacteriology* 189, 5903–5915 (2007).
- 277. Santucci, P. *et al.* Delineating the Physiological Roles of the PE and Catalytic Domains of LipY in Lipid Consumption in Mycobacterium-Infected Foamy Macrophages. *Infect. Immun.* 86, 591–45 (2018).
- 278. Santucci, P. *et al.* Dissecting the membrane lipid binding properties and lipase activity of *Mycobacterium tuberculosis* LipY domains. *FEBS J.* 78, 343–39 (2019).
- Mishra, K. C. *et al.* Functional Role of the PE Domain and Immunogenicity of the Mycobacterium tuberculosis Triacylglycerol Hydrolase LipY. *Infect. Immun.* 76, 127–140 (2007).
- Chung, C. T., Niemela, S. L. & Miller, R. H. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution.
 PNAS 86, 2172–2175 (1989).
- 281. Okonechnikov, K., Golosova, O. & Fursov, M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28, 1166–1167 (2012).
- 282. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425 (1987).
- 283. Guindon, S. & Gascuel, O. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* 52, 696–704 (2003).
- 284. Yang, J. & Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research* 43, W174–81 (2015).
- 285. Besra, G. S. Preparation of cell-wall fractions from mycobacteria. *Methods Mol. Biol.* 101, 91–107 (1998).
- 286. Ciucanu, I. & Kerek, F. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.* (1984).