Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Nicolas GERAUD

Le 23 novembre 2021

Développement d'un test phénotypique d'activité de la protéine O-mannosyltransférase de Mycobacterium tuberculosis pour la recherche d'inhibiteurs à fort potentiel thérapeutique

Ecole doctorale : **BSB - Biologie, Santé, Biotechnologies**

Spécialité : **BIOLOGIE STRUCTURALE ET FONCTIONNELLE**

Unité de recherche :
IPBS - Institut de Pharmacologie et Biologie Structurale

Thèse dirigée par
Michel RIVIèRE et Emeline FABRE

Jury

M. Marcelo GUERIN, Rapporteur
M. Jean-Louis HERRMANN, Rapporteur
Mme Marie-Ange KRZEWINSKI-RECCHI, Rapporteure
M. Laurent MAVEYRAUD, Examinateur
M. Michel RIVIÈRE, Directeur de thèse
Mme Emeline FABRE, Co-directrice de thèse

Résumé

L'antibiorésistance ou résistance acquise des bactéries aux antibiotiques est en passe de devenir l'une des principales causes de mortalité dans le monde. Le développement d'approches pharmacologiques antibactériennes novatrices constitue un enjeu vital face à cette « urgence sanitaire mondiale du XXIème siècle » (déclaration ONU du 21/09/2016).

Il apparait aujourd'hui que la recherche de composés ciblant les facteurs de virulence responsables des manifestations cliniques, mais non essentiels à la croissance de la bactérie pourrait constituer un des concepts des plus novateurs et prometteurs pour lutter contre l'antibiorésistance. De telles molécules dépourvues d'activité microbicide *per se*, exerçant une faible pression de sélection sur l'écosystème microbien, ne favoriseraient pas l'émergence de résistance et seraient dénuées des effets indésirables des antibiotiques conventionnels sur le microbiote commensal de l'hôte.

Dans ce contexte, lors de précédents travaux l'équipe a démontré que la mannosylation des protéines mycobactériennes n'est pas indispensable à la survie bactérienne, mais est essentielle à la virulence de *Mycobacterium tuberculosis*, agent étiologique de la tuberculose. Ceci nous a conduits à proposer l'enzyme Protéine O-mannosyltransférase PMTub, responsable de la mannosylation des protéines mycobactériennes, comme un « facteur de virulence non-essentiel » pouvant conduire au développement de nouvelles stratégies thérapeutiques contre la tuberculose.

En utilisant la modélisation moléculaire 3D, nous avons identifié un domaine fonctionnel spécifique de cette enzyme, appelé EL4, que nous prédisons être impliqués dans la reconnaissance des protéines substrats. Afin de valider cette hypothèse nous avons développé un test d'activité phénotypique original qui nous a permis d'en réaliser une analyse fonctionnelle par mutations ponctuelles. Les résultats obtenus démontrent pour la première fois l'implication de ce domaine spécifique dans l'activité des isoenzymes bactériennes.

De plus, le développement d'approches moléculaires complémentaires *in vitro* apportent des éléments préliminaires de compréhension du rôle de ce domaine dans l'activité de PMTub. Ces données nous ont permis d'initier la conception d'un test original de détection *in*-situ sur bactérie de l'interaction moléculaire entre ce domaine EL4 spécifique de PMTub et un substrat cible.

Ces travaux ont pour objectifs la mise au point d'un test phénotypique innovant de criblage de composés inhibiteurs dirigés contre la cible PMTub sur bactérie vivante. De tels composés spécifiquement actifs sur l'enzyme bactérienne et sans effets croisés sur les isoenzymes humaines pourraient ouvrir la voie au développement de nouvelles molécules à haut potentiel thérapeutique contre les infections mycobactériennes.

Abstract

The antimicrobial resistance has been declared as a “fundamental threat” to global health and safety at the first general assembly meeting on drug-resistant bacteria (UN, 21/09/2016). The development of innovative pharmacological anti-bacterial strategies constitutes a vital issue to face this worldwide health threat. In this regard, a novel promising concept is emerging which consist of targeting the so-called “non-essential bacterial virulence factors”. Such original drugs that would target *Mycobacterium tuberculosis* virulence factors during host colonization without directly impacting bacterial essential physiological processes are less likely to favor the development of drug resistance while still weakening the pathogen’s fitness and rendering it more sensitive to immune system attack. Moreover, anti-virulence drugs are not supposed to target the host microbiota, and therefore may show fewer side effects than classical antibiotics.

In this context, the team recently reported that the *Mycobacterium tuberculosis* Rv1002c protein O-mannosyltransferase (PMTub) is dispensable for *M. tuberculosis* growth but essential for its virulence in cellular and animal models. PMT activity impacts on a large number of bacterial glycosylated proteins that contribute to *M. tuberculosis* virulence and targeting PMT may thus result in a pleiotropic anti-virulence effect.

Using 3D molecular modeling, we have identified a specific functional domain of this enzyme, called EL4 that we predict to be involved in the recognition of protein substrates. In order to validate this hypothesis, we developed an original phenotypic activity assay taht allowed us to perform a functional analysis by point mutations. The results obtained demonstrate for the first time the involvement of this specific domain in the activity of bacterial enzymes.

In addition, the development of complementary *in vitro* molecular approaches provide preliminary elements for understanding the role of this domain in PMTub activity. These data allowed us to initiate the design of an original test for in-situ detection on bacteria of the molecular interaction between this specific EL4 domain of PMTub and a target substrate.

This project aims to develop an innovative assay for the detection of the molecular interaction between a reporter protein substrate of PMTub and the specific domain EL4. Such a strategy will then allow the selection of inhibitors of this recognition, which do not affect

the human iso-enzymes, and therefore suitable for the future development of new anti-tuberculous compounds harmless for the host, targeting the virulence factor PMTub.

Table des matières

Résumé 1
Abstract 3
Table des matières 5
Table des figures 8
Abréviations 11
Introduction bibliographique 13
Introduction : contexte et rationnel de l'étude 15
1. Tuberculose 15
2. Mycobacterium tuberculosis 24
Présentation générale des travaux 31
Rappels bibliographiques : Mannosylation des protéines et Glycosyltransférases 33
A. Protéine *O*-mannosylation et PMT 33
1. La O-mannosylation des protéines 33
2. Les Protéine O-mannosyltransférases 37
B. Propriétés structurales des glycosyltransférases 51
1. Classification des glycosyltransférases 51
2. Les glycosyltransférases de type A et B 51
3. Les glycosyltransférases de type C 53
Résultats et discussions 67
Partie 1 : Développement d'outils pour l'étude fonctionnelle de PMTub 69
A. Construction d'une souche rapportrice de l'activité de PMTub 69
B. Mise en évidence et quantification de la mannosylation de FasC 73
1. Par détection d'une altération du profil électrophorétique 73
2. Par spectrométrie de masse de protéine native 73
3. Par reconnaissance par une lectine 74
C. Contrôle de l'expression de PMTub pleine taille à la membrane 81
1. Détection de l'expression de PMTub par western blot 81
2. Détection de l'expression de PMTub sur bactéries 86
D. Conclusion 99
Partie 2 : Étude des relations structure-fonction de PMTub 101
A. Élaboration du modèle 3D de PMTub 101
1. Méthode 101
2. Modèle topologique de PMTub 105

3. Évaluation de la validité du modèle 105
4. Relation structure-fonction : Hypothèses de travail 107
5. Conclusion et perspectives de l'étude par modélisation 118
B. Caractérisation du site actif de PMTub 119
1. Mutagenèse 119
2. Chemical rescue 125
3. Conclusion et proposition d'un mécanisme catalytique 127
C. Identification de domaines fonctionnels 129
1. Étude structurale des GT-C et comparaison au modèle 3D de PMTub 129
2. Caractérisation du site d'interaction prédit avec le substrat protéique 135
3. EL4 comme "pseudo-WW domain" 143
Partie 3 : Stratégies d'identification d'inhibiteurs de PMTub 155
A. Criblage d'inhibiteurs en employant le test phénotypique d'activité 155
1. Preuve de faisabilité 155
2. Sélection rationnelle d'inhibiteurs putatifs 158
3. Perspectives : optimisation pour passage HTS 159
B. Développement d'une approche de biologie synthétique 167
1. Conception des outils génétiques 169
2. Construction de pSMjF 175
3. Insertion d'AAnn dans PMTub 179
1. Conclusion et perspectives 182
Conclusion générale 185
Matériels et méthodes 191
A. Cultures bactériennes et génération de souches mutées 193
A. Test d'activité phénotypique de PMTub 197
1. Étude de la mannosylation de la FasC 197
2. Contrôle de l'expression de PMTub à la membrane 201
3. Test d'inhibition de PMTub 209
B. Modélisation 211
C. Production et étude *in vitro* du peptide EL4 211
D. Approche de biologie synthétique 215
Annexe – Article en préparation 218
Bibliographie 241

Table des figures

Figure 1. Taux d'incidence estimé de la tuberculose en 2019. 14
Figure 2. Pathogénèse de la tuberculose. 16
Figure 3. Modèle de l'enveloppe mycobactérienne 26
Figure 4. Distribution des mannoprotéines identifiées à l'interface avec l'hôte 28
Figure 5. *O*-glycosylation des protéines dépendante de PMT 34
Figure 6 Structure des PMT de levure. 38
Figure 7. Les différentes familles de PMT 42
Figure 8. Les superfamilles de glycosyltransférases. 46
Figure 9. Les GT-Cs de structure résolues 50
Figure 10. Superposition des structures de GT-66 54
Figure 11 Comparaison des topologies des GT-C de structure résolues.. 56
Figure 12. Superposition des hélices horizontales 1 et 2 58
Figure 13. Le motif DxD (ou dérivé) des GT-Cs 60
Figure 14. Zone d'interaction avec le substrat lipidique donneur 62
Figure 15. Plusieurs GT-Cs décrivent une troisième hélice horizontale (HH3).. 64
Figure 16. Séquence annotée de la FasC recombinante utilisée dans cette étude. 68
Figure 17. Stratégies expérimentales pour suivre l'activité in vivo de la protéine O-mannosyltransférase PMTub 70
Figure 18. Validation du test d'activité phénotypique de PMTub 72
Figure 19. Validation du test d'activité phénotypique de PMTub 76
Figure 20. Suivi de la production de fascicline O-mannosylée en fonction des contenants utilisés pour la culture. 78
Figure 21. Contrôle de l'expression de PMTub par western blot 80
Figure 22. Visualisation de PMTub par microscopie à fluorescence 88
Figure 23. Détection de PMTub-mCherry par western blot. 90
Figure 24. Contrôle de l'expression de PMTub en utilisant la phosphatase alcaline. 94
Figure 25. Alignement de séquences de PMT 98
Figure 26. Modélisation 3D par homologie de PMTub 100
Figure 27. Topologie prédite de PMTub. 102
Figure 28. Comparaison du modèle de PMTub au socle commun des GT-Cs 104
Figure 29. Diagramme de Ramachandran des différents modèles de PMTub 106
Figure 30. Conservation de séquence et de localisation 3D de la poche catalytique entre PMTub et les PMT eucaryotes 109
Figure 31. Interaction avec le substrat donneur 110
Figure 32. Docking du substrat donneur dans le modèle structural de PMTub. 112
Figure 33. Interaction avec le substrat accepteur 114
Figure 34. Site actif prédit pour PMTub 116
Figure 35. Etude du site actif prédit de PMTub par mutagénèse dirigée 120
Figure 36. Résultats du « chemical rescue » de l'activité du mutant PMTub-R441A *in vivo*. 124
Figure 37. Proposition d'un mécanisme catalytique pour PMTub. 126
Figure 38. Visualisation de la boucle EL4 et de l'extrémité C-terminale des PMT 128
Figure 39. Distorsion de la bicouche lipidique autour de l'hétérocomplexe *Sc*PMT1-*Sc*PMT2 130
Figure 40. Positionnement du peptide accepteur dans *Sc*PMT2 132
Figure 41. Alignement de séquence des domaines fonctionnels prédits de plusieurs PMT. 134
Figure 42. Implication de EL4 et de l'extrémité C-terminale de PMTub 136

Figure 43. Activité PhoA relative mesurée pour les mutants non fonctionnels de PMTub .. 138
Figure 44. EL4 comme « pseudo-domaine WW » 142
Figure 45. Production de la boucle EL4 isolée dans *E. coli*..... 146
Figure 46. Spectre de dichroïsme circulaire du peptide $EL4_{346-394}$148
Figure 47. Suivi de la stabilité thermique du peptide $EL4_{346-394}$ par fluorescence..... 148
Figure 48. Analyse RMN ^{1}H du peptide $EL4_{346-394}$ 150
Figure 49. Titration calorimétrique de l'interaction entre le peptide EL4 de PMTub et un site peptide d'Apa de *M. tuberculosis,* porteur d'un site de mannosylation. 152
Figure 50. Mesure des effets inhibiteurs de dérivés de rhodanine-3-acide acétique sur PMTub. 156
Figure 51. Effet des composés sélectionnés rationnellement sur la croissance bactérienne161
Figure 52. Effet des composés sélectionnés rationnellement sur la concentration en FasC dans les surnageants de culture 163
Figure 53. Effet des composés sélectionnés rationnellement sur la mannosylation de la FasC. 165
Figure 54. Assemblage *in silico* de pSMjF (« long ») 170
Figure 55. Effet de la réduction de taille du promoteur pBlaF* sur l'activité de PMTub 172
Figure 56. Contrôle de la fonctionnalité du système orthogonal dans la souche de *M. smegmatis* exprimant les éléments génétiques du plasmide pSMJF 174
Figure 57. Contrôle de l'expression de la FasC 176
Figure 58. Contrôle de l'insertion d'un acide aminé non naturel dans PMTub..... 180
Figure 59. Séquence en acide aminé de PMTub présent sur le plasmide pWM158. 210
Figure 60. Séquence ADN insérée dans pGex-6P, correspondant au peptide EL4. 210
Figure 61. Séquence synthétique commandée pour la construction de pSMjF..... 214

Tableau 1. Liste des chimiothérapies antituberculeuses..... 20
Tableau 2. Les différentes familles de glycosyltransférase de type-C selon la base de donnée CAZy 48
Tableau 3. Les glycosyltransférase de type-C de structure 3D résolues..... 52
Tableau 4. Liste de peptides de mCherry-Nter-PMTub détectés par spectrométrie de masse 92
Tableau 5. Amorces sens utilisées pour la mutagénèse dirigée du vecteur codant pour PMTub.. 192
Tableau 6. Souches bactériennes utilisées dans cette étude et leurs principales caractéristiques. 194
Tableau 7. Antibiotiques utilisés pour les cultures bactériennes. 194
Tableau 8. Plasmides utilisés dans cette étude et leurs principales caractéristiques. 196
Tableau 9. Anticorps utilisés dans cette étude. 196
Tableau 10. Plasmides construits dans cette étude et leurs principales caractéristiques. ... 198
Tableau 11. Amorces utilisées pour la construction PMTub étiquetée. 202
Tableau 12. Amorces utilisées pour la construction de protéines de fusion..... 204
Tableau 13. Listes des composés inhibiteurs putatifs sélectionnés 208
Tableau 14. Amorces utilisées pour la construction du plasmide pSMjF et la mutagenèse du gène *gfp*..... 214
Tableau 15. Plasmides utilisés pour la construction du plasmide pSMjF 214
Tableau 16. Amorces utilisées pour l'introduction d'un codon ambre dans PMTub. 216

Abréviations

- OMS • Organisation Mondiale de la Santé
- PRR • Pattern recognition receptors
- PAMPS • Pathogen-Associated Molecular Patterns
- VIH • Virus de l'immunodéficience humaine
- BCG • Bacille de Calmette-Guérin
- PMT • Protéine *O*-mannosyltransférase
- PMTub • Protéine *O*-mannosyltransférase de *Mycobacterium tuberculosis*
- TMTC • Transmembrane and tetratricopeptide repeat-containing protein
- GT • Glycosyltransférase
- FasC • Fascicline
- ConA • Concanavaline A
- LM • Lipomannane
- PIM • phosphatidyl-inositol mannosides
- HRP • Horse Radish Peroxidase
- PhoA • Phosphatase alcaline
- BCIP • 5-Bromo-4-chloro-3-indolyl phosphate
- *p*Npp • Phosphate para-nitrophénylphosphate
- HH • Hélices horizontales
- EL • Boucle externe (« External Loop »)
- AAnn • Acide aminé non naturel

Introduction bibliographique

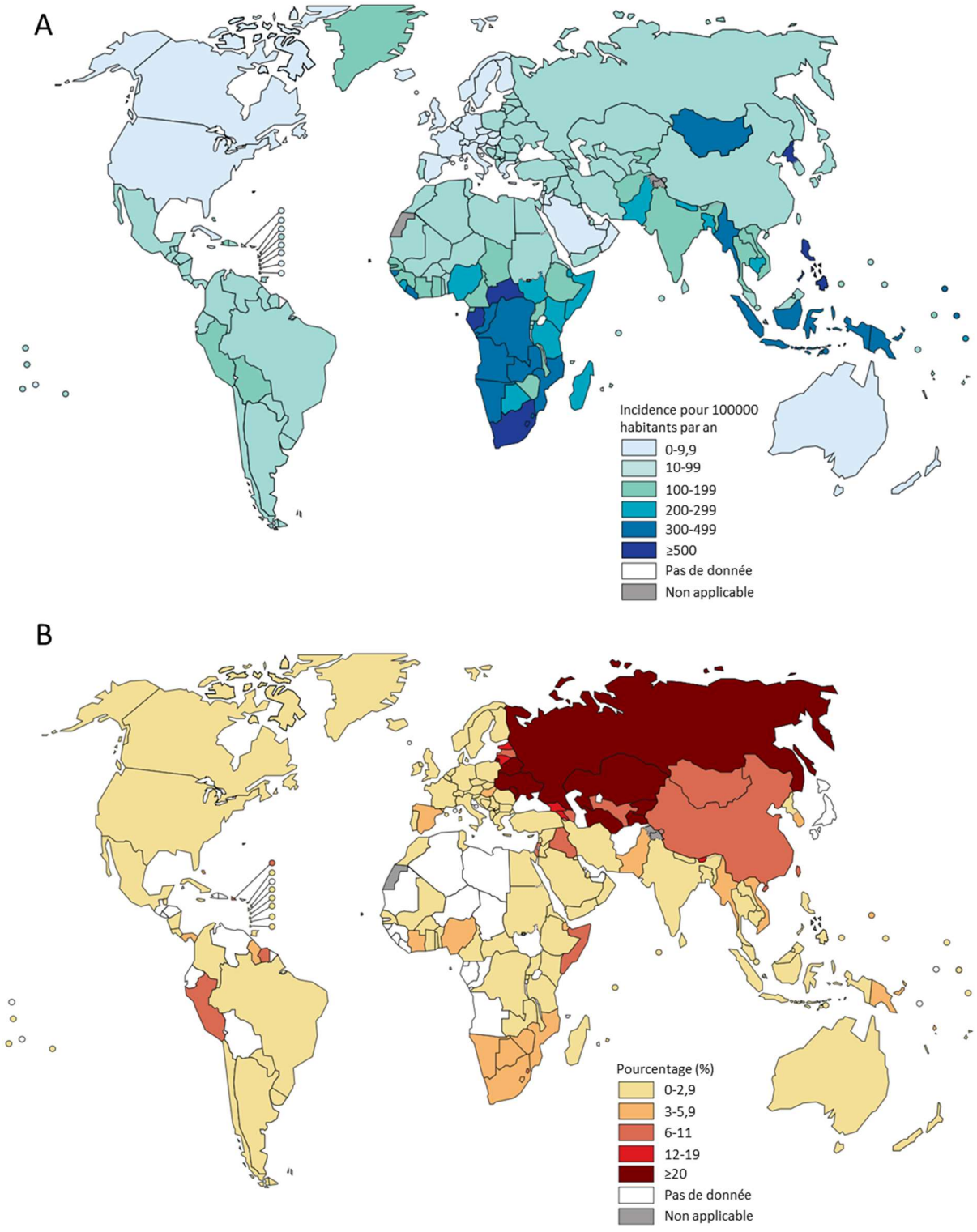


Figure 1. Taux d'incidence estimé de la tuberculose en 2019. (Adapté de [1]). **A)** Incidence de la tuberculose pour 100000 habitants en 2019 selon l'organisation mondiale de la santé. **B)** Pourcentage de nouveaux cas de tuberculose résistante (au moins à la rifampicine) en 2019. Les pourcentages sont basés sur les données disponible les plus récentes pour chaque pays.

Introduction : contexte et rationnel de l'étude

1. Tuberculose

La tuberculose pulmonaire humaine est une maladie infectieuse causée par la bactérie *Mycobacterium tuberculosis* également responsable de formes cliniques extrapulmonaires très minoritaires. Cette maladie qui a tué plus d'un milliard de personnes au cours des 200 dernières années[2] représente encore aujourd'hui un véritable enjeu de santé publique mondiale avec plus de 1,5 million de victimes (~3 morts par minutes) et 10 millions de nouveaux cas d'infection diagnostiqués en 2019 selon des estimations récentes de l'Organisation Mondiale de la Santé (OMS)[3].

a) Situation actuelle de la maladie

Bien que reléguée depuis peu au second rang par le SARS-CoV2, *M. tuberculosis* reste en effet toujours à l'échelle mondiale une des principales causes de décès par un agent infectieux unique. La tuberculose sévit principalement dans les pays pauvres où la prévalence de l'infection est très forte (**Figure 1A**) notamment du fait d'une comorbidité élevée avec le SIDA. Dans les pays développés, après une nette régression au début du XXème, on assiste depuis quelques années à une stagnation du recul de l'incidence de la maladie, essentiellement lié à une dégradation des conditions économiques et sociales associée à l'émergence croissante de souches de *M. tuberculosis* antibiorésistantes. Ce tableau est d'autant plus inquiétant qu'il pourrait être exacerbé par la pandémie de COVID dont l'impact sur le nombre de cas de tuberculose risque d'être tragique[1,4]. En effet, les mesures sanitaires mises en place comme le confinement, et la surcharge des hôpitaux pourraient retarder le dépistage et le traitement de la tuberculose, ce qui favoriserait la transmission du bacille à de nouveaux patients. L'OMS estime que la pandémie de COVID-19 pourrait mener à une augmentation du nombre de décès annuel par la tuberculose de 200.000 à 400.000 par rapport aux prédictions antérieures. Ces dernières années, d'importants progrès avaient pu être réalisés pour diminuer l'impact de la tuberculose à l'échelle mondiale mais la pandémie pourrait retarder le combat contre l'épidémie de tuberculose d'une décennie selon les prédictions de l'OMS[1].

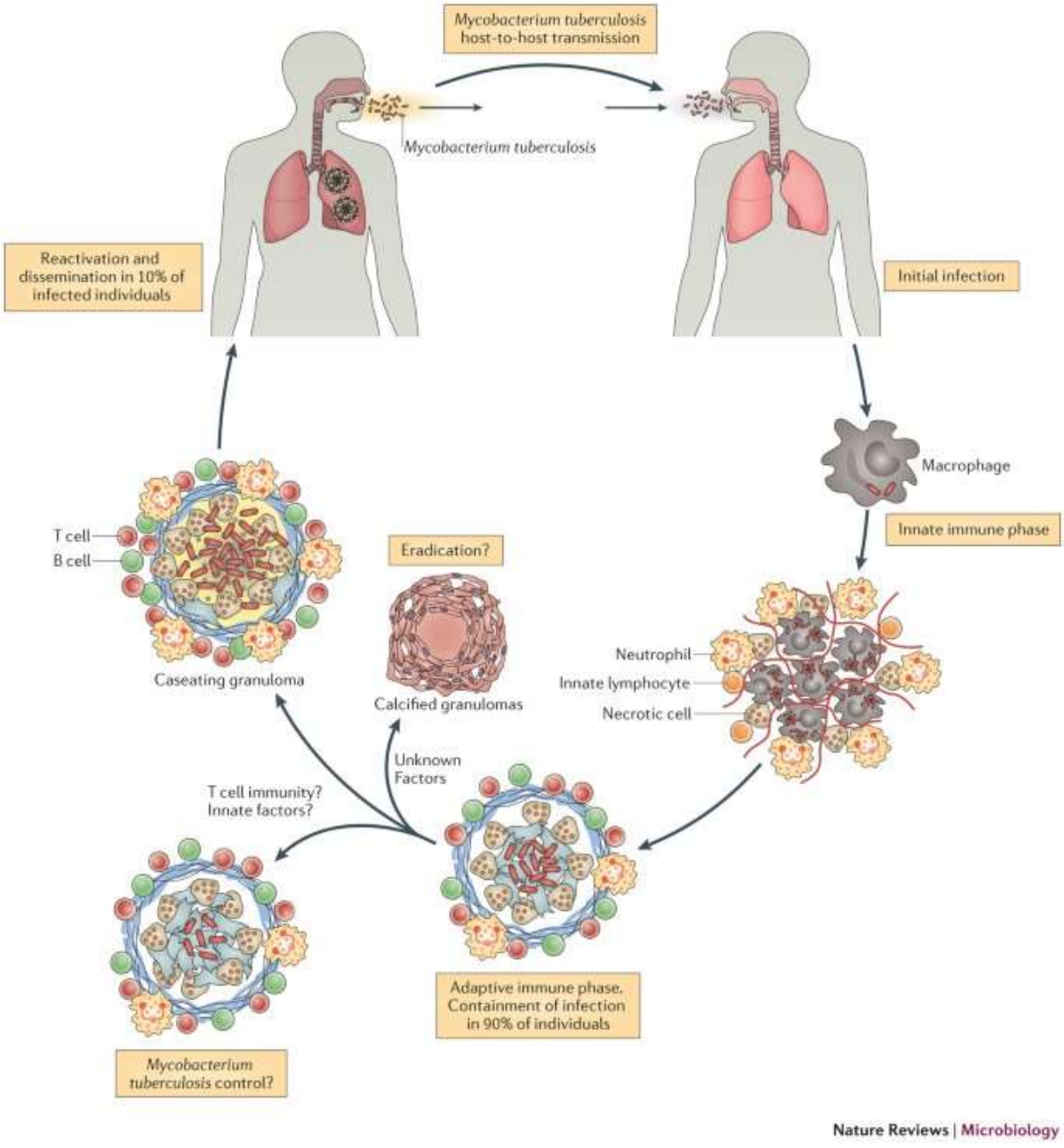


Figure 2. Pathogénèse de la tuberculose. (Adapté de[5]). L'infection est initiée par l'inhalation de gouttelettes libérées par un individu malade de tuberculose active. La réponse immunitaire innée débute par le recrutement de cellules phagocytaires au site d'infection. Survient ensuite la réponse immunitaire adaptative permettant de contenir l'infection dans un granulome. La majorité des infections mènent à une tuberculose latente qui peut se réactiver dans 10 % des cas.

b) Physiopathologie de la tuberculose

Le bacille de la tuberculose, identifié il y a plus d'un siècle par Robert Koch, affecte l'humanité depuis plusieurs milliers d'années ainsi qu'en témoignent les traces les plus anciennes de la tuberculose humaine connues à ce jour retrouvées sur un squelette datant de 9000 ans[6]. Cette longue relation entre l'homme et le bacille de Koch aurait conduit à la coévolution de l'agent pathogène et du système de défense immunitaire de son hôte[7]. L'incroyable capacité du bacille à exploiter à son avantage les mécanismes de défense de l'hôte infecté en se soustrayant à la réponse immune pour persister dans cet environnement hostile pendant des années est une des clés du succès de l'infection par *Mycobacterium tuberculosis*.

En effet, une des particularités de l'infection par le bacille de Koch est qu'elle se traduit rarement par un développement symptomatique de la maladie. Après son entrée dans l'organisme, la bactérie est contenue dans un granulome où elle persiste généralement dans un état de dormance. L'hôte est alors infecté, mais ne développe pas les symptômes de la tuberculose et n'est pas contagieux. Il est ainsi estimé qu'un tiers de la population mondiale est actuellement atteint de tuberculose dite latente[8], ce qui forme un formidable réservoir biologique pour la bactérie. On estime que seuls 10% des personnes infectées par le bacille tuberculeux risquent de développer une forme symptomatique lors de leur vie[9]. Cette prévalence augmente avec certaines comorbidités principalement lorsque celles-ci entrainent un affaiblissement du système immunitaire. En effet, le risque de développer une forme symptomatique de la tuberculose est fortement augmenté pour les personnes souffrant de malnutrition, de diabète ou du SIDA par exemple.

Les raisons de l'histoire naturelle particulière de cette maladie résident dans les mécanismes sous-jacents de l'infection par *M. tuberculosis* (**Figure 2**). Le bacille tuberculeux est inhalé par l'hôte sous forme de gouttelettes libérées dans l'atmosphère par la toux d'une personne atteinte d'une tuberculose active symptomatique. La bactérie atteint alors les poumons dans lesquels elle sera reconnue par les cellules du système immunitaire au niveau des alvéoles pulmonaires. *M. tuberculosis* est une bactérie intracellulaire, qui va alors résider dans les cellules phagocytaires professionnelles des poumons et en particulier dans les macrophages alvéolaires. La réponse immune innée se met alors en place : l'agent pathogène

va être reconnu via les « pattern recognition receptors » (PRR) de ces cellules entrainant *in fine* une cascade de signalisation cytokiniques[10]. Les cytokines vont promouvoir une inflammation localisée au site d'infection et vont participer au recrutement de nouveaux acteurs de la réponse immunitaire. Ces nouvelles populations de cellules vont comprendre des cellules présentatrices d'antigènes et notamment les cellules dendritiques qui vont pouvoir amorcer la réponse immunitaire adaptative.

Ces cellules dendritiques vont reconnaitre des motifs moléculaires bactériens conservés (ou PAMPS pour « Pathogen-Associated Molecular Patterns ») grâce à leur PRR, ce qui va entrainer leur maturation et leur migration vers les ganglions lymphatiques. Dans ces organes lymphoïdes secondaires, les cellules dendritiques induisent l'activation et la maturation de lymphocytes T naïfs. Les lymphocytes T vont alors se multiplier et migrer à leur tour vers les poumons[11]. Sur le site infectieux, les lymphocytes T vont alors produire des cytokines inflammatoires (interféron γ, TNFα) permettant l'induction des mécanismes microbicides des macrophages et le recrutement de nouvelles cellules immunitaires nécessaires à l'élimination du bacille ou à défaut, à sa mise sous contrôle.

Une des caractéristiques de l'infection par *M. tuberculosis* est la lenteur de la mise en place de cette réponse immunitaire adaptative. Alors que pour la majorité des infections, l'immunité adaptative se met en place après 3 à 5 jours, pour la tuberculose, la réponse par les lymphocytes T est initiée 10 jours après l'infection[12]. Cette réponse retardée est perçue comme l'une des nombreuses stratégies mises en place par *M. tuberculosis* et qui en font un pathogène si efficace puisqu'ainsi la bactérie peut fortement se multiplier avant que la réponse adaptative ne se déclenche[13].

Le recrutement important de ces cellules immunitaires dans les poumons mène à la formation du granulome. Le granulome est une formation multicellulaire qui permet de contenir l'infection en isolant le bacille du reste de l'organisme tout en en limitant la multiplication. Cette structure tissulaire se compose d'un cœur de macrophages infectés, entouré de cellules phagocytaires, ceinturé d'une capsule fibreuse et de lymphocytes[14].

L'intégrité du granulome contenant la bactérie est maintenue par le système immunitaire. Un affaiblissement de celui-ci peut mener à un échappement du bacille tuberculeux vis-à-vis de son contrôle conduisant à une reprise de la multiplication bacillaire et à une réactivation de la maladie. Il s'ensuit une perte de structuration du granulome qui

résulte en une libération, dans les poumons, des bactéries qui peuvent alors être expectorées dans l'environnement menant à l'infection d'autres individus.

La baisse des défenses immunitaires menant à la rupture du granulome peut être causée par plusieurs facteurs comme la malnutrition, le diabète ou le SIDA comme mentionné précédemment. On note en effet une forte synergie entre le virus de l'immunodéficience humaine (VIH) et la tuberculose[15]. Les deux pathogènes travaillent de concert accélérant le déclin du système immunitaire de l'hôte infecté. La coinfection par le VIH et la tuberculose augmente le risque de réactivation d'une tuberculose latente d'un facteur 20[16]. De même, la tuberculose semble exacerber l'infection par le VIH en créant un microenvironnement propice à la diffusion du virus[17].

c) Moyen de luttes contre la tuberculose

Malgré les progrès réalisés sur la compréhension de la physiopathologie de la maladie, les moyens de lutte disponibles, bien que d'une efficacité éprouvée, marquent le pas et montrent leurs limites devant l'augmentation des comorbidités complexes et l'émergence des souches infectieuses chimiorésistantes.

- Le vaccin BCG

À l'heure actuelle le seul traitement prophylactique contre la tuberculose est le vaccin BCG découvert en 1921, il y a précisément 100 ans cette année. Ce vaccin est préparé à partir d'une souche atténuée de *Mycobacterium bovis* vivante. Le bacille de Calmette-Guérin (BCG), du nom de ses développeurs, correspond donc à une bactérie qui a perdu de sa virulence sur l'homme à la suite de cultures successives. Cette souche bactérienne vaccinale est le seul vaccin préconisé par l'OMS, mais son efficacité fait controverse[18]. En effet, des études soulignent l'hétérogénéité d'efficacité du BCG, notamment dans la prévention de la tuberculose chez l'adulte. Cependant, un consensus existe quant à son efficacité pour la protection contre certaines formes graves de tuberculose extrapulmonaire chez l'enfant.

Tableau 1. Liste des chimiothérapies antituberculeuses. Adapté de [19].

Molécule	Cible	Mécanisme affecté	Effet
Antituberculeux de première ligne			
Rifampicine	Sous unité β de l'ARN polymérase	Inhibe la synthèse d'ARN	Bactéricide
Isoniazide	InhA	Inhibe la biosynthèse des acides mycoliques	Bactéricide
Pyrazinamide	Sous unité ribosomale 30S Membrane (transport et potentiel éngergétique) PanD mas et ppsA-E	Inhibe trans-traduction des protéines Inhibe les transport et énergie membranaires Inhibe la synthèse du co-enzyme A et du pantothenate Inhibe la synthèse des DIM	Bactéricide (pH 5,5-6) / Bactériostatique
Ethambutol	Arabinosyl transférases	Inhibition de la synthèse de l'arabinogalactane Biosynthèse des phospholipides Métabolisme des ARNs	Bactériostatique
Antituberculeux de deuxième ligne injectables			
Kanamycine, amikacine et Streptomycine	Unité ribosomale 30S	Inhibe la synthèse protéique	Bactéricide
Capréomycine	Unités ribosomales 30S et 50S	Inhibe la synthèse protéique	Bactéricide
Antituberculeux de deuxième ligne			
Ethionamide	InhA	Inhibe la biosynthèse des acides mycoliques	Bactériostatique
Thiocetazone	-	Inhibe la biosynthèse des acides mycoliques	Bactériostatique
Acide para-amino salicylique (PAS)	DHFR	Inhibition de la synthèse de l'acide folique Inhibition de la capture du fer	Bactériostatique
Fluoroquinolones			
Ofloxacine, Lévofloxacine et Moxifloxacine	ADN gyrase et ADN topoisomérase	Inhibe la réplication de l'ADN	Bactéricide

- Chimiothérapie

Le traitement de la tuberculose actuellement préconisé par l'OMS et les autorités de santé consiste en une thérapie à base d'antituberculeux pendant 6 à 9 mois. Ce traitement se compose généralement d'une phase intensive de 2 mois suivie d'une phase d'entretien de 4 à 7 mois. La phase initiale consiste en une combinaison de 4 antituberculeux de première intention : l'isoniazide, l'éthambutol, la rifampicine et la pyrazinamide (**Tableau 1**). Au terme de cette première phase, la contagiosité du malade diminue fortement ainsi que les symptômes de la maladie. La deuxième phase va se composer d'un traitement plus léger composé de rifampicine et d'isoniazide seulement. L'objectif de cette phase d'entretien est l'élimination des bacilles restants pour prévenir toute rechute.

Ce schéma thérapeutique contre la tuberculose peut cependant varier. En effet, il peut être adapté en fonction des comorbidités du malade. Par exemple, dans le cas d'une co-infection au VIH, des alternatives à la rifampicine peuvent être envisagées du fait d'interaction médicamenteuse possible avec les agents antiviraux[20]. Aussi, un des problèmes majeurs dans le traitement de la tuberculose actuellement est l'apparition de souches résistance à plusieurs des traitements utilisés[1]. Ceci mène à l'utilisation d'antibiotiques alternatifs qui présentent parfois des effets secondaires importants (**Tableau 1**).

d) Chimiorésistance : Limite de l'arsenal thérapeutique.

Par définition, une tuberculose multirésistante est une tuberculose qui ne répond pas à au moins l'isoniazide et la rifampicine. L'OMS estime qu'en 2019, 0,5 million de personnes ont développé une tuberculose multirésistante représentant près de 5% des nouveaux cas de tuberculose détectés[1]. Ce taux est variable puisqu'il peut atteindre jusqu'à 35% des nouveaux cas dans certains pays comme la Russie ou la Moldavie par exemple (**Figure 1B**).

L'émergence de ces résistances est principalement due à une mauvaise observance du traitement contre la tuberculose. Du fait de sa durée, de sa lourdeur et de son coût, le traitement antituberculeux est parfois incorrectement suivi ou interrompu avant son terme par le patient. Ce type de comportement participe à la sélection de souches résistantes qui peuvent alors être transmises.

Pour faire face à ce problème grandissant, des antituberculeux de seconde ligne sont utilisés parmi lesquels figurent les fluoroquinolones, l'amikacine, la kanamycine ou la capréomycine (**Tableau 1**). Cependant, des souches ultrarésistantes qui ne sont plus sensibles à ces traitements de seconde ligne sont maintenant également identifiées[1].

e) Prospectives thérapeutiques

En 2014, un ambitieux plan international « End TB » a été voté par l'OMS pour contrôler la maladie et mettre fin à l'épidémie mondiale de tuberculose. Ce programme repose sur trois points : une meilleure prévention et prise en charge des patients, un engagement politique dans le soin de la tuberculose et des efforts dans le développement de nouvelles stratégies thérapeutiques. Ce dernier point repose sur la mise au point de nouvelles molécules actives contre *Mycobacterium tuberculosis*.

(1) Limites à l'identification de nouvelles molécules thérapeutiques

Toutefois, de manière générale la découverte de nouveaux composés antibactériens a ralenti ces dernières années. Cette diminution n'est pas due à une chute des efforts de recherche et développement, mais à une difficulté croissante d'identifier de nouvelles molécules utilisables en thérapie anti-infectieuse. Une campagne d'identification d'antibactériens classiques repose sur le criblage haut débit de bibliothèques de composés au moyen de tests conçus pour détecter les molécules possédant les propriétés désirées (d'inhibition de la croissance cellulaire ou de la pathogénicité de la bactérie par exemple). Les molécules actives initiales ou « hits » sont ensuite caractérisées et optimisées par ingénierie chimique pour obtenir le meilleur composé utilisable en clinique. Cette stratégie montre des taux de succès de plus en plus réduits pour diverses raisons parmi lesquelles la faible solubilité des hits identifiés *in vitro* qui atteignent difficilement leur cible *in vivo*, ou une faible spécificité avec des composés actifs ciblant l'organisme pathogène, mais également des cellules de l'hôte…

Pour améliorer les chances de succès de telles stratégies de criblage, il est essentiel d'avoir une meilleure compréhension des cibles afin de développer des tests de criblages plus résolutifs pour la découverte de nouvelles molécules thérapeutiques.

(2) Nouvelles stratégies

Bien que la recherche de nouvelles antibiothérapies pour le traitement de la tuberculose constitue encore l'approche majoritaire, de nouvelles stratégies non conventionnelles qui ne généreraient pas de résistances sont explorées, notamment les thérapies ciblées sur l'hôte ou les stratégies anti-virulences.

Les thérapies ciblées sur l'hôte vont avoir pour rôle de médier et d'accompagner la réponse immunitaire face au pathogène. Cela va se traduire par l'utilisation de composés permettant de réduire l'inflammation, de moduler l'autophagie ou de potentialiser la réponse immunitaire adaptative. Plusieurs essais cliniques pour le développement de tels composés efficaces lors de la tuberculose sont en cours[21].

Aujourd'hui, le ciblage médicamenteux des facteurs de virulence mycobactériens non essentiels pour la survie de la bactérie semble être une voie prometteuse pour réduire la pathogénicité de *M. tuberculosis* sans induire de pression de sélection responsable de la résistance aux antibiotiques. Les traitements anti-virulence présentent donc comme principal avantage de diminuer la sélection de souches résistantes. De plus, ce type de thérapie pourrait présenter des effets indésirables moins importants que les antibiotiques classiques. En effet, de telles molécules ciblant les facteurs de virulence de *M. tuberculosis* seraient spécifiques et n'auraient pas d'effets néfastes sur le microbiote humain puisque la majorité des bactéries commensales n'expriment pas ces mêmes facteurs de virulences.

Les stratégies anti-virulence ont déjà été utilisées avec succès pour d'autres maladies telles que le botulisme infantile[22]. Dans ce cas-ci, le traitement réduit fortement le temps d'hospitalisation, diminue la gravité de la maladie et diminue les coûts médicaux pour guérir de la maladie. Il existe également des médicaments anti-virulence qui ciblent la toxine du charbon et qui sont utilisés en tant qu'adjuvants de traitements antibiotiques[23,24]. Malgré l'existence de ces traitements anti-virulences, ce type de stratégie thérapeutique est encore très innovante et est donc peu répandue en clinique. Il existe peu de recul quant à leur

utilisation et la majorité des avantages escomptés de ce type de thérapie doivent donc toujours être vérifiés. Ils offrent cependant une alternative prometteuse face aux nombreux désavantages des antibiotiques.

La multiplicité des facteurs de virulence chez *M. tuberculosis* offre un formidable potentiel de cibles de traitements anti-virulence[25]. Actuellement, des molécules anti-virulence ciblant *M. tuberculosis* sont en essais précliniques[21,26]. Celles-ci ciblent des facteurs de virulence bien identifiés de *M. tuberculosis* comme le lipoarabinomannane : un lipoglycanne majeur de la paroi mycobactérienne[27].

2. *Mycobacterium tuberculosis*

a) Les mycobactéries

Mycobacterium tuberculosis appartient au genre *Mycobacterium*, famille des *mycobacteriaceae*, ordre des *actinomycetales*. Outre *M. tuberculosis*, le genre des *Mycobacterium* regroupe de nombreuses espèces dont certaines pathogènes chez l'homme à l'origine de pathologies graves telles que la lèpre provoquée par *Mycobacterium leprae,* l'ulcère de Buruli provoqué par *Mycobacterium ulcerans* ou le granulome des aquariums transmis par *Mycobacterium marinum*. À côté de celles-ci, d'autres espèces mycobactériennes pathogènes opportunistes telles que *M. Avium, M. kansasii* ou encore *M. abcessus* sont responsables de pathologies pulmonaires graves. Ainsi, la mycobactérie opportuniste *M. abcessus* est très régulièrement retrouvée chez des patients atteints de mucoviscidose entrainant une dégradation des fonctions respiratoires de ces patients alors que *M. Avium, M. kansasii* sont responsables de pathologies cliniquement et radiologiquement indiscernables de la tuberculose et généralement associées à des contextes d'immunodépression.

Ces espèces ont été regroupées, avec plus d'une centaine d'autres sous la qualification de Mycobactéries Non Tuberculeuses par opposition aux mycobactéries dites du « complexe *tuberculosis* » comprenant toutes les espèces responsables de tuberculose chez l'homme ou l'animal et phylogénétiquement associées à *M. tuberculosis* telles que *Mycobacterium bovis, Mycobacterium africanum* ou *Mycobacterium canetti* pour les plus représentatives.

Les mycobactéries sont également communément classées selon leur temps de génération définissant leur vitesse de croissance. Alors que les espèces pathogènes strictes sont toutes des bactéries à croissance lente (avec un temps de génération de 20h), certaines espèces de mycobactéries non tuberculeuses peuvent être à croissance rapide (temps de génération de 4-6h) comme le pathogène opportuniste *M. abcessus* ou la mycobactérie saprophyte non pathogène *Mycobacterium smegmatis* régulièrement utilisée en laboratoire comme modèle de substitution pour l'étude de *M. tuberculosis* en raison de ses propriétés de croissance rapide et de son caractère non pathogène.

Parmi les caractéristiques des mycobactéries, il est important de souligner 2 points importants qui les distinguent particulièrement des autres bactéries : du point de vue génomique, les mycobactéries se caractérisent par un taux de GC élevé pouvant aller de 61 à 71% ; du point de vue phénotypique les mycobactéries se distinguent des autres bactéries Gram positive et négative par une paroi à l'architecture et à la composition unique (**Figure 3**).

b) La paroi mycobactérienne et facteurs de virulence

La paroi mycobactérienne est, en effet, particulièrement riche en lipides et contient des molécules uniques dans le monde bactérien telles que les acides mycoliques, acides gras à très longues chaines (C80) propres aux *Mycobacteriaceae*. L'organisation unique de cette paroi et sa richesse en composés singuliers fortement hydrophobes lui confèrent une grande imperméabilité à de nombreux antibiotiques et participe à la virulence des mycobactéries.

L'enveloppe mycobactérienne se compose de la membrane plasmique, du périplasme, de la paroi et d'une capsule (**Figure 3**). La membrane plasmique mycobactérienne a une composition classique en phospholipides alors que la paroi et la capsule ont une composition complexe en lipides, glycoconjugués et protéines.

La paroi mycobactérienne se définit par différentes strates qui se composent successivement de peptidoglycane, d'arabinogalactane puis de la mycomembrane (**Figure 3**). La mycomembrane présente une épaisseur similaire à la membrane plasmique. Son feuillet interne est composé d'acides mycoliques liés à l'arabinogalactane, lui-même lié au peptidoglycane. Cette structure forme ainsi le complexe mAGP (mycoloyl-arabinogalactan-

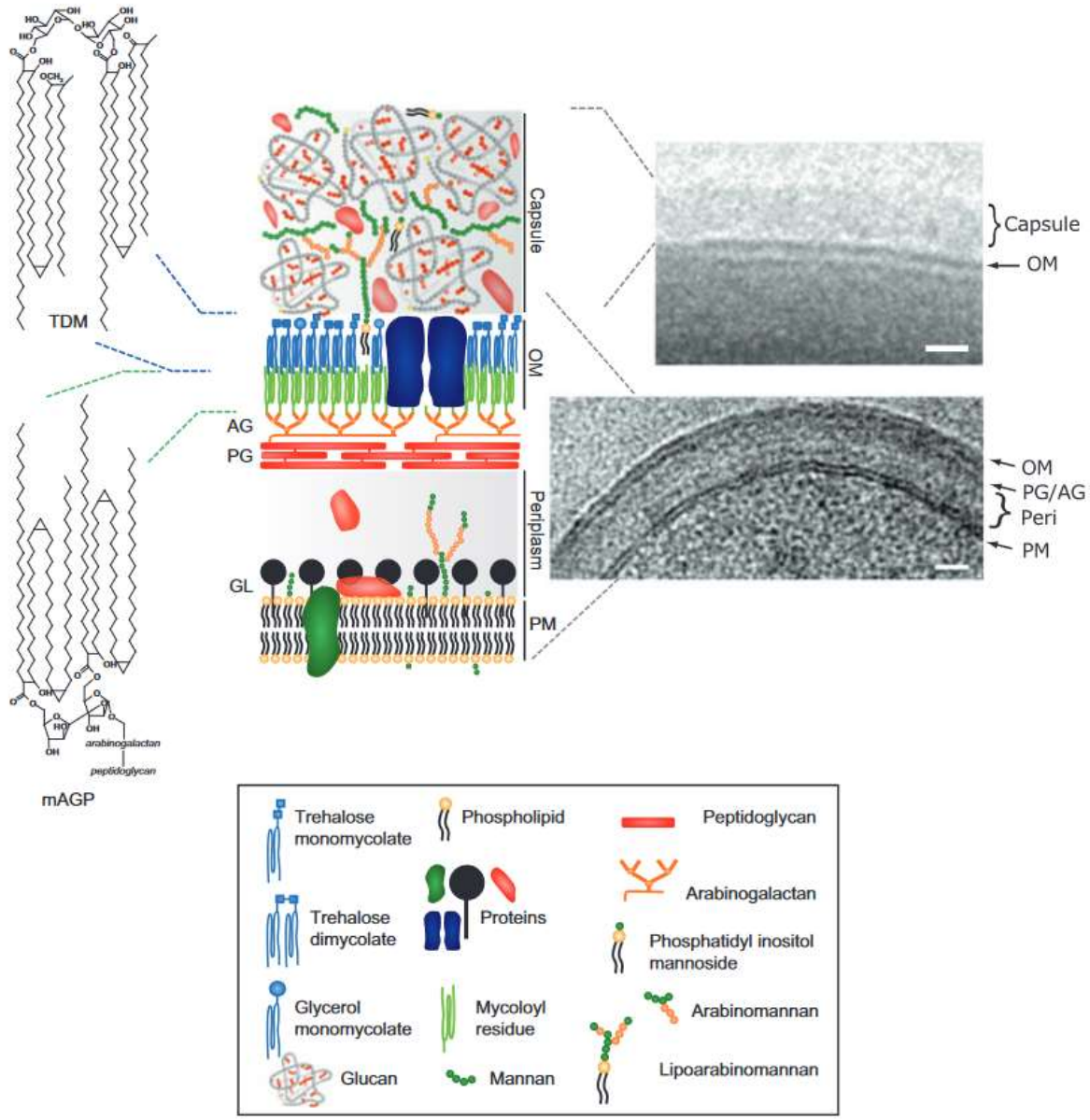


Figure 3. Modèle de l'enveloppe mycobactérienne. Adapté de [28]. L'enveloppe mycobactérienne se compose de la membrane plasmique (PM), du périplasme (Peri), de la mycomembrane (OM) et d'une capsule. La membrane plasmique mycobactérienne a une composition classique en phospholipides alors que la paroi (composé du mAGP et de la mycomembrane) et la capsule ont une composition complexe en lipides, glycoconjugués et protéines. Au-dessus de la membrane plasmique, le périplasme (Peri) contient une couche granuleuse (GL) composée de protéines. A droite, images de microscopie électronique sur mycobactérie entière mettant en évidence les différentes structures de l'enveloppe (adapté de [29] pour l'image du haut et [30] pour l'image du bas), Barre d'échelle = 20nm. A gauche, représentation du complexe mAGP (mycoloyl-arabinogalactan-peptidoglycan), qui se compose de peptidoglycane (PG) et d'arabinogalactane (AG), formant la partie inférieure de la mycomembrane, et représentation du tréhalose dimycolate (TDM) qui est un des nombreux composants de la partie supérieure de la mycomembrane.

peptidoglycan) qui est la couche basale de la paroi mycobactérienne. Le feuillet externe de la mycomembrane présente une plus grande richesse en composés : on y retrouve des molécules telles que des mono- ou dimycolates de tréhalose (TDM) et du glycérol monomycolate.

Au-delà de la mycomembrane, on retrouve alors la capsule qui est une composante spécifique des espèces pathogènes et n'est, par exemple, pas retrouvée chez *M. smegmatis*[30]. La capsule se compose majoritairement de composés polysaccharidiques parmi lesquels on trouve également protéines et lipides.

Plusieurs composés de cet ensemble ont été démontrés comme participant à la virulence de *M. tuberculosis*[25]. En effet, la paroi mycobactérienne se situe à l'interface avec l'hôte et représente un réservoir de facteurs de virulence. Le rôle des acides mycoliques dans la virulence de *M. tuberculosis* est par exemple très bien documenté. Il a été démontré que leurs fonctions oxygénées étaient importantes pour le maintien de la perméabilité de la mycomembrane et que leur perte entrainait une atténuation de la virulence mycobactérienne chez la souris[31]. Les acides mycoliques peuvent aussi être décorés de motifs cyclopropanes et la perte de la machinerie cellulaire responsable de cette modification entraine une atténuation sévère de la virulence de la bactérie chez la souris [32]. De façon similaire, la mutation de gènes impliqués dans la synthèse d'un autre composant de la paroi, le phosphatidylinositol mannoside, entraine un effet notable sur la pathogénicité de *M. tuberculosis*[25].

On peut aussi noter d'autres molécules composant la paroi mycobactérienne et impliqué dans la virulence de la bactérie. Le lipoarabinomannane et son précurseur synthétique, le lipomannane, sont tous deux des composant de la paroi impliqué dans différents mécanismes tel que l'inhibition de la maturation du phagosome[33–35]. Par ailleurs, le rôle dans la virulence du lipoarabinomannane semble lié à LprG, une mannoprotéine impliquée dans l'export du LAM[36]. Les mannoprotéines sont des protéines glycosylées qui peuvent être associées à la paroi mycobactérienne et plusieurs d'entre elles ont été identifiées comme impliquées dans la virulence de *M. tuberculosis* (**Figure 4**).

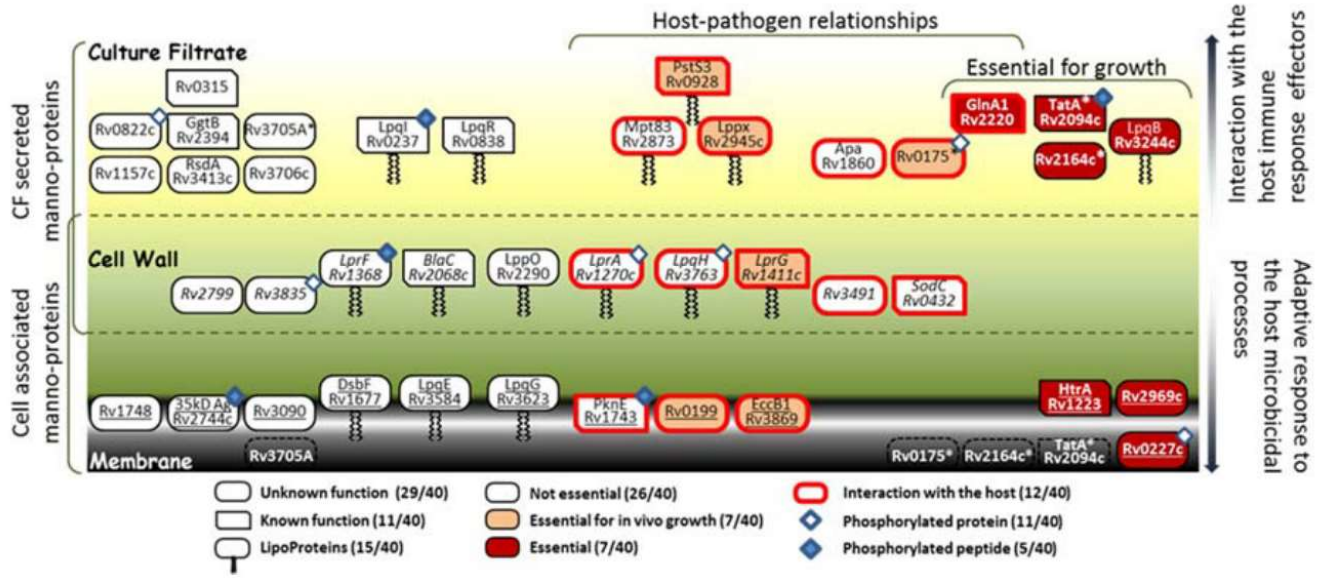


Figure 4. Distribution des mannoprotéines identifiées à l'interface avec l'hôte. Adapté de [37]. La distribution topologique des mannoprotéines est représentatives des résultats obtenus par Tonini *et al,* 2020[37]. L'essentialité et les propriétés physiologiques ou immunomodulatrices de chaque protéine sont identifiées.

c) Les mannoprotéines de M. tuberculosis comme facteur de virulence.

Un nombre important de mannoprotéines sont recensées chez les mycobactéries[37,38]. Bien que les connaissances sur le rôle de cette modification traductionnelle restent parcellaires, plusieurs glycoprotéines mycobactériennes ont été identifiées comme impliquées dans des processus de virulence.

À ce jour, plus de 60 mannoprotéines ont été identifiées chez *Mycobacterium tuberculosis*[37,38], dont moins d'une dizaine pour lesquelles les motifs et profils glycaniques ont pu être spécifiquement caractérisés (Apa[39–44], MPT83[45], LpqH[46,47], SodC[48], PstS-1[49] et LprG[50]).

La mannoprotéine la plus étudiée à ce jour est Apa (alanine and proline rich protein). Cette protéine est un antigène immunodominant sécrété par la bactérie[39]. La présence de mannose sur cette protéine a tout d'abord été démontrée par sa capacité à être reconnu par la lectine concanavaline A[40]. Cette modification post-traductionnelle est importante pour le rôle immunologique de Apa puisque le changement du profil de glycosylation de cette protéine, ou sa déglycosylation totale entraine une diminution de ses capacités à engendrer une réponse immunitaire des lymphocytes T chez le cochon d'inde[41,42].

De plus, Apa a été prouvée comme capable de lier des lectines de type-C, comme DC-SIGN et SP-A, contribuant ainsi à l'invasion cellulaire par la bactérie[43,44]. Ce type de comportement a pu être mis à jour pour d'autres mannoprotéines. Elles pourraient ainsi avoir un rôle d'adhésine favorisant l'internalisation de *M. tuberculosis* dans les macrophages et facilitant par conséquent l'infection de ces bactéries qui résident et se multiplient dans les macrophages. Il a par exemple été décrit que plusieurs mannoprotéines mycobactériennes sont capables d'interagir avec le récepteur au mannose (comme LpqH[46] ou PstS-1[51]), médiant ainsi la phagocytose bactérienne. Dans le cas de LpqH par exemple, cette glycolipoprotéine a été démontrée comme une adhésine majeure qui lie le récepteur au mannose des monocytes THP-1 contribuant ainsi à l'internalisation par phagocytose du bacille de la tuberculose[46].

Certaines mannoprotéines peuvent être impliquées dans la capacité des bactéries à survivre lors d'une infection, par exemple SodC est une mannoprotéine présentant une activité superoxyde dismutase assurant une protection contre les radicaux libres d'oxygène produit par le macrophage pour combattre les bactéries[48].

Au final, la compréhension du rôle de cette modification post-traductionnelle reste élusive. Un lien entre virulence et mannosylation semble exister, mais le faible nombre de protéines étudiées ne permet pas d'affirmer cette corrélation. La preuve formelle de leur rôle dans la virulence fut apportée par l'équipe. La mannosylation de toutes ces protéines est initiée par une enzyme unique : la protéine *O*-mannosyltransférase de *Mycobacterium tuberculosis* (PMTub). Dans ce contexte, l'équipe a démontré que la réaction catalysée par PMTub n'est pas essentielle pour la croissance bactérienne, mais cruciale pour la virulence[52]. De ce fait, PMTub constitue une cible attractive pour le développement de nouveaux médicaments anti-virulence. L'effet observé après interruption du gène de PMTub n'est pas expliqué à l'échelle moléculaire, mais il est estimé que la perte de la mannosylation des protéines pourrait avoir un effet pléiotrope important sur la virulence au travers des nombreuses mannoprotéines mycobactériennes.

Les données expérimentales montrent une conservation évolutive des voies de *O*-mannosylation des protéines entre les eucaryotes et certaines bactéries[53–55]. La forte conservation des protéine *O*-mannosyltransférases (PMT) dans le monde vivant (animaux, champignons et certaines bactéries) est un défi pour l'identification d'inhibiteurs de PMTub à potentiel thérapeutique. En effet, il y a une forte identité de séquence entre PMTub et ses homologues humains essentiels, POMT1 et POMT2, pour lesquels des mutations sont à l'origine de dystrophies sévères et de décès prématurés[56]. Le défi consiste donc à identifier des inhibiteurs de PMTub qui n'affectent pas l'activité des PMT humaines.

Présentation générale des travaux

Nous avons vu dans l'introduction l'importance que revêt l'identification de nouvelles stratégies thérapeutiques contre *M. tuberculosis*. Dans le contexte actuel d'émergence de souches multirésistantes aux antibiotiques, le développement de stratégies alternatives aux antibiotiques telles que les thérapies anti-virulences semblent donc une piste prometteuse. Ainsi, l'équipe a mis en évidence que PMTub, la protéine O-mannosyltransférase de *M. tuberculosis*, est un enzyme non essentielle à la survie de la bactérie, mais cruciale pour sa virulence[52]. PMTub semble donc une cible anti-virulence attractive. Or, les PMT humaines sont des enzymes essentielles et elles sont fortement conservées avec PMTub. Pour que PMTub constitue une cible thérapeutique acceptable, il faut donc être capable d'identifier des inhibiteurs spécifiques de l'enzyme mycobactérienne et qui ne ciblent pas les isoenzymes humaines.

Pour cela, l'objectif de cette thèse est de développer et d'appliquer les outils nécessaires à l'identification d'inhibiteurs spécifiques de PMTub. Précédemment à cette thèse, peu de données fonctionnelles ont été recueillies sur PMTub. Or ce type d'informations est nécessaire au développement d'un test de criblage performant d'inhibiteurs spécifiques de PMTub. Une étude bibliographique des protéine O-mannosyltransférases, et plus généralement de glycosyltransférases partageant des caractéristiques proches, a permis de mettre en évidence des similarités qui pourraient s'appliquer à PMTub (Rappels bibliographiques, p.33).

Afin de pouvoir confirmer ces similarités, je me suis, dans un premier temps, appliqué à l'étude des relations structure/fonction de PMTub. Pour cela, nous verrons dans une première partie de ce travail le développement du test d'activité nécessaire à ce type d'étude (Partie 1, p.68). Nous avons décidé de développer un test d'activité quantitatif et phénotypique permettant ainsi de suivre l'activité de l'enzyme dans le contexte de la bactérie vivante et qui pourra donc être adapté au crible d'inhibiteurs capables de traverser l'enveloppe mycobactérienne et d'atteindre PMTub à la membrane plasmique. Ceci passe par la mise au point d'une souche mycobactérienne rapportrice de l'activité de PMTub et le développement d'un système de détection de la mannosylation d'une protéine témoin. Ce

test a été pensé pour permettre l'étude des relations structure-fonction de PMTub, mais aussi pour être adaptable au crible d'inhibiteurs de PMTub.

L'étude des relations structure/fonction de PMTub a été réalisée dans une seconde partie (Partie 2, p.98). Cette étude passe par l'élaboration d'un modèle 3D de PMTub. Ce modèle a permis d'identifier un site catalytique putatif et des domaines d'intérêt pour le ciblage spécifique de PMTub par des inhibiteurs. L'étude de ces motifs identifiés a été réalisée à l'aide du test d'activité phénotypique préalablement développé, mais aussi par un ensemble d'approches *in vitro* complémentaires.

Dans une dernière partie, nous avons utilisé le test construit en partie 1 et les connaissances acquises en partie 2 pour développer des approches de crible d'inhibiteur de PMTub (Partie 3, p.155). La meilleure compréhension des relations structure/fonction nous a permis de sélectionner rationnellement des molécules pouvant avoir un effet inhibiteur sur PMTub. Après une optimisation du test mis en place en partie 1 nous avons pu tester ces composés pour leur capacité à affecter la mannosylation des protéines mycobactérienne. En parallèle de cette approche, nous avons développé les outils pour un test de criblage plus innovant et résolutif. Ce test utilise des outils de biologie synthétique et vise à identifier des inhibiteurs spécifiques de PMTub.

L'ensemble des approches utilisées lors de cette thèse vise donc à l'approfondissement des connaissances fonctionnelles de PMTub, au développement de nouveaux outils pour l'étude de la O-mannosylation des protéines mycobactériennes et à la mise au point de stratégies de cible d'inhibiteurs de PMTub ayant un fort potentiel thérapeutique.

Rappels bibliographiques : Mannosylation des protéines et Glycosyltransférases

A. Protéine *O*-mannosylation et PMT

1. La O-mannosylation des protéines

a) Biosynthèse des mannoprotéines

Parmi les processus de glycosylation post-traductionnels, la *O*-mannosylation des protéines est conservée depuis les bactéries jusqu'à l'homme. Cette modification consiste en la formation d'une liaison α-glycosidique entre la fonction hydroxyle de la chaine latérale d'une sérine ou d'une thréonine et l'hydroxyle anomérique d'un résidu mannose.

Le mannose transféré provient toujours d'un donneur lipidique activé de type polyisoprenoyl–phospho-mannose[57,58]. Ce substrat lipidique-phospho-β-mannosyl, donneur de mannose est synthétisé par une polyprenyl phospho-mannosyl synthase à partir de GDP-mannose[57], lui-même dérivé du fructose-6-phosphate, intermédiaire de la voie de la glycolyse, par une série de réactions enzymatiques impliquant successivement une phospho-mannose isomérase, une phospho-mannose-mutase et une GDP-mannose[57].

Il a longtemps été pensé que la O-mannosylation des protéines résultant du transfert du résidu mannose à partir du polyprenyl phospho-mannosyl sur la protéine était catalysé par un type particulier de glycosyltransférases: les protéine-O-mannosyltransférases (PMT) présentes dans la majorité des opisthocontes et chez certaines bactéries du groupe des actinobacteries[52,55,59,60]. Toutefois, très récemment d'autres voies de O-mannosylation des protéines ont été mises en évidence[61,62] chez les vertébrés, notamment la voie de O-mannosylation des cadhérines catalysée par une nouvelle famille de protéines les TMTC (transmembrane and tetratricopeptide repeat-containing protein). Cette voie présente des similarités avec la voie O-mannosylation dépendante des PMT notamment en termes de localisation subcellulaire du processus de transfert dans la lumière du réticulum endoplasmique, et de nature du substrat donneur de mannose utilisé (**Figure 5**).

En revanche, ces 2 voies semblent se distinguer sur la suite des processus impliquant les résidus mannoses transférés. En effet, les modifications des cadhérines par les TMTC ne semblent donner lieu qu'à des motifs mono-mannosidiques suggérant que le transfert du mannose initial ne soit pas suivi de substitutions supplémentaires.

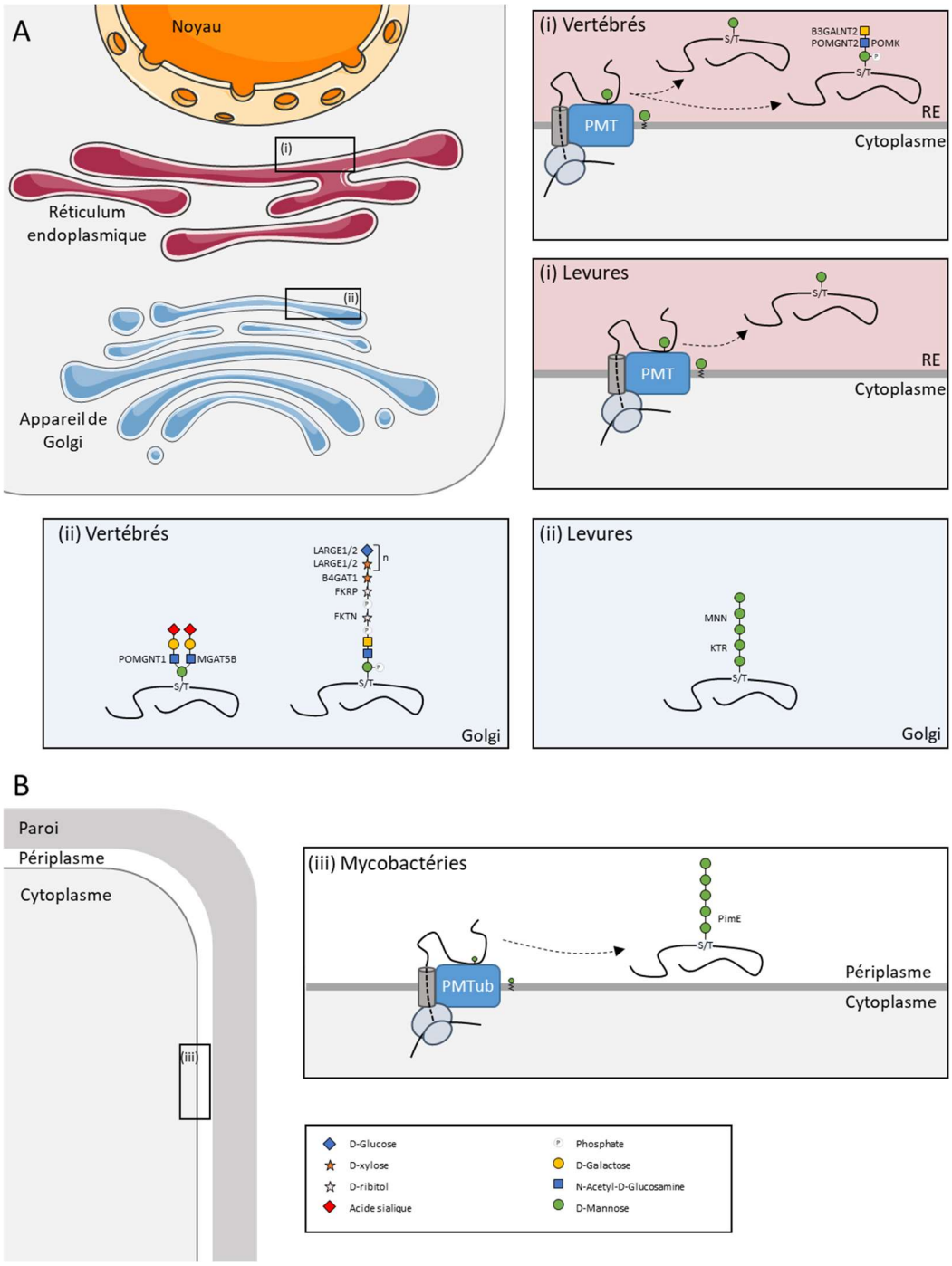


Figure 5. *O*-glycosylation des protéines dépendante de PMT. Les images ont été adaptées de [63] et [64]. Les principales étapes du processus de O-glycosylation dépendante des PMT sont représentées dans les différents compartiments cellulaires dans lesquels elles ont lieu chez **A)** les eucaryote et **B)** les bactéries. La voie de O-glycosylation des cadhérines chez les vertébrés, dépendantes de TMTC, n'est pas représentés ici mais se déroule à la membrane du réticulum endoplasmique (RE). Les enzymes impliquées dans l'élongation des chaines glycosidiques sont nommées[52,64].

À l'inverse la *O*-mannosylation dépendante des PMT donne lieu à des structures polysaccharidiques complexes[65,66]. En effet, le mannose initial transféré par la PMT dans le réticulum endoplasmique est lui-même substitué pour servir de base à la génération de structures saccharidiques plus élaborés par additions successives d'unités monosaccharidiques par des glycosyltransférases présentes dans l'appareil de Golgi[67–69] (**Figure 5**).

Les structures des glycanes venant substituer le mannose initial sont radicalement différentes entre les eucaryotes supérieurs et les levures. Chez les eucaryotes supérieurs comme les mammifères, ces glycanes présentent une grande variabilité en termes de taille, de séquence et de nature des sucres constitutifs. On peut y trouver des sucres peu fréquents comme le ribitol 5-P, des déterminants antigéniques comme l'épitope HNK-1 (SO4-GlcA[β1-3]Gal[β1-4]GlcNAc) ou des répétitions de motifs itératifs comme le disaccharide GlcA[β1 -3]Xyl[α1-3] présent sur l'α-dystroglycane[70,71].

En revanche, chez les levures, ces glycanes sont principalement des poly-mannanes plus ou moins ramifiés constitués de chaines linéaires d'unités Man avec certaines variations spécifiques d'espèce, telles que des branchements phospho-mannosyl (P-Man) chez *Saccharomyces cerevisiae* ou des résidus galactosyl (Gal) terminaux chez *Schizosaccharomyces pombe*[72] (**Figure 5**).

Chez la levure, les protéines *O*-mannosylées sont des protéines sécrétées suggérant que le processus de O-mannosylation soit associé au processus de sécrétion des protéines du réticulum endoplasmique via l'appareil de Golgi et les granules de sécrétion[73,74].

Chez les bactéries, la O-mannosylation des protéines a été détectée à ce jour seulement dans le groupe des actinobactéries[52,55,59,60]. On y observe un processus similaire à celui de la levure. La principale différence réside dans la localisation des enzymes prenant part à ce mécanisme : les bactéries ne possédant pas de réticulum endoplasmique ni d'appareil de Golgi, la O-glycosylation des protéines a lieu à la surface externe de la membrane plasmique, dans le périplasme[75] (**Figure 5**). Comme chez la levure, le mannose initial est transféré par une PMT pour être ensuite allongé par une autre mannosyltransférase[52] en une chaine oligomannosidique linéaire pouvant comporter jusqu'à 7 unités mannosyl[47]. Par ailleurs il a été montré que le processus de *O*-mannosylation chez les bactéries est également lié à la voie de sécrétion Sec des protéines[55]. Actuellement, seule la voie de *O*-mannosylation dépendante

des PMT a été étudiée chez les bactéries, mais des orthologues potentiels de TMTC chez certaines bactéries ont pu être identifié sans pour le moment avoir été fonctionnellement caractérisées.

b) Impact biologique de la O-mannosylation des protéines

Chez les eucaryotes, les défauts de O-mannosylation post-traductionnelle de protéines sont souvent liés à des manifestations physiopathologiques graves. Ainsi, des défauts fonctionnels des TMTC chez l'homme sont associés à des maladies telles que l'épilepsie, des malformations cérébrales ou la schizophrénie[76–78].

De même, des défauts congénitaux de *O*-mannosylation dépendante des PMT sont responsables de dystrophies musculaires sévères chez l'homme telles que les α-dystroglycanopathies de type syndrome de Walker-Warkurg[56]. Chez les mammifères, les PMT sont des enzymes essentielles au développement embryonnaire et leur interruption mène à une mort de l'embryon après une semaine chez la souris[79]. Des études récentes montrent aussi un rôle important de la O-mannosylation des protéines dans le maintien de la cohésion structurale des tissus des photorécepteurs de la rétine chez la souris[80]. Chez la drosophile, l'activité des protéines O-mannosyltransférase est essentiel à un développement musculaire normal, la perte de cette fonction entrainant un phénotype « twisted abdomen»[81]. Chez *Saccharomyces cerevisiae*, organisme dans lequel les PMT ont été initialement identifiées[82] et qui constitue depuis le principal modèle pour l'étude de leur fonctionnement, le processus de *O*-mannosylation des protéines a été défini comme vital. En effet, la dérégulation de l'activité PMT peut être létale pour la cellule ou peut affecter radicalement divers mécanismes importants pour la stabilité, la localisation ou la fonction de certaines protéines impliquées dans des processus physiologiques essentiels tels que la polarité cellulaire, l'intégrité de la paroi cellulaire ou la reproduction[83].

Enfin, il est à noter que les protéines O-manosylées semblent être liées à la virulence des organismes pathogènes qui expriment les PMT. Plusieurs exemples de perte de virulence lors de l'interruption des gènes des PMT ont été rapportés, notamment chez *Candida albicans*, *Cryptococcus neoformans*, ou *Mycobacterium tuberculosis*[52,84,85].

2. Les Protéine O-mannosyltransférases

Les protéine *O*-mannosyltransférases (PMT) sont décrites dans la base de données de classification de domaine protéique, Pfam[86], comme contenant au moins les domaines « PMT » (PF02366, Pfam) et « PMT_4TMC » (PF16192, Pfam). Ces deux domaines se localisent aux extrémités N- et C-terminale des PMT respectivement. Un troisième domaine intermédiaire caractéristique: le domaine MIR (PF02815, Pfam) est également généralement présent dans les séquences des PMT eucaryotes. Ce motif MIR, est ainsi nommé, car originellement décrit dans des Mannosyltransferases, les récepteurs à l'Inositol triphosphate et les récepteurs à la Ryanodine. À ce jour, les domaines MIR ont été décrits uniquement dans des protéines eucaryotes, et leur fonction reste encore largement indéterminée même si une étude récente vient d'apporter quelques indices quant à leurs rôles dans le fonctionnement du complexe PMT1-PMT2 de *Saccharomyces cerevisiae*[87,88] (décrit plus en détails p.129).

La recherche de ces différents domaines caractéristiques définissant les PMT a permis le dénombrement de plus de 5000 gènes putatifs de PMT répertoriés dans les bases de données génomiques couvrant l'ensemble des règnes du vivant. Ce type d'enzyme semble donc prévalant dans le monde vivant, mais non ubiquitaire, puisque non conservé dans les plantes, dans de nombreuses bactéries et certains opisthocontes tels que les nématodes qui sont dépourvus de gènes de PMT. De manière intéressante, même s'ils ne possèdent pas de PMT, les nématodes ne sont cependant pas dénués de *O*-mannosylation des protéines puisqu'ils expriment des TMTC[89].

Enfin, il est important de souligner que malgré le grand nombre de gènes identifiés et l'importance physiologique de ces Protéines-O-mannosyltransférases, peu d'entre elles ont pu être fonctionnellement caractérisées à ce jour.

La caractérisation fonctionnelle des PMT est généralement basée sur la mise en évidence du rôle essentiel d'un doublet caractéristique d'acides aminés acide aspartique et acide glutamique (DE), présent dans toutes les PMT étudiées[54]. Ce motif DE est localisé dans le domaine « PMT » (PF02366, Pfam) et est considéré, du fait de son essentialité, comme participant à la catalyse de la réaction de transfert de mannose.

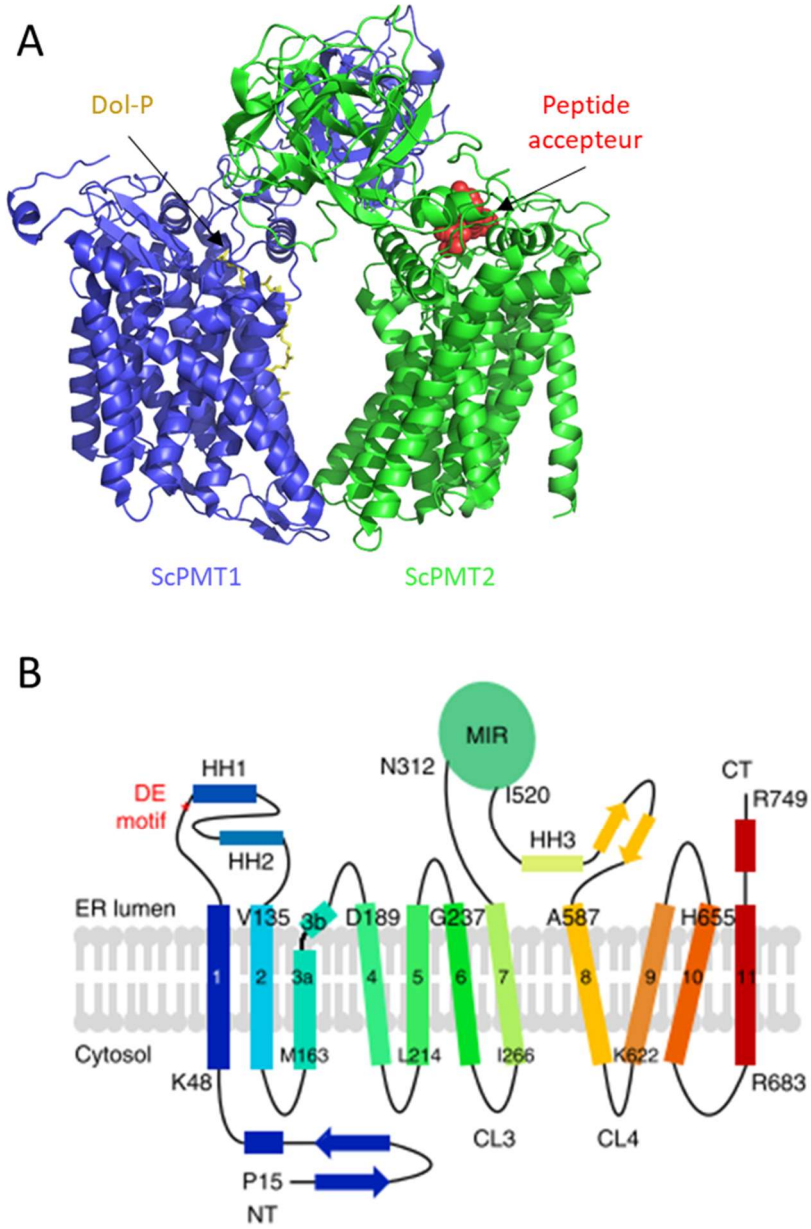


Figure 6 Structure des PMT de levure. A) Structure 3D de l'hétérodimère de *Sc*PMT1 (en bleu) et *Sc*PMT2 (en vert) en représentation en ruban. Le substrat donneur (Dol-P) est représenté en jaune dans la structure de ScPMT1 et le substrat accepteur est représenté en rouge dans ScPMT2. Structure PDB « 6P25 »[88] **B)** Topologie de ScPMT1 définissant 11 domaines transmembranaire. Le domaine MIR, les hélices horizontales (HH) et les hélices terminales sont représentées. Les premiers et derniers résidus de chaque domaine transmembranaire et du domaine MIR sont indiqués. Le motif DE essentiel est représenté par une astérisque rouge. Selon Bai et al., 2019[88].

a) Les PMT Eucaryotes

Les Protéine O-mannosyltransférases ont été identifiées pour la première fois en 1993 chez la levure[82] et décrites seulement six années plus tard chez l'homme[90]. Chez *Saccharomyces cerevisiae*, un total de sept gènes paralogues de PMT ont été identifiés numérotés de *Sc*PMT1 à *Sc*PMT7. L'activité *O*-mannosyltransférase a pu être démontrée pour *Sc*PMT1 à 4 et *Sc*PMT6[54,82,91–93]. La présence de plusieurs paralogues de PMT est une caractéristique commune aux eucaryotes. Par exemple, la levure *Candida albicans* possède 5 gènes de PMT (*Ca*PMT1–2 et *Ca*PMT4–6)[94] tout comme *S. cerevisiae*, alors que l'on en trouve généralement que 2 chez les animaux comme *Drosophila melanogaster* (*Dm*POMT1-2)[81], la souris (*Mm*POMT1-2)[95,96] ou l'homme (*h*POMT1-2)[97].

Les différentes isoformes de protéines *O*-mannosyltransférases peuvent être classifiés en trois sous-familles PMT1, PMT2 et PMT4 en fonction de leur proximité de séquences avec les protéines orthologues *Sc*PMT1, *Sc*PMT2 et *Sc*PMT4 de *S. cerevisiae*. Malgré de fortes analogies, des distinctions claires existent entre les sous-familles PMT1, PMT2 et PMT4, incluant des caractéristiques de séquences, mais aussi de spécificité de substrat protéique[53,98]. Aussi, il est à noter que les membres de la sous-famille de PMT1 sont spécifiques des *Fungi* et absents des autres règnes. De plus, les membres des sous-familles PMT1 et PMT2 sont fortement redondants chez les *Fungi* alors que la sous-famille PMT4 ne possède qu'un seul représentant par espèce de *Fungi*[53,99,100].

Du point de vue fonctionnel, il a été démontré que les PMT eucaryotes forment des dimères [88,101]. La dimérisation des PMT de levures suit une logique définie : les membres de la sous-famille PMT1 (*Sc*PMT1 et *Sc*PMT5) forment *in vivo* des hétérocomplexes avec des membres de la sous-famille PMT2 (*Sc*PMT2, *Sc*PMT3 et *Sc*PMT6), tandis que *Sc*PMT4 forme des homodimères. Ce type d'organisation semble conservé dans l'ensemble des *Fungi* exprimant les différents paralogues de ces sous-familles de PMT[102]. Comme précédemment indiqué, il n'existe pas d'orthologues de la sous-classe PMT1, chez les animaux. Ainsi, chez l'homme, les deux paralogues de PMT, *h*POMT1 et *h*POMT2 appartiennent aux sous-familles PMT4 et PMT2 respectivement. De manière intéressante, il a été démontré que l'expression des deux isoenzymes, *h*POMT1 et *h*POMT2, est nécessaire à l'activité *O*-mannosyltransférase sous-entendant une probable hétérodimérisation comme celle observée dans le cas des

complexes hétérodimériques de la levure *Sc*PMT1-*Sc*PMT2[97] récemment résolue par cryomicroscopie électronique (**Figure 6**)[88].

b) Les PMT bactériennes

Historiquement, les premières observations faisant état d'une probable activité de synthèse de mannoprotéines chez les bactéries ont été rapportées chez *Mycobacterium smegmatis* en 1975[103]. Cette observation princeps ne fut confirmée que plus de 25 ans plus tard, via l'utilisation de substrat donneur de mannose radiomarqué (GDP-Man*) et de peptides accepteurs dérivés d'une mannoprotéine mycobactérienne (nommée Apa)[75]. Le suivi *in vitro* du transfert du mannose radiomarqué (Man*) du substrat donneur aux peptides accepteurs a alors permis de détecter et quantifier une activité protéine *O*-mannosyltransférase dans des extraits acellulaires de *M. smegmatis*. De plus cette activité glycosyltransférase absente de la fraction cytosolique et principalement associée aux fractions membranaires de l'extrait acellulaire a confirmé la localisation membranaire probable de l'enzyme responsable de ce transfert.

Ce n'est toutefois qu'en 2005 qu'un gène orthologue des PMT de levure a pu être identifié chez *Mycobacterium tuberculosis* par une approche bio-informatique[55]. Ce gène Rv1002c code pour une protéine de 522 acides aminés présentant environ 20% d'identité en acides aminés et un profil hydropathique similaire aux PMT de *S. cerevisiae*. L'implication de ce gène dans la *O*-mannosylation des protéines de *Mycobacterium tuberculosis* a été étudiée[55]. Il a été montré que la surexpression du gène Rv1002c chez *Mycobacterium smegmatis* se traduit *in vitro* par une augmentation de l'activité mannosyltransférase d'extrait membranaire, mesurée comme précédemment par le taux de transfert de mannose radiomarqué sur un peptide prototypique de la protéine Apa de *M. tuberculosis*[55,75]. De plus il a pu être montré que l'expression hétérologue chez *M. smegmatis* du gène rv1002c muté pour le motif DE essentiel des PMT, s'accompagne d'une abolition totale de l'augmentation du taux de mannosylation des protéines en le ramenant au niveau basal du contrôle négatif. Enfin, cette étude a permis de mettre en évidence que le peptide signal d'adressage au complexe de translocation Sec est nécessaire à la glycosylation des protéines bactériennes, attestant d'un couplage des processus de O-mannosylation post-traductionnelle et de translocation membranaire des mannoprotéines[55]. Ces observations ont conduit les auteurs à définir que la protéine PMTub, codée par le gène rv1002c, est une

protéine membranaire associée au translocon Sec et qui initie la *O*-mannosylation des protéines par un processus analogue à celui décrit pour les PMT de levure[55].

Chez la mycobactérie non pathogène à croissance rapide, *Mycobacterium smegmatis*, le gène de la PMT a été identifié et caractérisé comme étant Msmeg_5447[52]. La délétion de ce gène révéla sa nature non essentielle dans cette bactérie. La souche de *M. smegmatis* délétée pour le gène Msmeg_5447 ne montre pas d'augmentation de sensibilité à la ciprofloxacine suggérant que l'absence de mannoprotéines n'a pas d'impact sur le métabolisme et la perméabilité de la paroi à cet antibiotique. Cependant, cette souche s'est révélée être plus sensible au SDS suggérant des altérations probables de la composition ou de la structuration de la paroi, sans pour autant avoir d'effet visible sur la bactérie en microscopie électronique[52]. Il aussi été montré que la perte d'expression de la PMT chez *M. smegmatis* n'influençait pas le niveau de phagocytose par des macrophages THP-1, mais s'accompagnait d'une diminution de la survie bactérienne dans des conditions mimant l'environnement intraphagolysosomal (culture en milieu acide (pH 5.0) ou en présence de lysozymes)[104]. L'ensemble de ces observations semble indiquer que, bien que non essentielle, la PMT mycobactérienne pourrait avoir un rôle dans la survie du pathogène dans les conditions acides du phagolysosome lors de l'infection.

Chez *M. tuberculosis*, le gène rv1022c codant pour PMTub avait été précédemment proposé comme essentiel[105]. Toutefois, des résultats du laboratoire ont permis de démontrer que la souche délétée du gène rv1002c est pourtant viable même si sa croissance *in vitro* en milieu solide était très fortement impactée[52]. Ce déficit de croissance ne semble pas pouvoir s'expliquer par une altération de la paroi mycobactérienne puisque la souche dépourvue de PMT ne montre pas de sensibilité accrue à la ciprofloxacine ou au SDS. De façon plus intéressante, cette souche se révèle surtout présenter un défaut de multiplication dans les macrophages alvéolaires et une virulence réduite chez les souris immunodéprimées[52]. À ce jour, les bases moléculaires de ce phénotype original restent encore incomprises. Des travaux récents ont pu montrer que les souches mutantes de *M. tuberculosis* et de *M. smegmatis* dépourvues de PMT présentent une augmentation de la sécrétion de lipoarabinomannane, un lipoglycane immunomodulateur de la réponse immunitaire de l'hôte[50]. La sécrétion accrue de lipoarabinomannane, en partie liée à LprG, une mannolipoprotéine impliquée dans la translocation des lypoglycanes à la surface cellulaire, conduit à une augmentation de la

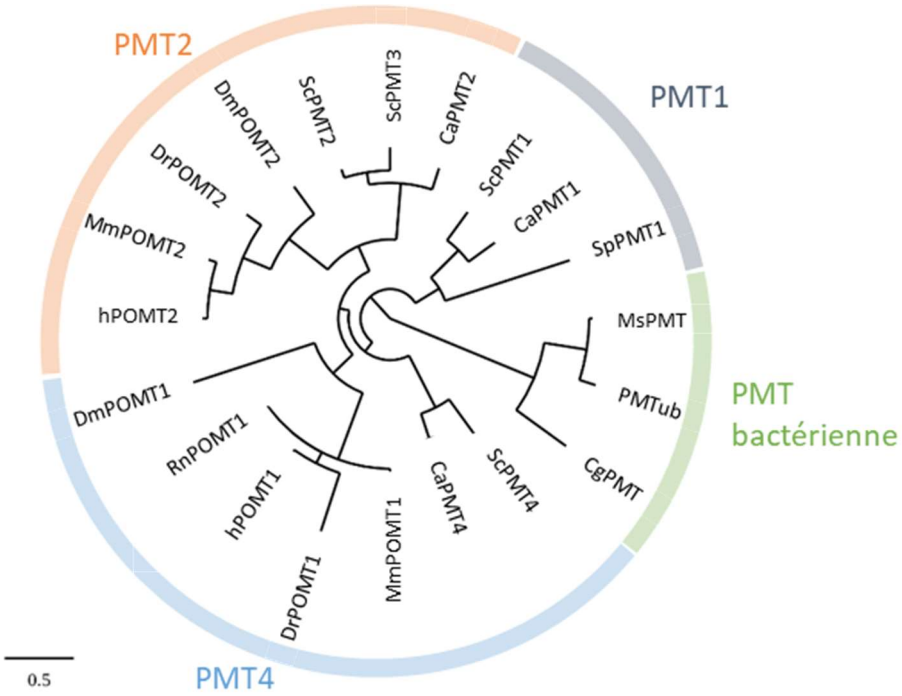


Figure 7. Les différentes familles de PMT. Arbre phylogénétique de PMT définissant les différentes sous-familles de PMT. Les séquences choisies correspondent à des PMT démontrées fonctionnelles et représentatives de la diversité des organismes exprimant des PMT : PMT d'organismes unicellulaires eucaryotes (*S. cerevisiae*[91,101], *S. pombe*[106] et *C. albicans*[94]), de procaryotes (*M. tuberculosis*[55], *M. smegmatis*[52] et *C. glutamicum*[59]), de mammifères (homme[97], *M. musculus*[95,96] et *R. norvegicus*[107]), d'insectes (*D. melanogaster*[81]) et de poissons (*D. rerio*[108]). Arbre construit après alignement des séquences entière de PMT par Muscle[109] grâce au logiciel phylogeny.fr[110]

réponse inflammatoire dépendante de TLR-2 pouvant contribuer à expliquer le phénotype observé.

Chez *Mycobacterium abscessus,* une mycobactérie pathogène à croissance rapide, la suppression du gène MAB_1122c, codant pour une PMT putative, se traduit également par une perte de la mannosylation des protéines[111]. Cette souche dépourvue de PMT ne présente pas de défaut de croissance *in vitro*, comme *M. smegmatis*, toutefois elle présente *in cellulo* un phénotype analogue au mutant rv1002c de *M. tuberculosis*. La perméabilité de la paroi bactérienne semblerait affectée par la perte de la O-mannosylation des protéines chez *M. abscessus* comme le montrent des essais d'absorption de bromure d'éthidium. Cette propriété pourrait expliquer la plus grande sensibilité aux antibiotiques de la souche dépourvue de PMT comparée à la souche sauvage de *M. abscessus*[111].

Enfin des PMT ont également été mises en évidence expérimentalement dans d'autres actinobactéries telles que *Corynebacterium glutamicum*[59] et *Streptomyces coelicolor*[112]*.* Ainsi la délétion chez *C. glutamicum* ou *S. coelicolor* des gènes Cg1004 ou CSO3154 orthologues du gène Rv1002c de *M. tuberculosis,* provoque une perte complète de la mannosylation des protéines sécrétées et en particulier du facteur de ressuscitation Rpf2 de *C. glutamicum*[59] ou de la protéine de liaison au phosphate PstS de *S. coelicolor*[112]*.*

Finalement, à ce jour, dans le monde bactérien seuls les actinomycètes semblent avoir été démontrés comme capables d'accomplir la O-mannosylation de protéines, via des Protéine *O*-mannosyltransférases fortement conservées présentant des similarités structurales indéniables avec les PMT eucaryotes.

Toutefois certaines divergences importantes existent entre les PMT eucaryotes et bactériennes. Ainsi, les PMT bactériennes n'appartiennent à aucune des trois sous-familles PMT1, PMT2 et PMT4 classifiant les PMT eucaryotes et constituent un groupe distinct sur un arbre phylogénétique (**Figure 7**).

Une autre différence majeure réside dans le fait que les bactéries n'expriment qu'un seul gène de PMT[55,59,111,112] contrairement aux eucaryotes qui peuvent exprimer jusqu'à 7 gènes paralogues de PMT comme *S. cerevisiae* par exemple. D'un point de vue fonctionnel, l'unicité des gènes de PMT facilite l'étude de l'impact physiologique de leur délétion chez les

bactéries, mais n'a pas permis à ce jour de mettre en évidence la formation des complexes dimériques comme observés dans le cas des PMT eucaryotes.

Enfin, du point de vue structural, les PMT bactériennes se distinguent foncièrement par l'absence de domaines MIR caractéristiques de la grande majorité des PMT eucaryotes.

Ces différents points de divergences phylogénique, génomique (absence de duplication des gènes) et structurale (absence de motif MIR) ont ainsi conduit certains auteurs à proposer que les PMT bactériennes pourraient constituer une forme ancestrale des PMT eucaryotes[99].

c) Suivi de l'activité des PMT

Comme souligné plus haut, à ce jour peu de ces Protéines *O*-mannosyltransférases ont pu être expérimentalement caractérisées d'un point de vue fonctionnel et ceci malgré la forte conservation de ces protéines dans le vivant et leurs rôles physiologiques importants voire essentiels pour les organismes qui les expriment.

Une des raisons à cela est leur difficulté d'étude liée à leurs propriétés intrinsèques de protéines membranaires intégrales, qui gênent leur purification et la caractérisation de leurs activités enzymatiques.

La plupart des méthodes mises en place pour la détection de l'activité des PMT sont basées sur la démonstration de la présence de substitutions glycosyl, voire de mannose, sur des peptides ou des protéines substrats accepteurs de PMT. Pour cela, 2 grands types d'approches ont été proposées dans la littérature, plus ou moins quantitatives, voire uniquement qualitatives, et qui sont basées soit sur l'utilisation de composés radiomarqués, soit sur la détection de variations de masse, associée à la glycosylation de protéines cibles. Historiquement, la première stratégie développée pour la mise en évidence d'activité Protéine *O*-mannosyltransférase chez les bactéries[55,75,103] ou chez la levure[93], a été le suivi du transfert de mannose radiomarqué (Man*) à partir du substrat donneur ou de son précurseur GDP-mannose vers des peptides ou protéines acceptrices. Ce test acellulaire *in vitro* consiste à incuber des dérivés polyprenol-phospho-mannose (pPr-P-Man* ou Dol-P-Man*) ou du GDP-Man* en présence d'extraits membranaires de bactéries, levures ou autres organismes censés exprimer des PMT, et éventuellement la machinerie enzymatique permettant la biosynthèse

du pPr-Man* à partir du GDP-Man*. L'activité mannosyltransférase peut alors être révélée par la détection du transfert de la radioactivité du Man* du substrat donneur vers des protéines endogènes[103] ou un peptide cible exogène[55,75,93]. Initialement développée pour mettre en évidence une activité protéine mannosyltransférase *chez M. Smegmatis*[103], ce test a permis de caractériser les PMT de levure et de déterminer des activités relatives de transfert de mannose sur différents peptides accepteurs[93]. Appliqué à l'étude cinétique de la PMT de *M. Smegmatis*, cette approche a permis de quantifier son activité enzymatique et de déterminer un Km 0,56 mM pour le peptide accepteur DPEPAPPVPTTAASPP correspondant au peptide amino terminal de la protéine Apa (Rv1860) de M. tuberculosis[75]. Toutefois cette approche présente certaines limitations intrinsèques principalement liées à la reproductibilité de la préparation d'extraits membranaires acellulaires notamment en termes de quantité de PMT présente, mais également de substrat donneur de mannose radiomarqué (dans le cas de l'utilisation du précurseur GDP-Man*). Ces paramètres variables, difficilement contrôlables et quantifiables d'une préparation à l'autre constituent une faiblesse inhérente à ces approches acellulaires en termes de robustesse et reproductibilité des résultats.

Une autre limitation inhérente de cette stratégie réside dans les contraintes expérimentales associées aux mesures de sécurité imposées pour l'utilisation de radionucléides.

Parallèlement à ces approches, les stratégies basées sur la détection des variations de masse associées à la glycosylation de protéines cibles ont été largement utilisées et présentent l'avantage de pouvoir être applicables au suivi de l'activité de mannosylation des protéines *in cellulo* sur cellules en culture. Par exemple, la détection de la glycosylation des protéines peut être réalisée par analyse SDS-PAGE: la décoration saccharidique entraine une augmentation de la masse moléculaire de la protéine glycosylée comparée à sa version non glycosylée, qui se traduit par un retard de migration sur gel SDS-PAGE. Dans le cas de protéines cibles purifiées ou traçables par des anticorps (ou toutes autres approches permettant de l'identifier au sein d'un mélange protéique), ce décalage de migration sur gel permet ainsi rapidement de détecter une activité de *O*-glycosylation[53] . Dans le cas de mannoprotéines, cette approche peut être complétée et renforcée par l'utilisation de lectines spécifiques telle que la Concanavaline A permettant de détecter la présence de mannose terminaux et de confirmer ainsi la mannosylation de la protéine[113].

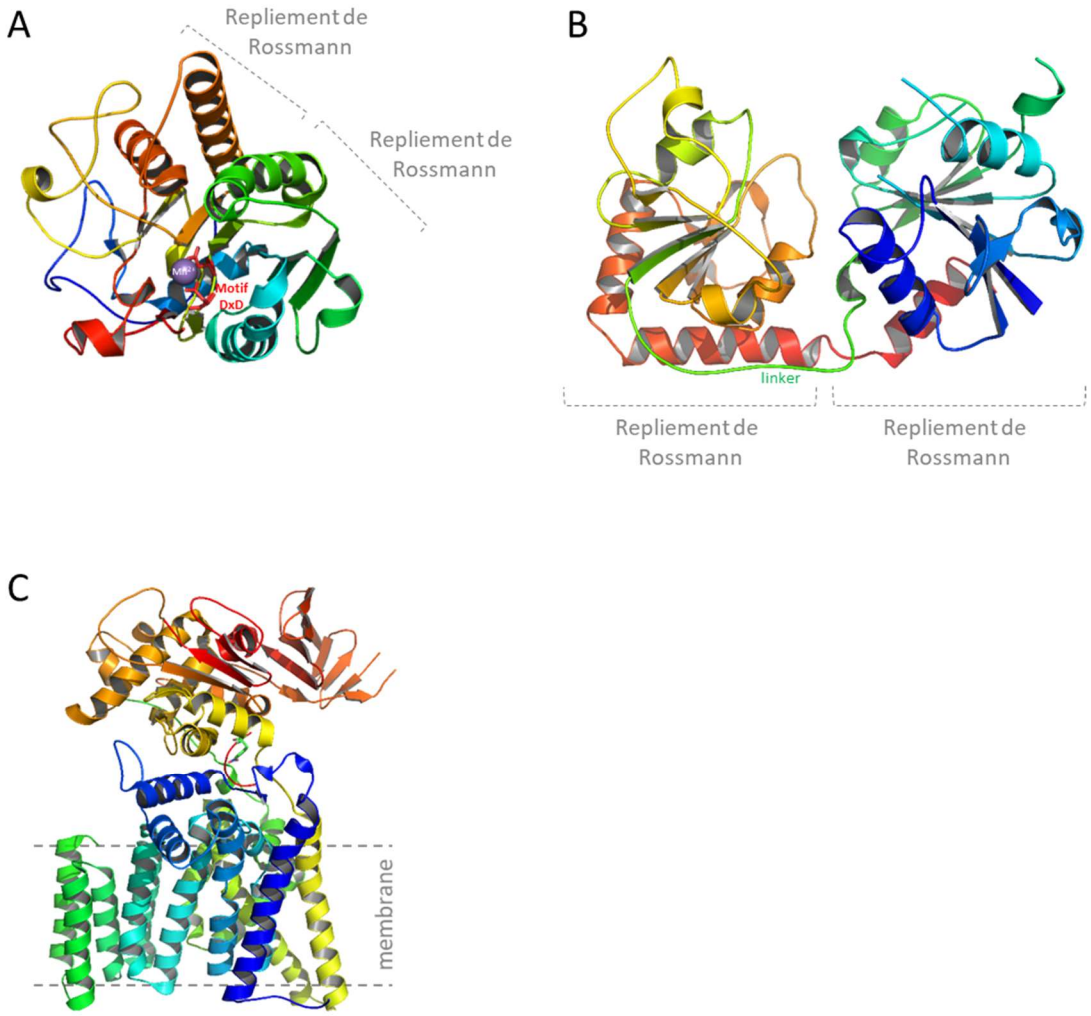


Figure 8. Les superfamilles de glycosyltransférases. A) Structure de GT-A résolue (PDB 3IOJ), la transférase du groupe sanguin ABO de l'homme[114], décrivant deux repliements de Rossmann asymétrique entrecroisés. L'ion divalent est représenté par une boule violette et les acides aspartiques du motif DxD sont représentés en bâtons rouges. **B)** Structure de GT-B résolue (PDB 2BGT), la DNA beta-glucosyltransferase du bactériophage T4[115], décrivant deux repliements de Rossmann symétriques séparés par un linker. **C)** Première structure de GT-C résolue (PDB 2BGT), PglB de *Campylobacter lari*[116], montrant un ancrage membranaire polytopique.

La spectrométrie de masse est aussi un outil régulièrement utilisé pour ce genre d'analyse et a notamment permis de confirmer le gène Rv1002c comme codant pour la PMT de *M. tuberculosis* en détectant la présence ou non d'hexose sur un peptide issu de la protéine Apa (Rv1860) [55]. De plus les approches de spectrométrie de masse de type « bottom-up » utilisées en protéomique, permettant la détection sélective et l'identification de glycopeptides, constituent une alternative performante et efficace pour détecter à large échelle des modifications post-traductionnelles associées à une activité Protéine O-mannosyltransférase[37,38,117]. Parallèlement l'analyse par spectrométrie de masse dite « Top-Down » de glycoprotéines entières purifiées permet d'accéder à un niveau d'information supplémentaire jamais atteint auparavant et représente un outil prometteur pour l'analyse du profil de glycosylation et l'étude fine du déterminisme des sites de O-mannosylation préférentielle des PMT comme cela a été décrit dans le cas de la mannosylation de la protéine LpqH de *M. tuberculosis*[47].

Toutefois il apparait clairement que ces approches performantes nécessitent des mises en œuvre complexes et contraignantes peu adaptées à des analyses quantitatives de routines.

Enfin, on se doit de noter le développement original d'un test dit « phénotypique » pour le suivi de l'activité de la PMT1 de *Candida albicans.* En se basant sur l'observation selon laquelle une souche de *C. albicans* délétée pour *Ca*PMT1 était sensible aux antibiotiques aminoglycosides, le groupe de Orchard *et al.* a conçu un test d'activité pour le criblage d'inhibiteurs ciblant cette enzyme[118], conduisant à l'identification d'inhibiteurs spécifiques de type rhodanine-3-acide acétique.

Toutefois ces approches phénotypiques restent difficiles à développer du fait de l'absence, pour la plupart des organismes, de phénotype spécifique exploitable après interruption de l'activité protéine O-mannosyltransférase. De plus, comme pour les méthodes basées sur la détection de mannose sur les protéines substrats de PMT, ces approches ne permettent pas d'obtenir une mesure quantitative de l'activité mannosyltransférase.

En conclusion, il apparait de cet inventaire abrégé, mais exhaustif, qu'il n'existe pas de méthode « gold standard » pour suivre l'activité de glycosylation de ce type de glycosyltransférase membranaire intégrale. Ce constat nous a donc logiquement contraints à développer au cours de ma thèse un nouveau bio-essai robuste, reproductible et compatible

Tableau 2. Les différentes familles de glycosyltransférase de type-C selon la base de donnée CAZy [119,120].

Famille GT	Activités	Nombre gènes identifiés
GT-22	Dol-P-Man: Man6GlcNAc2-PP-Dol α-1,2-mannosyltransferase (EC 2.4.1.259); Dol-P-Man: Man8GlcNAc2-PP-Dol α-1,2-mannosyltransferase (EC 2.4.1.261); Dol-P-Man: Man2-GlcNAc-phosphatidylinositol α-1,2-mannosyltransferase (EC 2.4.1.-); Dol-P-Man: Man3-GlcNAc-phosphatidylinositol α-1,2-mannosyltransferase (EC 2.4.1.-)	1504
GT-39	Dol-P-Man: protein α-mannosyltransferase (EC 2.4.1.109)	3804
GT-50	Dol-P-Man α-1,4-mannosyltransferase (EC 2.4.1.-)	368
GT-53	DecapreP-Araf: α-(1,3)-arabinosyltransferase (EC 2.4.2.-); DecapreP-Araf: Lipoarabinomannan α-(1,5)-arabinosyltransferase (EC 2.4.2.-)	2174
GT-57	Dol-P-Glc: α-1,3-glucosyltransferase (EC 2.4.1.-)	635
GT-58	Dol-P-Man: Man5GlcNAc2-PP-Dol α-1,3-mannosyltransferase (EC 2.4.1.258)	539
GT-59	Dol-P-Glc: Glc2Man9GlcNAc2-PP-Dol α-1,2-glucosyltransferase (EC 2.4.1.256)	308
GT-66	dolichyl-diphosphooligosaccharide-protein glycotransferase (EC 2.4.99.18); undecaprenyl-diphosphooligosaccharide-protein glycotransferase (EC 2.4.99.19)	1858
GT-83	undecaprenyl phosphate-α-L-Ara4N: 4-amino-4-deoxy-β-L-arabinosyltransferase (EC 2.4.2.43); dodecaprenyl phosphate-β-galacturonic acid: lipopolysaccharide core α-galacturonosyl transferase (EC 2.4.1.-)	10706
GT-85	β-D-arabinofuranosyl monophosphoryldecaprenol: galactan α-D-arabinofuranosyltransferase (EC 2.4.2.-)	887
GT-87	polyprenol-P-Man: α-1,2-mannosyltransferase (EC 2.4.1.-)	3524
GT-105	Dol-P-β-D-Man: protein α-D-mannosyltransferase (EC 2.4.1.109)	470

avec des analyses bas-moyen débit pour suivre et quantifier de façon multiplexée l'activité *in cellulo* de la protéine *O*-mannosyltransférase de *M. tuberculosis*.

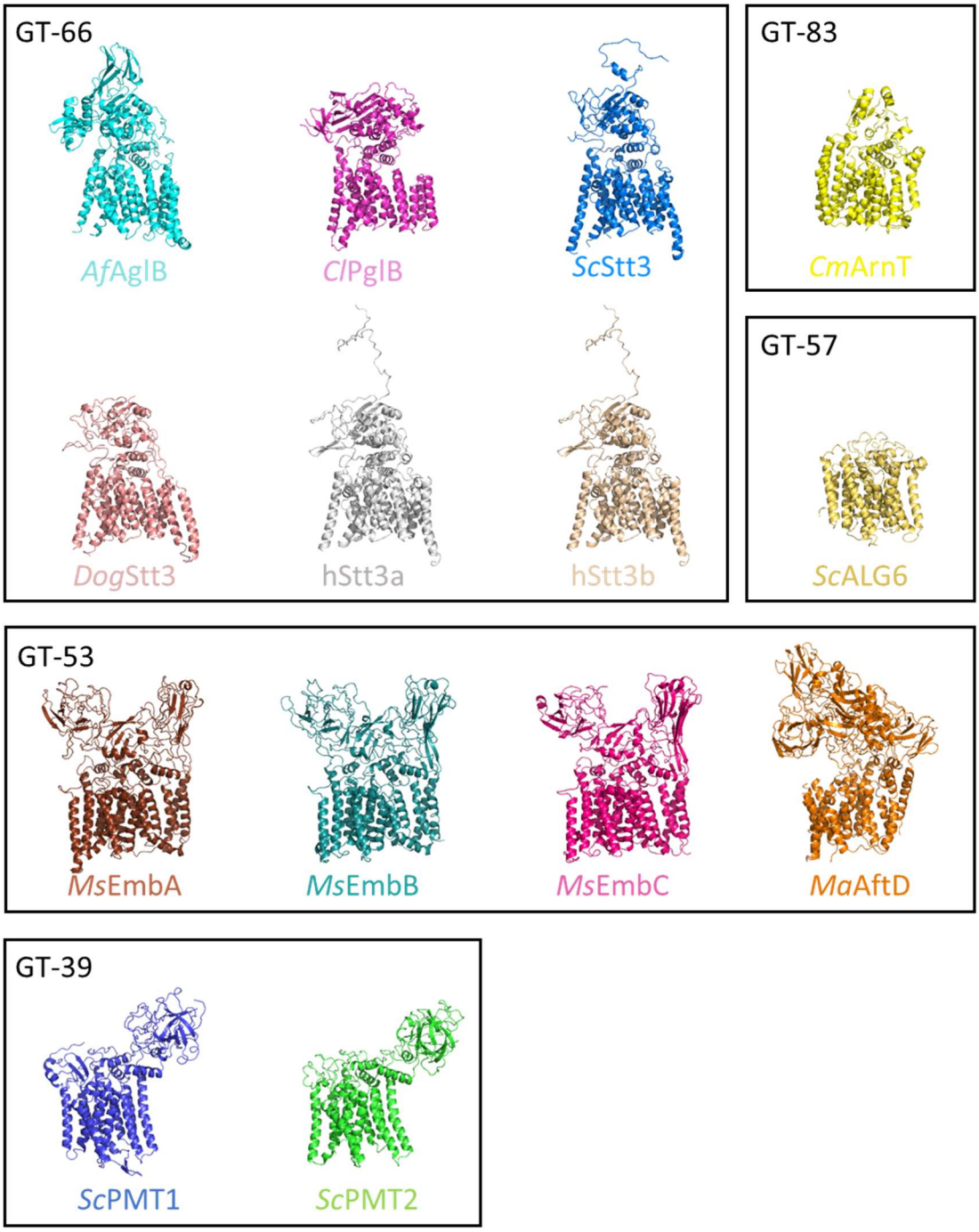


Figure 9. Les GT-Cs de structure résolues. Représentations en ruban des glycosyltransférases de type-C de structures connues classé selon leur famille CAZy : GT-39 (*Sc*PMT1 et 2[88]), GT-53 (*Ms*EmbA, B et C[121], et *Ma*AftD[122]), GT-57 (*Sc*ALG6[123]), GT-66 (*Af*AglB[124], *Cl*PglB[116], *Sc*Stt3[125], *Dog*Stt3[126], *h*Stt3a et b[127]) et GT-83 (*Cm*ArnT[128]). *Sc : S. cerevisiae ; Ms : M. smegmatis ; Ma : M. abscessus ; Af : A. fulgidus ; Cl : C. lari ; h : human ; Cm : C. metalidurans.*

B. Propriétés structurales des glycosyltransférases

1. Classification des glycosyltransférases

Les PMT font partie de la vaste famille des glycosyltransférases (GT) impliquées dans la synthèse de toute une variété de composés saccharidiques, allant de l'oligosaccharide à la protéine glycosylée. Ces enzymes catalysent le transfert de sucres depuis un substrat donneur activé vers un substrat accepteur. Cette réaction résulte en la formation d'une liaison glycosidique. Il existe une grande diversité de substrats donneurs et accepteurs reconnus par les GT à l'origine de la grande variété des membres de cette superfamille d'enzyme. En effet, plusieurs milliers de GT ont pu être identifiés au travers de tous les règnes du vivant dont un nombre important sont impliquées dans des processus essentiels de modifications post-traductionnelles des protéines par glycosylation.

La nomenclature EC, établie par l'Union internationale de biochimie et de biologie moléculaire (en anglais, International Union for biochemistry and molecular biology ou IUBMB), ne permet pas de rendre compte de la diversité de la superfamille des glycosyltransférases. Une classification basée sur des alignements de séquence en acides aminés a permis de définir 111 familles de GT nommées de GT1 à GT111 (CAZy database[119,120]). Toutefois, de manière intéressante, la résolution des structures de GT met en évidence trois grandes classes de repliements structuraux caractéristiques, définissant alors trois grandes superfamilles structurales de GT : GT-A, GT-B et GT-C (**Figure 8**). Il est cependant à noter que plusieurs glycosyltransférases présentent des repliements ne correspondant à aucune de ces 3 superfamilles, comme cela a été récemment mis en évidence pour la glycosyltransférase bactérienne PBP2 de *Staphylococcus aureus*[129].

2. Les glycosyltransférases de type A et B

Les glycosyltransférases de type A et B sont aussi appelées enzyme de Leloir, du nom de Louis Leloir, prix Nobel de chimie de 1970 pour ses travaux sur la biosynthèse des hydrates de carbone. Ces enzymes ont comme point commun d'utiliser un dérivé nucléotidique soluble comme substrat donneur de sucre. Le sucre est transféré du donneur nucléotide-sucre vers le substrat accepteur selon un mécanisme par inversion ou par rétention de l'anomérie. Les enzymes de Leloir possèdent plusieurs autres points de convergence : ce sont des enzymes avec un cœur globulaire soluble arborant un repliement dit de Rossmann

Tableau 3. Les glycosyltransférase de type-C de structure 3D résolues.

Famille GT	Nom de l'enzyme	Espèce	PDB	Résolution	Méthode	Référence
GT-39	ScPMT1	*Saccharomyces Cerevisiae*	6P25	3.20 Å	Cryo-ME	Bai *et* a, 2019[88]
	ScPMT2	*Saccharomyces Cerevisiae*	6P25	3.20 Å	Cryo-ME	
GT-53	MsEmbA	*Mycobacterium smegmatis*	7BVG	3.10 Å	Cryo-ME	Zhang *et al*, 2020 [121]
	MsEmbB	*Mycobacterium smegmatis*	7BVG	3.10 Å	Cryo-ME	
	MsEmbC	*Mycobacterium smegmatis*	7BVE	2.81 Å	Cryo-ME	
	MaAftD	*Mycobacterium abscessus*	6W98	2.90 Å	Cryo-ME	Tan *et al*, 2020[122]
GT-57	ScALG6	*Saccharomyces Cerevisiae*	6SNI	3.00 Å	Cryo-ME	Bloch *et al*, 2020[123]
GT-66	AfAglB	*Archaeoglobus fulgidus*	3WAK	3.41 Å	Rayon X	Matsumoto *et al*, 2013[124]
	ClPglB	*Campylobacter lari*	3RCE	3.40 Å	Rayon X	Lizak *et al*, 2011[116]
	ScStt3	*Saccharomyces Cerevisiae*	6EZN	3.30 Å	Cryo-ME	Wild *et al*, 2018[125]
	DogStt3	*Canis lupus familiaris*	6FTI	4.20 Å	Cryo-ME	Braunger *et al*, 2018[126]
	hStt3a	*Homo sapiens*	6S7O	3.50 Å	Cryo-ME	Ramirez *et al*, 2019[127]
	hStt3b	*Homo sapiens*	6S7T	3.50 Å	Cryo-ME	
GT-83	CmArnT	*Cupriavidus metallidurans*	5EZM	2.70 Å	Rayon X	Petrou *et al*, 2016[128]

(**Figure 8A et B**). Ces repliements de Rossmann comportent plusieurs hélices α entourées de feuillets β.

Les GT-A et les GT-B se distinguent par l'agencement différent de ces domaines. Les GT-A possèdent généralement deux repliements de Rossmann asymétriques qui s'entrecroisent pour former un seul domaine globulaire (**Figure 8A**). Au cœur de ce domaine se trouve le site catalytique identifiable par la présence d'un motif DxD (où D est un acide aspartique et x est n'importe quel acide aminé) et d'un ion divalent impliqué dans la catalyse. Les GT-B sont composées de deux repliements de Rossmann identiques séparés par un linker flexible formant alors deux domaines globulaires distincts (**Figure 8B**). Contrairement aux GT-A, les GT-B ne présentent pas de motif caractéristique au niveau de leur site actif et leur activité est indépendante de la présence d'ion divalent.

3. Les glycosyltransférases de type C

a) Localisation cellulaire

La troisième grande famille structurale des glycosyltransférases, les GT-C, se distingue des GT-A et des GT-B essentiellement par deux caractéristiques : la nature de leur substrat donneur et leur conformation structurale. En effet contrairement aux glycosyltransférases de type Leloir (GT-A et GT-B), les GT-C sont des enzymes membranaires polytopiques qui n'utilisent pas de nucléotide-sucre, mais un lipidique activé de type polyisoprenoyl–phospho-sucre comme substrat donneur de sucre (**Figure 8C**).

Chez les eucaryotes, ces glycosyltransférases de type-C utilisent des substrats liés au dolichol-phosphate (Dol-P) ou au dolichol-diphosphate (Dol-PP) et se localisent aux membranes de la cellule telles que la membrane plasmique ou la membrane de compartiments intra cellulaire comme le réticulum endoplasmique ou l'appareil de Golgi. De façon similaire chez les procaryotes, les GT-C bactériennes utilisent un polyprénol-phosphate comme substrat donneur et sont retrouvées à la membrane plasmique. Au-delà de la similarité de substrat, les GT-C fonctionnent selon un mécanisme catalytique conduisant à une inversion de l'anomérie du sucre transféré. Cependant, malgré ces points de convergence, les GT-C présentent une grande diversité de séquences et se répartissent donc dans de nombreuses familles CAZy (**Tableau 2**).

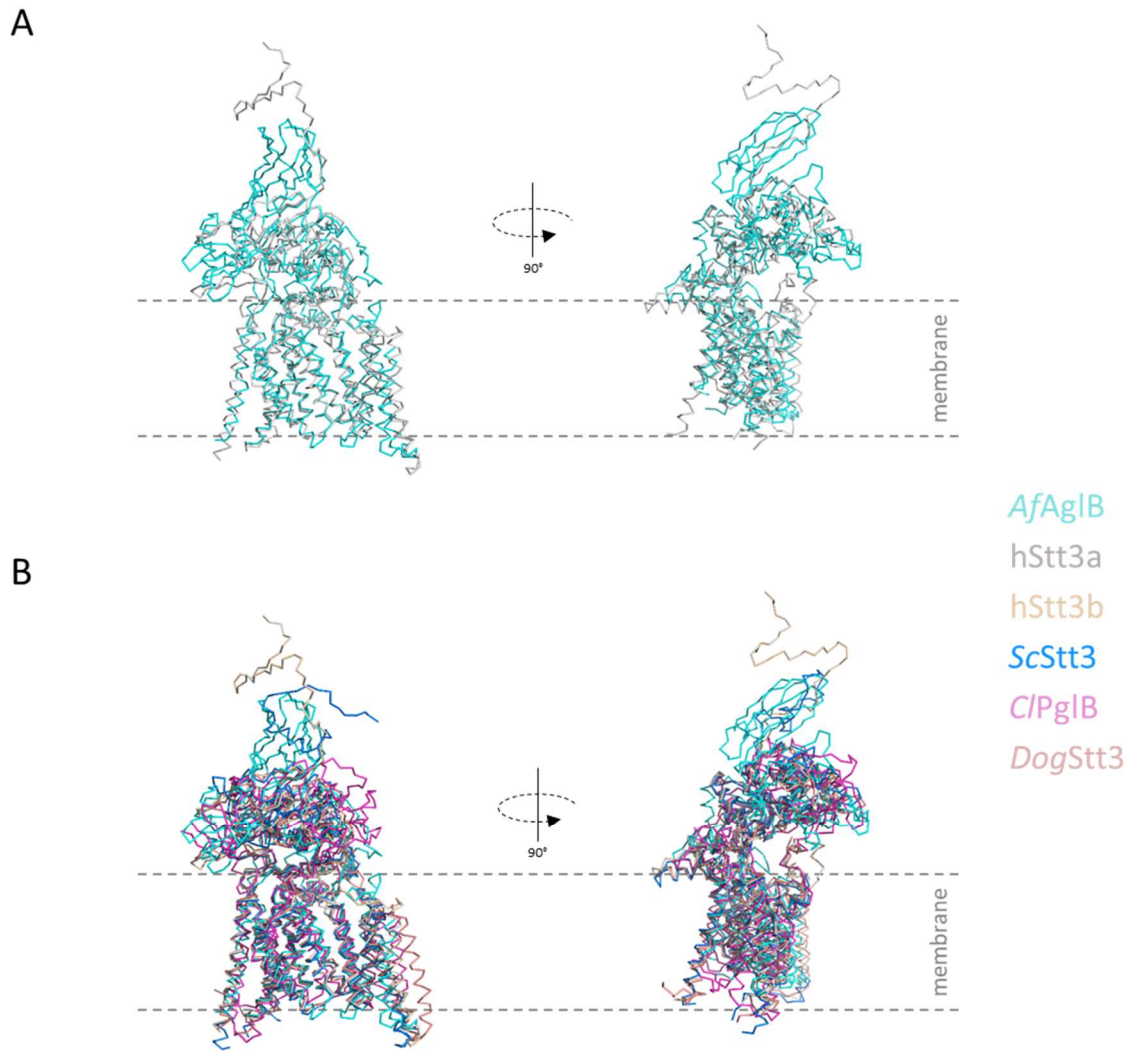


Figure 10. Superposition des structures de GT-66. A) Superposition de *Af*AglB[124] et *h*Stt3a[127] en représentation en fil de fer vue de face et de côté. La coloration est selon la légende de gauche et suit les couleurs utilisées Figure 9. **B)** Superposition de toutes les enzymes de la famille CAZy GT-66 de structure connue[116,124–127] en représentation en fil de fer vue de face et de côté. La coloration est selon la légende de gauche et suit les couleurs utilisées Figure 9.

Ainsi, les GT-C sont impliquées dans divers mécanismes comme la *N*-glycosylation, la *O*-glycosylation, et *C*-mannosylation des protéines ou la biosynthèse de glycolipides ou polysaccharides par exemple. Parmi ces glycosyltransférases de type C, figurent les Protéine *O*-mannosyltransférase PMT et TMTC qui correspondent respectivement aux familles GT-39 et GT-105 de la classification CAZy.

b) Topologie membranaire

Une particularité importante des GT-C face aux GT-A et GT-B est donc leur intégration membranaire polytopique. Les premières caractérisations expérimentales de la topologie membranaire des GT-C ont été réalisées par étiquetage protéique de la Protéine O-mannosyltransférase *Sc*PMT1 de *S. cerevisiae*, et de la sous-unité catalytique *Sc*Stt3 de l'oligosaccharyltransférase de la levure [130,131].

Dans l'étude princeps de la topologie de *Sc*PMT1 de *S.cerevisiae*, les auteurs se sont basés sur le profil hydropathique de *Sc*PMT1 pour construire plusieurs protéines recombinantes rapportrices de la topologie en fusionnants divers domaines de *Sc*PMT1 avec l'histidinol déshydrogénase His4C, protéine active dans le cytoplasme[131]. Sur milieu minimum supplémenté en histidinol, l'expression des protéines de fusion *Sc*PMT1-His4C dans le cytoplasme d'une souche de levure auxotrophe permet la métabolisation de l'histidinol en histidine et la survie du microorganisme. En revanche la translocation de *Sc*PMT1-His4C dans la lumière du réticulum endoplasmique ne permettant pas l'utilisation de l'histidinol supplémenté, conduit à un phénotype létal. En construisant diverses protéines de fusion où His4C a été localisée aux extrémités de *Sc*PMT1 ou dans différentes boucles solubles prédites par le profil hydropathique, les auteurs ont ainsi pu préciser la topologie de *Sc*PMT1. Ils ont donc pu définir expérimentalement une topologie à 7 domaines transmembranaires, divergente de la topologie prédite par l'analyse du profil hydropathique suggérant 11 domaines transmembranaires[131].

Une approche expérimentale similaire utilisant des protéines de fusion avec des insertions de *Suc2/Hic4s* dans des domaines stratégiques a permis de définir 11 domaines transmembranaires pour la protéine *Sc*Stt3 du complexe oligosaccharyltransférase de *S. cerevisiae*[130]. Il est intéressant de noter que dans ces deux cas les résultats ont permis de conclure à des topologies de type trans avec des domaines N-terminaux émergeant coté

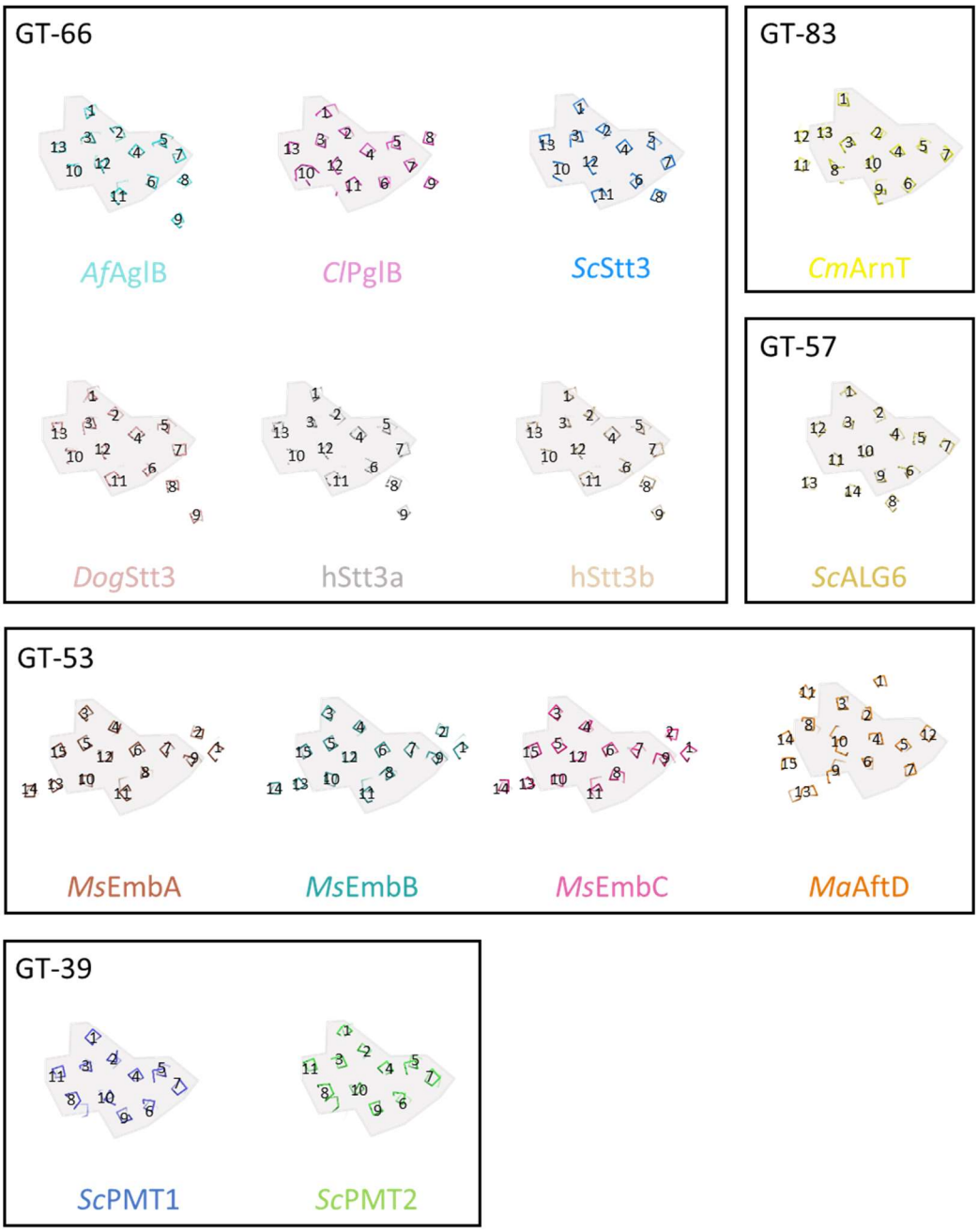


Figure 11 Comparaison des topologies des GT-C de structure résolues. Représentation des structures selon une coupe au niveau de la membrane vue de dessus. Le module GT-C conservé de 11 domaines transmembranaires est grisé. Les hélices transmembranaires sont numérotées. Pour *Sc*Stt3 la structure du 9^{e} domaine transmembranaire n'est pas résolu, il n'est donc pas représenté ici. *Ma*AftD ne semble pas suivre strictement le socle commun de GT-C.

cytoplasmique de la membrane du réticulum endoplasmique et des extrémités carboxy-terminales pointant dans la lumière du réticulum.

Les approches prédictives *in silico* et les études expérimentales se basant sur des étiquettes protéiques ont donc permis de progresser dans la compréhension des relations structure/fonction de ces différentes enzymes, mais seule la résolution de la structure 3D de ces enzymes a vraiment permis d'avancer et d'avoir une vision précise de l'organisation et de la topologie de ces enzymes membranaires. Toutefois, du fait de la large taille et de la forte hydrophobicité de ces protéines polytopiques intégrales, peu de structures de GT-C ont pu être résolues à ce jour. Actuellement seules 14 structures de GT-C ont été publiées[88,116,121–125,127,128] (**Figure 9**) dont la majorité ont été établies grâce aux avancées récentes de la cryomicroscopie électronique (Cryo-ME) de haute résolution à l'exception des structures de *Cl*PglB[116], *Af*AglB[124] et *Cm*ArnT[128] qui ont été résolues par diffraction aux rayons X après cristallographie (**Tableau 3**).

La première structure d'une GT-C publiée en 2011 a été celle de la protéine oligosaccharyltransférase *Cl*PglB de la bactérie *Campylobacter lari*[116]. Deux ans plus tard, la structure de son équivalent chez l'archaebactérie *Archaeoglobus flugidus*, *Af*AglB, est aussi résolue[124]. Ces deux enzymes équivalentes de Stt3 chez les eucaryotes appartiennent à la famille GT-66 et présentent des structures faisant apparaitre des topologies à 13 domaines transmembranaires. Depuis, les structures de *Sc*Stt3 de *S. cerevisiae* et de *h*Stt3A et *h*Stt3B de l'homme ont été décrites[125,127] qui confirment le repliement commun à 13 domaines transmembranaires pour tous les membres de la famille GT-66 caractérisés à ce jour. De manière intéressante, comme dans le cas de ScPMT1 discuté plus haut, il est à noter que les topologies expérimentalement définies par la résolution des structures de ces oligosaccharyl-transférase diffèrent légèrement des topologies prédites par l'analyse bio-informatique des profils hydropathiques[125,127,132]. Ainsi, pour *Sc*Stt3, évoqué plus haut, un domaine transmembranaire avait été décrit dans la boucle intra-luminale 1 et le cinquième domaine transmembranaire n'avait pas été détecté lors de l'étude *in silico*[130].

Finalement, malgré des divergences importantes de séquences primaires, la comparaison des structures résolues des différentes Stt3 montre une convergence architecturale remarquable de l'organisation tridimensionnelle des différents membres de cette famille GT-66. Ainsi il existe seulement 26% d'identité de séquence entre *h*Stt3a et

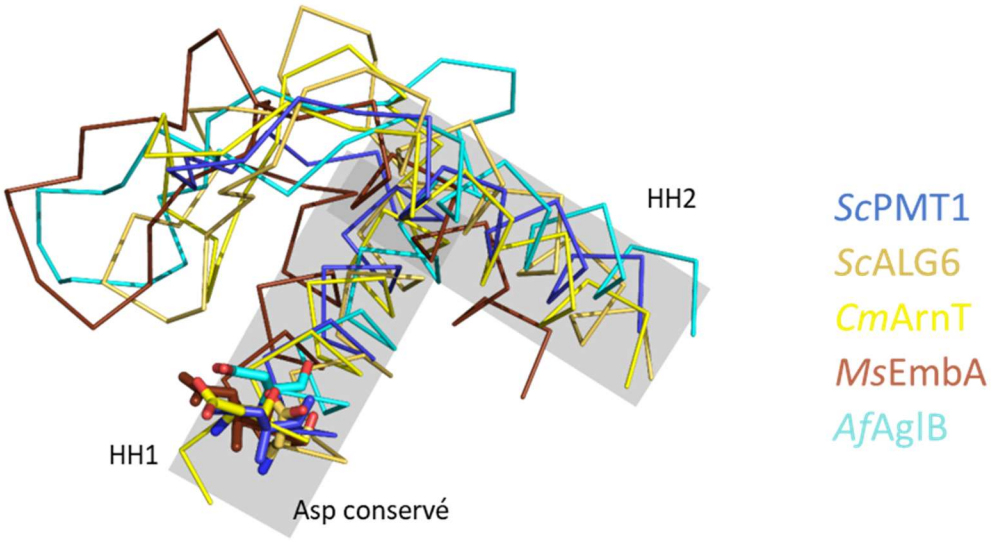


Figure 12. Superposition des hélices horizontales 1 et 2. Représentation fil de fer du squelette des hélices horizontales 1 et 2 (HH1 et HH2) de GT-Cs des 5 familles CAZy pour lesquelles des structures ont été résolues : GT-53 (*Ms*EmbA[121]), GT-57 (*Sc*ALG6[123]), GT-39 (*Sc*PMT1[88]), GT-66 (*Af*AglB[124]) et GT-83 (*Cm*ArnT[128]). Les acides aspartiques situés en amont de HH1 pour chaque enzyme sont mis en évidence en représentation bâtons. Les hélices sont encadrées en gris. La coloration est selon la légende à droite et reprend les couleurs de la Figure 9.

*Af*AglB alors que l'alignement des deux structures montre une superposition importante des segments transmembranaires et une forte convergence structurale du segment protubérant soluble (**Figure 10A**). Le même constat peut être réalisé sur le reste de la famille GT-66 où presque tous les membres, à l'exception de *Cl*PglB, semblent montrer une conservation de l'organisation de leurs 13 domaines transmembranaires et une disposition très similaire de leur domaine globulaire soluble (**Figure 10B**). *Cl*PglB, en revanche, diverge dans le positionnement de ses domaines transmembranaires 8 et 9 et ne présente donc une identité de positionnement avec le reste de la famille que pour 11 de ses 13 segments transmembranaires (segment 1 à 7 et 10 à 13)(**Figure 10B et Figure 11**).

De manière intéressante, la comparaison de l'ensemble des structures de GT-C résolues permet de mettre en évidence que bien que décrivant des topologies comptant de 11 à 15 domaines transmembranaires[88,122], ces protéines partagent un socle structural commun organisé autour de 11 domaines transmembranaires (**Figure 11**). Ce socle se traduit par un agencement à la membrane similaire de 11 segments transmembranaires auxquels peuvent être ajoutées jusqu'à 4 hélices transmembranaires supplémentaires. Une autre caractéristique de ce socle structural partagé par toutes les GT-C est la présence dans la boucle soluble délimitée par les deux premiers segments transmembranaires de ce socle commun, de deux hélices conservées (appelées hélices HH1 et HH2) disposées horizontalement parallèlement à la surface de la membrane (**Figure 12**)[88,116,123–125,128].

En conséquence, il apparait de l'analyse des différentes structures résolues que les 13 segments caractéristiques de ce socle structural commun (11 domaines transmembranaires et 2 hélices horizontales) présentent une disposition similaire d'une GT-C à une autre ainsi qu'une relative conservation de séquences comparée aux autres domaines tels que les boucles solubles localisées de part et d'autre de la membrane ou les segments transmembranaires additionnels.

Il est intéressant de souligner que ce socle structurellement conservé partagé par toutes les GT-C caractérisées à ce jour est le support de la conservation fonctionnelle du mécanisme catalytique caractéristique des glycosyltransférases de cette superfamille. En effet, quelle que soit leur classification CAZy, toutes les glycosyltransférases de type GT-C, y compris les PMT, ont pour caractéristiques communes d'utiliser un substrat donneur de sucre

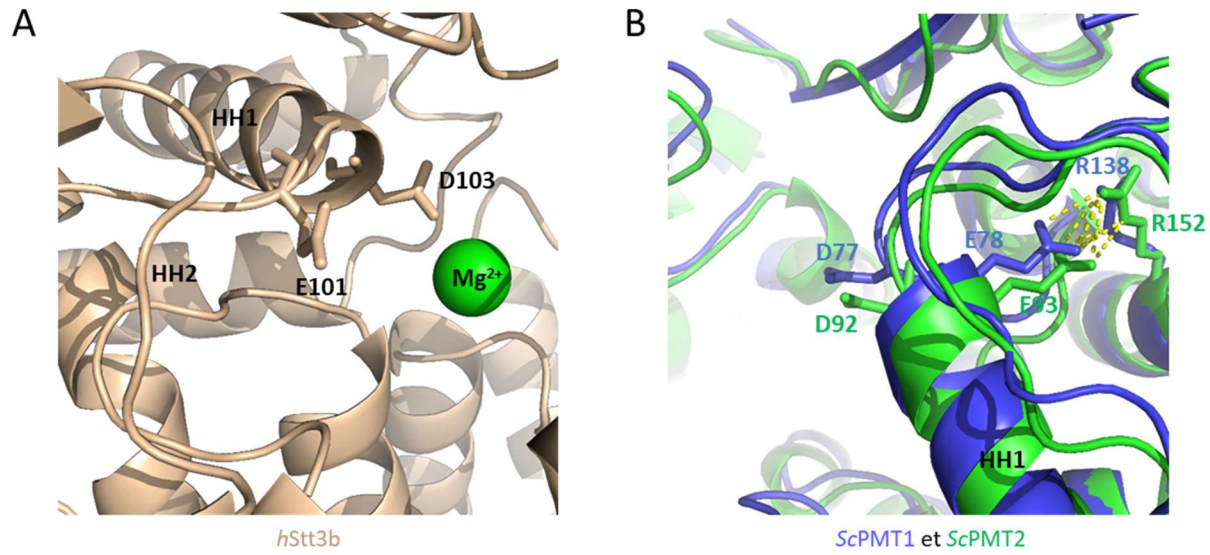


Figure 13. Le motif DxD (ou dérivé) des GT-Cs. **A)** Représentation ruban du motif DxD de *h*Stt3a[127]. Les acides aspartiques du motif DxD sont numérotés et représentés en bâtons de part et d'autre de l'ion divalent représenté par une boule verte. Les hélices HH1 et HH2 sont indiquées. **B)** Superposition des représentations ruban du motif DE de *Sc*PMT1 et 2[88]. Les acides aminés acides du motif DE sont représentés en bâton et numérotés. L'arginine conservée formant un pont salin avec l'acide glutamique du motif DE est représentée en bâton et numérotée. Les pointillés jaunes représentent les possibles interactions entre les acide glutamique et les arginines. L'hélice HH1 est indiquée.

de nature lipidique (poly-isoprenoyl-phosphate) et d'entrainer une inversion de la configuration anomérique du sucre possiblement grâce à un mécanisme de transfert de type Sn2 (ou Sn2-like) impliquant un résidu acide servant de catalyseur basique à la formation d'un intermédiaire oxocarbenium[133].

c) Conservation fonctionnelle

Le lien de conservation entre la structure et la fonction des GT-C est particulièrement mis en évidence par la présence de cet acide aminé acide essentiel (Asp ou Glu) conservé en amont de HH1 décrit dans toutes les GT-C de structure connue (**Figure 12**). Cet acide aminé peut être considéré comme une forme dégénérée du motif DxD reminiscent des GT-A. Alors que dans les GT-A, ce motif est bien conservé, dans les GT-C, la conservation est moins stricte avec une éventuelle substitution des aspartates en glutamates et une perte de l'acide aminé intermédiaire. Tout comme dans les GT-A, ce motif DxD, et ses différentes déclinaisons, est essentiel à l'activité glycosyltransférase des GT-C. Enfin, à la différence des GT-A, il n'a pas pu être établi avec certitude si cet acide aminé acide est impliqué ou non dans la complexation de cation divalent. Ainsi la structure résolue de *h*Stt3b révèle que les acides aminés du motif conservé ExD sont impliqués dans la coordination d'un cation métallique divalent[97,127] (**Figure 13A**) alors que l'activité des protéine *O*-mannosyltransférase humaines *h*POMT ne semble pas requérir la présence d'ion [97,127].

Un autre aspect de la conservation des relations structure-fonctions au travers de la superfamille des GT-C est la préservation de l'interface d'interaction et de reconnaissance du substrat lipidique donneur de sucre. L'homologie des dérivés phospho-isoprénol donneurs utilisés par les GT-C sous-tend que les sites d'interaction aux substrats phospholipidiques présentent des analogies structurales (**Figure 14**). C'est ce que révèlent les différentes structures résolues de *Sc*ALG6, *Sc*PMT1, *Ms*EmbB, *Cl*PglB et *Cm*ArnT[88,121,123,128,134] en mettant en évidence le long de la sixième hélice transmembranaire du socle commun, la présence d'un sillon hydrophobe orienté perpendiculairement au plan de la membrane et débouchant à sa surface dans une cavité à proximité immédiate du site catalytique. Cette interface structurellement conservée permet d'accommoder la partie aglycone aliphatique

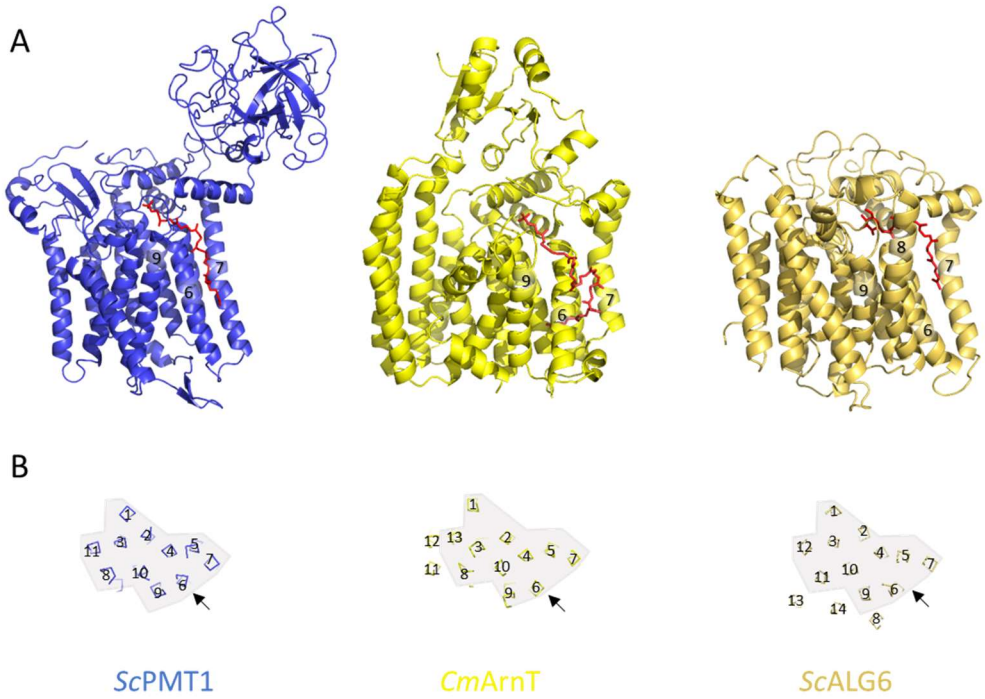


Figure 14. Zone d'interaction avec le substrat lipidique donneur. **A)** Représentations ruban de *Sc*PMT1[88], *Cm*ArnT[27] et *Sc*ALG6[22]. Ces structures ont été résolues avec un lipide-phosphate qui est représenté en rouge dans chaque structure. **B)** Représentation schématique des structures selon une coupe au niveau de la membrane vue de dessus des mêmes enzymes. La flèche indique le point d'entrée du lipide substrat donneur pour chaque enzyme. Le module GT-C conservé de 11 domaines transmembranaires est grisé.

polyprenoïque du substrat donneur le long du segment transmembranaire en positionnant l'extrémité polaire phospho-glycanne dans la cavité catalytique riche en charges positives permettant de stabiliser le groupement phosphate et d'orienter correctement la liaison phospho-sucre en regard du carboxylate de l'acide aminé catalytique[123,134,135].

Au-delà de ce socle conservé, toutes les GT-C affichent cependant plusieurs domaines de séquences et de structures variables au niveau des différentes boucles solubles localisées de part et d'autre de la membrane. Certaines de ces boucles notamment les boucles protubérantes de grande taille, délimitées par les segments transmembranaires 1-2 et 7-8 (3-4 et 9-10 pour les membres de la famille GT-53) du socle commun ont été montrées comme potentiellement impliquées dans la sélection du motif glycannique du substrat donneur ou dans les interactions avec le substrat accepteur[123,125]. La variabilité de séquence de ces domaines au sein des GT-C permettrait ainsi d'accommoder les différents substrats et serait à l'origine de la grande spécificité enzymatique des membres de cette superfamille de glycosyltransférases pour des motifs glycanniques (Glc, GlcNac, Man, Ara, oligosaccharides complexes...) et des substrats accepteurs (lipides, oligosaccharides, protéines) d'une grande diversité.

d) Les PMT : GT-C de la famille 39

Sur la base de leur homologie de séquence et de fonction, les PMT ont été regroupées au sein de la famille 39 de la superfamille des GT-C mais ce n'est qu'en 2019, que les premières structures tridimensionnelles de PMT ont été rendues accessibles avec la publication de la structure du dimère *Sc*PMT1 et *Sc*PMT2 de *S. cerevisiae* en complexe avec un tétrapeptide accepteur et un résidu Dol-P. Ces structures déterminées par cryomicroscopie électronique à une résolution de 3,2Å[88] révèlent une topologie à 11 domaines transmembranaires ainsi que la présence des deux hélices HH1 et HH2 parallèles au plan de la membrane et situées dans la première boucle soluble délimitée par les segments transmembranaires 1 et 2. On retrouve également en amont immédiat de l'hélice HH1, la dyade catalytique DE conservée dans les PMT et démontrée comme essentielle chez la levure avec une perte totale d'activité lors de la substitution de ce motif par deux alanines[54]. L'étude de la structure résolue du complexe *Sc*PMT1-*Sc*PMT2 permet d'éclaircir le rôle de ce doublet DE en démontrant que l'acide aspartique (*Sc*PMT1-D77 ou *Sc*PMT2-D92) est idéalement positionné pour activer l'hydroxyle

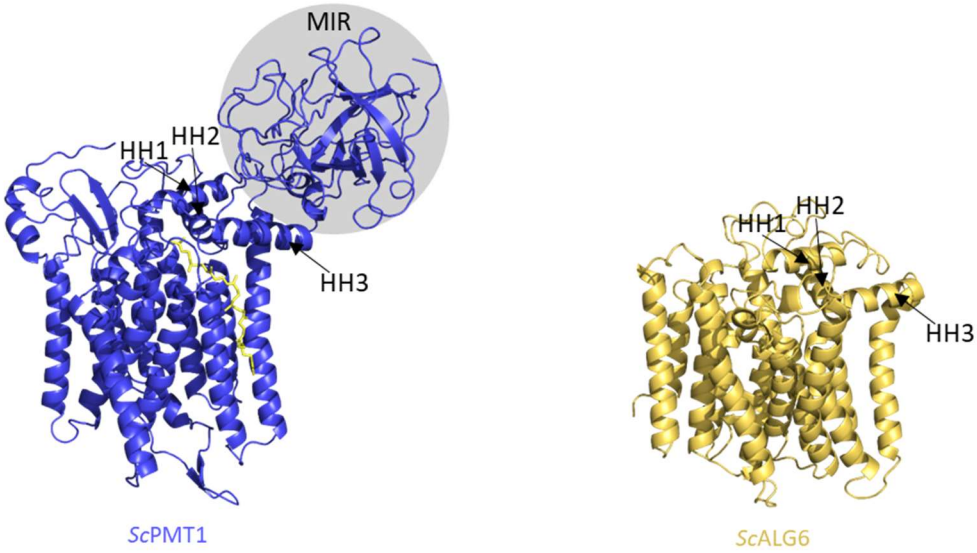


Figure 15. Plusieurs GT-Cs décrivent une troisième hélice horizontale (HH3). Représentation en ruban de différentes GT-Cs possédant une troisième hélice horizontale : *Sc*ALG6[123] et *Sc*PMT1[88]. HH1, 2 et 3 sont indiqués sur les 2 structures. Le domaine MIR est indiqué sur ScPMT1. Le substrat lipidique donneur de ScPMT1 est représenté en jaune dans la structure.

de l'acide aminé accepteur du peptide cible alors que l'acide glutamique (*Sc*PMT1-E78 ou *Sc*PMT2-E93) forme un pont salin avec une lysine conservée (*Sc*PMT1-R138 et *Sc*PMT2-R152) (**Figure 13B**). Ce pont salin aurait pour effet de maintenir la cohérence structurale du site actif, forçant l'acide aspartique du motif DE à orienter sa chaine latérale de sorte à pouvoir interagir avec l'hydroxyle du substrat accepteur.

Outre les éléments structuraux communs à toutes les GT-C, cette étude bibliographique permet d'appréhender l'organisation de la 4^{e} boucle intra-luminale délimitée par les segments transmembranaires 7 et 8 et comportant les motifs MIR caractéristique des PMT eucaryotes. Elles révèlent aussi la présence d'une 3^{e} hélice α horizontale (HH3) en aval des domaines MIR décrite également dans d'autres GT-C telle que *Sc*ALG6 [123] (**Figure 15**).

Enfin, cette étude a aussi permis de démontrer que les PMT adoptent une conformation membranaire caractéristique basée sur le cœur structural commun à toutes les GT-C. Il est cependant intéressant de noter qu'à l'exception des domaines MIR et de l'hélice HH3 portés par la quatrième boucle intra-luminale, les PMT de S. cerevisiae semblent être structuralement constituées quasi exclusivement du socle commun des GT-C et apparaissent comme les représentants les plus simples des GT-C dont les structures ont été résolues à ce jour.

Les structures de ces PMT eucaryotes ont été publiées dans le courant de ma thèse et ont servi de base pour la construction du modèle tridimensionnel de la protéine O-mannosyltransférase bactérienne sur lequel je me suis appuyé pour réaliser l'étude mutationnelle détaillée des relations structures-fonctions de PMTub de *M. tuberculosis* présentée dans la suite de ce document.

Résultats et discussions

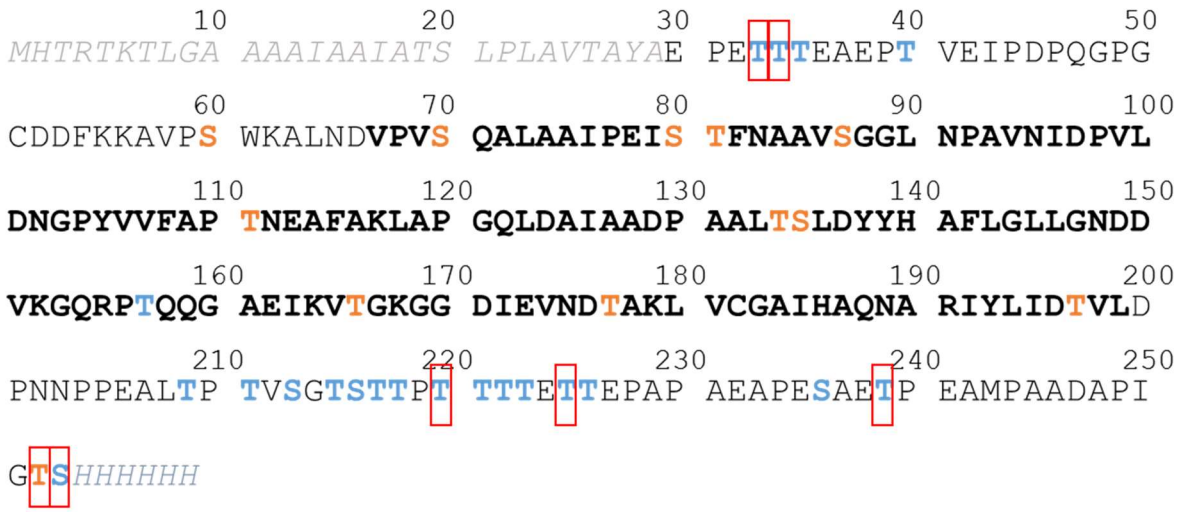


Figure 16. Séquence annotée de la FasC recombinante utilisée dans cette étude. En italique et gris clair sont annotés le peptide signal et l'étiquette histidine respectivement. En gras est représenté le domaine fascicline. Les sites de glycosylations prédits par le logiciel NetOglyc 4[136] sont indiqués en bleu. En orange sont les Thr et les Ser non détectées par NetOglyc 4. Encadré en rouge sont les acides aminés détectés par spectrométrie de masse comme mannosylés sur de la FasC recombinante purifiée[137].

Partie 1 : Développement d'outils pour l'étude fonctionnelle de PMTub

Comme discuté plus haut la première étape *sine qua non* de mon projet d'étude des relations structure/fonction de PMTub pour le criblage d'inhibiteurs a été le développement d'un bio essai quantitatif robuste pour suivre l'activité de PMTub en parallèle dans des échantillons multiples.

Pour suivre l'activité de PMTub par un test phénotypique, les conditions optimales auraient été de travailler directement sur la bactérie *M. tuberculosis*. Dans cette configuration, nous aurions pu analyser l'activité de l'enzyme dans son environnement natif. Toutefois, du fait de la pathogénicité de cette bactérie, de sa croissance lente et de la conséquence de la suppression du gène de rv1002c sur sa croissance, l'utilisation d'un organisme modèle est préférée. Pour cela nous avons choisi de construire par ingénierie génétique une souche mycobactérienne recombinante exprimant la protéine *O*-mannosyltransférase de *M. tuberculosis*, PMTub, conjointement à une mannoprotéine cible dont nous pourrons quantifier le taux de mannosylation.

A. Construction d'une souche rapportrice de l'activité de PMTub

L'organisme modèle choisi est dérivé de *Mycobacterium smegmatis* en raison de son phénotype à croissance rapide, sa non-pathogénicité et de sa proximité génétique avec *M. tuberculosis*.

Pour étudier l'activité de PMTub, une souche recombinante de *M. smegmatis* a été conçue pour exprimer PMTub en lieu et place de sa propre PMT. Pour cela le gène Msmeg_5447 codant pour la PMT de *M. smegmatis* a été interrompu par échange allélique tel que décrit précédemment, conduisant à une bactérie ci-après nommé « ΔPMT » dépourvue d'activité protéine *O*-mannosyltransférase endogène[52]. Cette souche a alors été transformée par un plasmide intégratif pour surexprimer le gène de PMTub (rv1002c) sous le contrôle du promoteur constitutif fort pBlaF*.

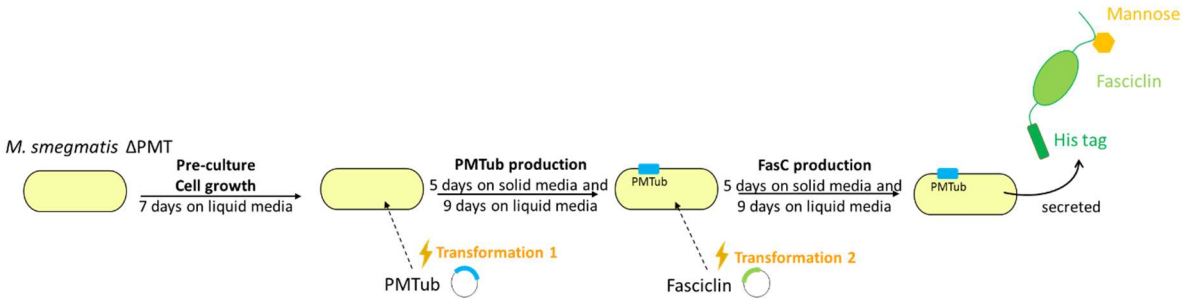


Figure 17. Stratégies expérimentales pour suivre l'activité in vivo de la protéine O-mannosyltransférase PMTub. Construction d'une souche de *M. smegmatis* rapportrice de l'activité PMTub. La souche ΔPMT subit deux transformations successives pour exprimer PMTub et secréter la FasC^{6His}.

Pour évaluer la protéine *O*-mannosylation, nous avons choisi de suivre le taux de mannosylation d'une mannoprotéine mycobactérienne rapportrice : la « fasciclin domain protein » de *M. smegmatis* (gène MSMEG_5196) aussi dénommée fascicline ou FasC. La FasC est un substrat naturel de la PMT endogène de *M. smegmatis* et des études antérieures ont montré qu'elle comporte un grand nombre de sites de mannosylation[52,137] (**Figure 16**). En outre cette protéine est sécrétée dans le milieu de culture et peut donc être facilement récupérable. Pour faciliter son isolement, une étiquette histidine a été rajoutée à l'extrémité C-terminal du gène de la FasC porté par un plasmide multi-copie utilisé pour transformer *M. smegmatis*. Enfin, des analyses par spectrométrie de masse ont permis de démontrer précédemment que le profil de glycosylation de la FasC est similaire lorsqu'elle est mannosylée par la PMT endogène de *M. smegmatis* ou par PMTub exprimée dans la souche, ΔPMT:PMTub[137].

En résumé, la construction de la souche rapportrice comprend deux transformations successives de la souche ΔPMT afin d'exprimer parallèlement PMTub et FasC^{6His} (**Figure 17**).

Pour mettre au point et optimiser le test d'activité phénotypique, nous avons choisi d'analyser le taux de mannosylation de la FasC issue de la souche ΔPMT:PMTub exprimant le gène sauvage de PMTub que nous avons comparé à celui de la protéine produite par la souche « contrôle négatif » ΔPMT transformée avec un vecteur contrôle vide et dénommée « No PMT ». Enfin une souche ΔPMT:PMTub* transformée par un plasmide portant une mutation ponctuelle sur la séquence de PMTub a été également utilisée dans le but d'évaluer le caractère quantitatif du bio-essai.

Le suivi du taux de mannosylation de la FasC devant permettre d'évaluer l'activité de transfert de mannose catalysée par PMTub, plusieurs approches méthodologiques ont été appliquées pour détecter la présence de mannose et quantifier le taux de mannosylation de la FasC.

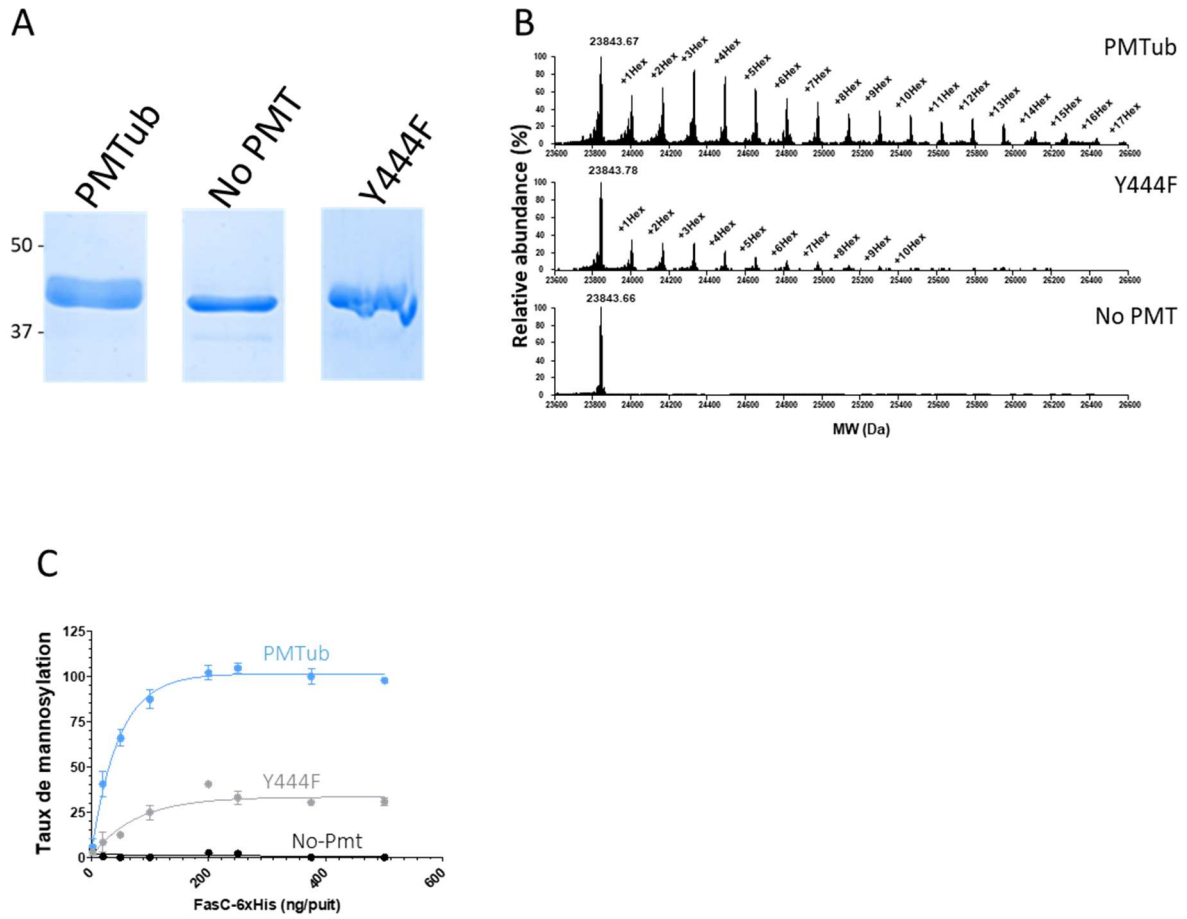


Figure 18. Validation du test d'activité phénotypique de PMTub. Activité de PMTub mesurée pour trois souches de *M. smegmatis* ΔPMT, complémentées avec les plasmides PMTub, Y444F ou No-Pmt, et le plasmide FasC^{6His}. **A)** SDS-PAGE de FasC^{6His} purifiée pour les différentes souches. **B)** Analyse par spectrométrie de masse du profil de glycosylation de FasC6His purifiée, par LC-ESI-native-MS. Les spectres ont été normalisés en fonction de l'intensité du signal observé pour la forme non glycosylée. **C)** Courbes dose-réponse obtenues avec le dosage employant la lectine ConA, pour des concentrations croissantes de FasC^{6His}. Les activités ont été mesurées à partir des surnageants de culture déduites de l'intensité du signal obtenu par l'ELISA lectinique, normalisée par le signal le plus fort obtenu pour PMTub. Des duplicats biologiques (analysés en triplicats techniques) sont présentés.

B. Mise en évidence et quantification de la mannosylation de FasC

1. Par détection d'une altération du profil électrophorétique

En première intention, la glycosylation de la FasC peut être mise en évidence par l'observation d'une modification du profil de migration de la fascicline purifiée sur SDS-PAGE (**Figure 18A**). On peut en effet observer une bande fine pour la fascicline non glycosylée produite par la souche No PMT alors que la fascicline produite par la souche exprimant PMTub native montre une bande plus large résultant de la présence de glycoformes de masses apparentes différentes. Toutefois cette différence se manifeste uniquement sur la largeur de la bande. L'absence de bandes discrètes distinctes pour les formes glycosylées et non glycosylées de la protéine empêche toute évaluation quantitative de ces 2 états. Ceci est clairement illustré par le profil obtenu pour la FasC produite par la souche rapportrice exprimant un mutant (Y444F) de PMTub dont nous avons pu par ailleurs démontrer qu'il avait une activité de mannosylation réduite, mais non nulle (voir ci-après et p.123). On observe dans ce cas, une bande large peu différente de celle obtenue avec la souche exprimant la protéine PMTub sauvage. Cette approche de type « tout ou rien » est donc purement qualitative. Elle permet de mettre en évidence l'activité de l'enzyme, mais manque de résolution pour détecter des variations quantitatives d'activité que nous souhaitons pouvoir mesurer dans la suite de notre projet.

2. Par spectrométrie de masse de protéine native

Des mesures ont également été réalisées par chromatographie liquide (LC) couplée à une analyse par spectrométrie de masse (MS) de protéines entières après purification de la $FasC^{6His}$ sur colonnes Ni-NTA (**Figure 18B**). L'étude du profil de glycosylation de la FasC par cette approche permet d'obtenir un aperçu semi-quantitatif de la variation du degré de mannosylation de la protéine rapportrice en fonction des souches.

Le spectre de masse obtenu pour la FasC issue de la souche exprimant PMTub sauvage, montre plusieurs pics d'intensité décroissante et distants de 162 Da correspondant à la masse d'un résidu d'hexose, mettant en évidence la présence de plusieurs glycoformes de la FasC. La masse de 23 843,67 Da observée pour le pic le plus intense est en accord avec la masse moléculaire monoisotopique attendue pour la $FasC^{6His}$ non glycosylée ($M_{calculée}$ 23 844 Da). Cette analyse confirme que la $FasC^{6His}$ produite par la souche de *M. smegmatis* exprimant

PMTub sauvage est fortement glycosylée, comportant jusqu'à 17 unités glycosidiques détectables (**Figure 18B)** Comme attendu, le spectre obtenu pour la FasC^{6His} purifiée produit par la souche n'exprimant pas de PMT montre un seul pic, correspondant à la protéine non glycosylée, en accord avec la perte totale de la glycosylation de protéines en l'absence de PMT[52]. Les résultats obtenus pour ces deux souches témoins sont fiables en termes de présence ou d'absence de glycosylation de la FasC^{6His} et peuvent donc être utilisés comme contrôle positif et négatif pour la mannosylation des protéines. Il est intéressant de noter que l'analyse du spectre de masse obtenue pour la FasC produite par la souche exprimant le gène muté de PMTub confirme clairement que la mutation Y444F affecte la glycosylation de la FasC^{6His}. En effet le spectre de masse montre un nombre réduit de glycoformes détectées (ne comportant que jusqu'à 10 unités glycosyl liées) dont les intensités relatives cumulées sont nettement plus faibles que celles observées pour la FasC issue de la souche exprimant la forme sauvage de PMTub.

L'analyse par spectrométrie de masse permet donc d'obtenir des résultats plus complets que l'approche par SDS PAGE. Cependant, la quantification précise par cette approche de l'impact de la mutation sur l'activité de PMTub reste difficile du fait de la variabilité des mesures liée aux propriétés intrinsèques de la technique qui nécessite un nombre significatif de réplicats techniques. De plus, une autre limite de l'approche est la nécessité de purifier la FasC ce qui introduit une autre source de variabilité pouvant impacter la quantification.

3. Par reconnaissance par une lectine

Nous avons donc développé un protocole permettant de déterminer de manière quantitative le taux de mannosylation de la FasC et qui ne nécessite pas d'étape de purification préalable. Le surnageant de culture brut contenant la FasC sécrétée est donc collecté et analysé par un protocole en deux étapes.

Le principe de ce protocole est de quantifier le taux de mannoses terminaux présents sur la FasC et de mesurer parallèlement la quantité de FasC pour évaluer le taux de mannosylation de la protéine FasC.

Pour mesurer la quantité de mannose sur la FasC, nous avons choisi d'effectuer un dosage par la lectine Concanavaline A (ConA) qui reconnait et se fixe sur les résidus mannose terminaux des oligosaccharides. Compte tenu de l'absence de ramification des motifs oligomannosidiques des mannoprotéines mycobactériennes, la ConA va ainsi permettre d'évaluer le nombre mannose terminaux et par conséquent le nombre de site de mannosylation de la FasC qui refléteront indirectement le niveau d'activité de transfert de PMTub.

Les mycobactéries sécrètent cependant de nombreux glycoconjugués à mannose tel que le lipomannane (LM) ou les phosphatidyl-inositol mannosides (PIM) potentiellement présents dans le milieu de culture. Pour nous affranchir de ces composés sans pour autant devoir effectuer une étape préalable de purification complémentaire de la FasC, nous avons choisi d'utiliser des plaques multipuits greffées avec des anticorps anti-HisTag pour isoler la $FasC^{6His}$ des autres composés mannosylés présents dans le surnageant de culture. La mannosylation de la $FasC^{6His}$ ainsi retenue par les anticorps greffés sur la plaque, peut alors être évaluée par la lectine ConA couplée à la HRP (Horse Radish Peroxidase).

Cette configuration en sandwich choisie pour cet ELISA lectinique réalisable en plaque multipuits présente l'avantage de requérir une faible quantité de produits et d'être facilement adaptable à un protocole de crible d'inhibiteurs à haut débit.

Enfin, la concentration de $FasC^{6His}$ dans les échantillons de surnageant mesurée en parallèle par un test ELISA direct avec des anticorps anti-HisTag permet de normaliser les mesures rendant possible la comparaison du taux de mannosylation relatif de la FasC dans des échantillons différents.

En pratique, pour faciliter l'étape de normalisation de la concentration en FasC, nous avons choisi d'utiliser, pour l'ELISA lectinique, des plaques multipuits possédant un seuil de saturation de 200ng de protéines étiquetées His par puits. L'ELISA direct nous permet donc de nous assurer que nous sommes au-dessus de ce seuil de saturation de 200ng/puits pour normaliser la quantité de FasC et rendre possible l'analyse comparative d'échantillons multiples.

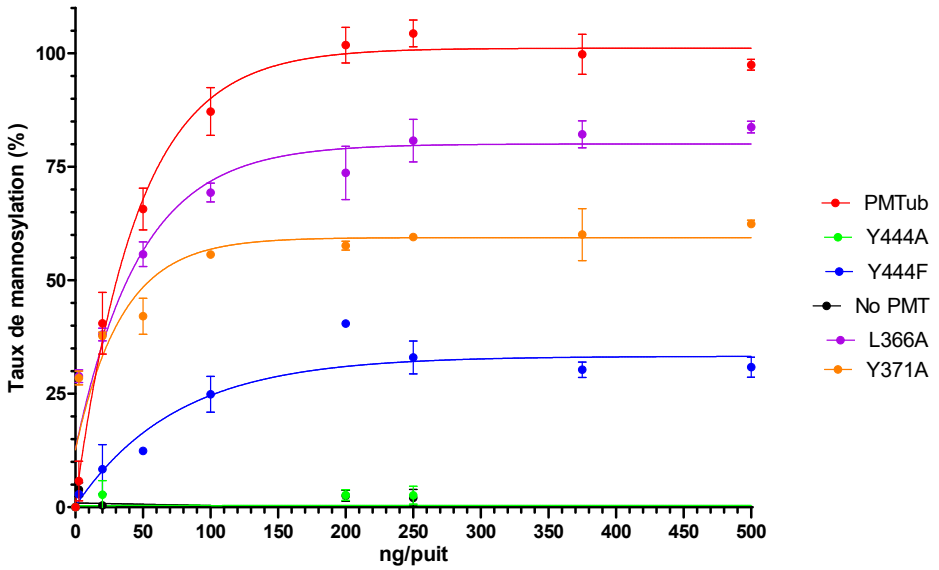


Figure 19. Validation du test d'activité phénotypique de PMTub. Activité de PMTub mesurée pour différentes souches de *M. smegmatis* ΔPMT, complémentées avec les plasmides PMTub, Y444F, Y444A, L366A, Y371A ou No-Pmt, et le plasmide FasC^{6His}. Courbes dose-réponse obtenues avec le dosage employant la lectine ConA, pour des concentrations croissantes de FasC^{6His}. Les activités ont été mesurées à partir des surnageants de culture déduites de l'intensité du signal obtenu par l'ELISA lectinique, normalisée par le signal le plus fort obtenu pour PMTub. Des duplicats biologiques (analysés en triplicats techniques) sont présentés.

Comme cet ELISA lectinique doit être suffisamment sensible et fiable pour comparer un large panel de surnageants de *M. smegmatis* (qui peuvent varier en termes de concentration de FasC et de mannosylation), nous avons vérifié sa robustesse en mesurant sa réponse à différentes doses de FasC. Nous avons utilisé différentes concentrations de FasC (allant de 1 à 500 ng/puits) produites par les différentes souches décrites précédemment (PMTub native, No PMT et Y444F) (**Figure 18C**).

Premièrement, comme attendu, nous constatons que la FasC non glycosylée produite par la souche « No PMT », n'est pas reconnue par la Concanavaline A ce qui se traduit par une absence de signal. Ce résultat permet d'évaluer le niveau basal du bruit de fond du test et de constater le niveau de signal pratiquement nul attribuable à une éventuelle contamination de l'échantillon par d'autres glycoconjugués à mannose présents dans le milieu de culture.

A l'inverse, les résultats obtenus à partir des surnageants de culture de la souche exprimant PMTub sauvage permettent de mettre en évidence une effet dose-réponse avec saturation. En effet, la valeur du signal mesuré augmente avec la quantité de FasC dans le puit, jusqu'à atteindre un plateau de saturation pour des concentrations en FasC supérieures à 200 ng par puit. Cette valeur mesurée correspondant à la capacité maximale de fixation de protéines étiquetées His dans le puit est en accord avec celle annoncée par le fabriquant des plaques utilisées dans notre test.

Finalement, les valeurs de la réponse obtenue pour la FasC produite par la souche exprimant la forme mutée de PMTub, permettent de mettre en évidence et de quantifier une réduction de l'ordre de 70% du taux de mannosylation relatif comparé aux valeurs obtenues pour la FasC produite par la souche exprimant PMTub sauvage. Ce test permet donc de quantifier l'effet d'une mutation sur le taux de mannosylation de la FasC. De plus, on peut observer que quelle que soit la concentration de FasC, la proportion relative de fixation de ConA détectée dans les surnageants de culture de la souche exprimant PMTub Y444F reste constante avec un effet dose et un niveau de saturation à 200 ng de FasC correspondant à 30% de la valeur du taux de mannosylation de la FasC produite par la souche exprimant PMTub sauvage.

Cette analyse confirme ainsi un effet dose-réponse de l'ELISA lectinique et prouve la fiabilité de la technique de quantification de la mannosylation des protéines. Nous confirmons aussi la présence du seuil de saturation à 200ng/puit, ce qui permet de valider notre protocole

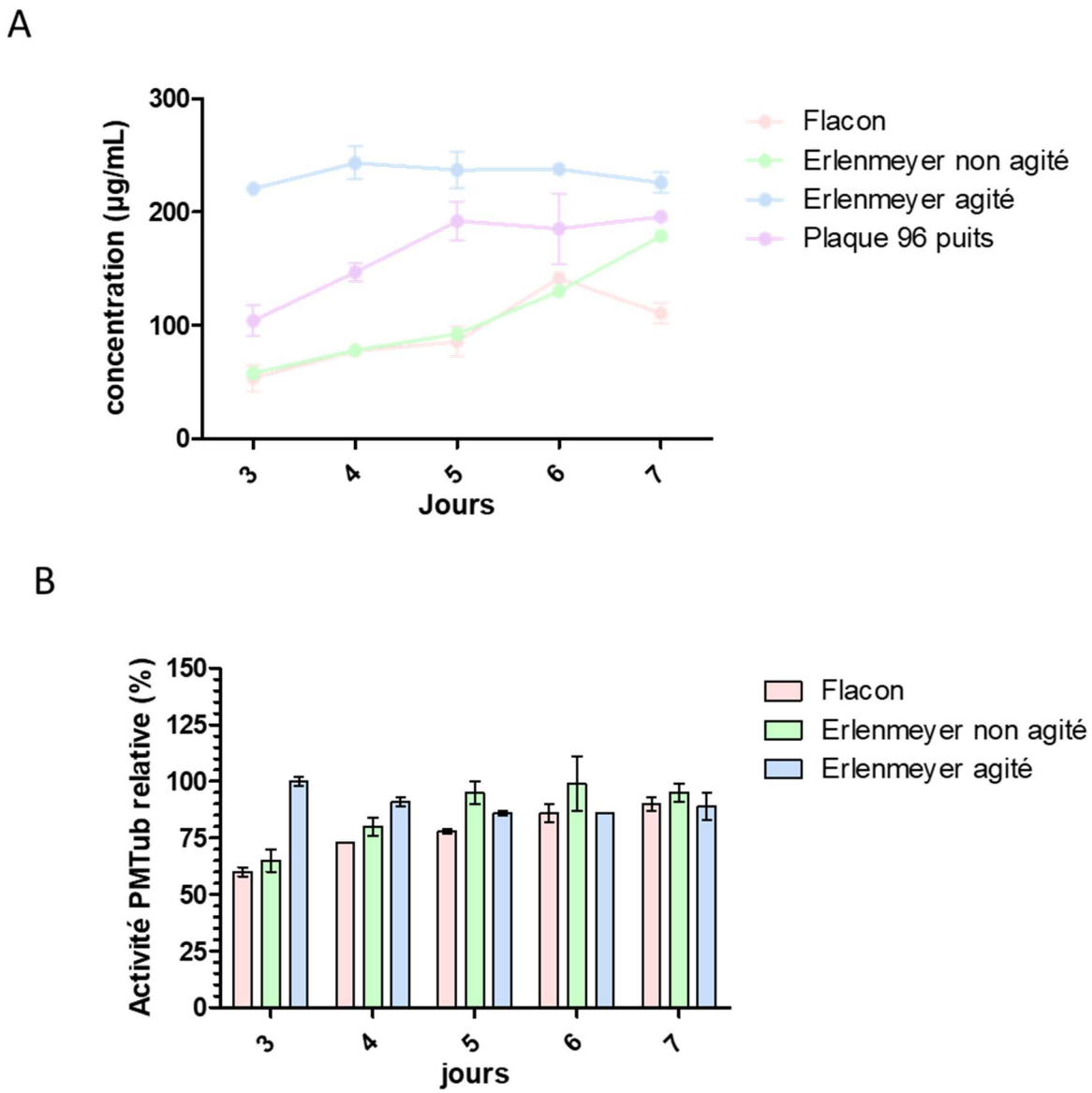


Figure 20. Suivi de la production de fascicline O-mannosylée en fonction des contenants utilisés pour la culture. **A)** Mesure de la concentration en fascicline au cours du temps ($FasC^{6His}$) dans les surnageants de culture **B)** Mesure du taux de mannosylation de la fascicline ($FasC^{6His}$) au cours du temps dans les surnageants de culture. Les résultats présentés sont des triplicats techniques de replicats biologiques.

de normalisation de la FasC. Cette analyse a aussi été effectuée pour un nombre de mutants plus important de PMTub consolidant les résultats obtenus (**Figure 19**).

Il a été démontré que les conditions de culture peuvent impacter la *O*-glycosylation des protéines mycobactériennes[37]. Nous avons donc entrepris de comparer l'effet de différentes conditions de culture sur l'activité mannosyltransférase de PMTub. La production de FasC et son taux de mannosylation ont été comparés lors de cultures en flacon ou en erlenmeyer, avec ou sans agitation. Ces différentes méthodes de culture nous ont ainsi permis de déterminer l'effet de la taille de la culture et du contenant sur l'activité *O*-mannosyltransférase.

Nous avons observé une augmentation de la concentration en Fascicline au cours du temps de façon continue pour tous les modes de culture à l'exception de la condition agitée qui atteint son maximum dès le jour 3 (**Figure 20A**). Ceci s'explique probablement par un taux d'oxygénation accru favorisant une croissance bactérienne plus rapide de cette condition. D'après ces observations, on peut conclure que la fascicline est produite en quantité suffisante pour effectuer un ELISA lectinique dans toutes les conditions dès le troisième jour de culture.

La mesure de la *O*-mannosylation de la FasC par ELISA lectinique montre une augmentation relative de l'activité *O*-mannosyltransférase dans le temps pour certaines conditions de culture (**Figure 20B**). Malgré une normalisation de la concentration en FasC, nous observons une variation de la quantité de ConA fixée, traduisant une variation de la quantité de mannose portée par la FasC au cours du temps. Ce résultat suggère une régulation possible du niveau de modification post traductionnelle, probablement lié à la variation phénotypique des bactéries et à un processus de régulation de l'expression de l'enzyme lors des différentes phases de croissance exponentielle ou stationnaire de la culture. En effet, des travaux récents ont permis de mettre en évidence chez *M. tuberculosis*, une régulation transcriptionnelle du gène Rv1002c lors de l'infection[37] renforçant l'hypothèse d'une régulation possible de l'expression de PMTub selon les conditions de culture de la bactérie.

Afin de s'affranchir de ces biais méthodologiques éventuels, un protocole de culture homogène devra donc être appliqué rigoureusement pour que toutes les mesures d'activité relative soient effectuées dans des conditions comparables.

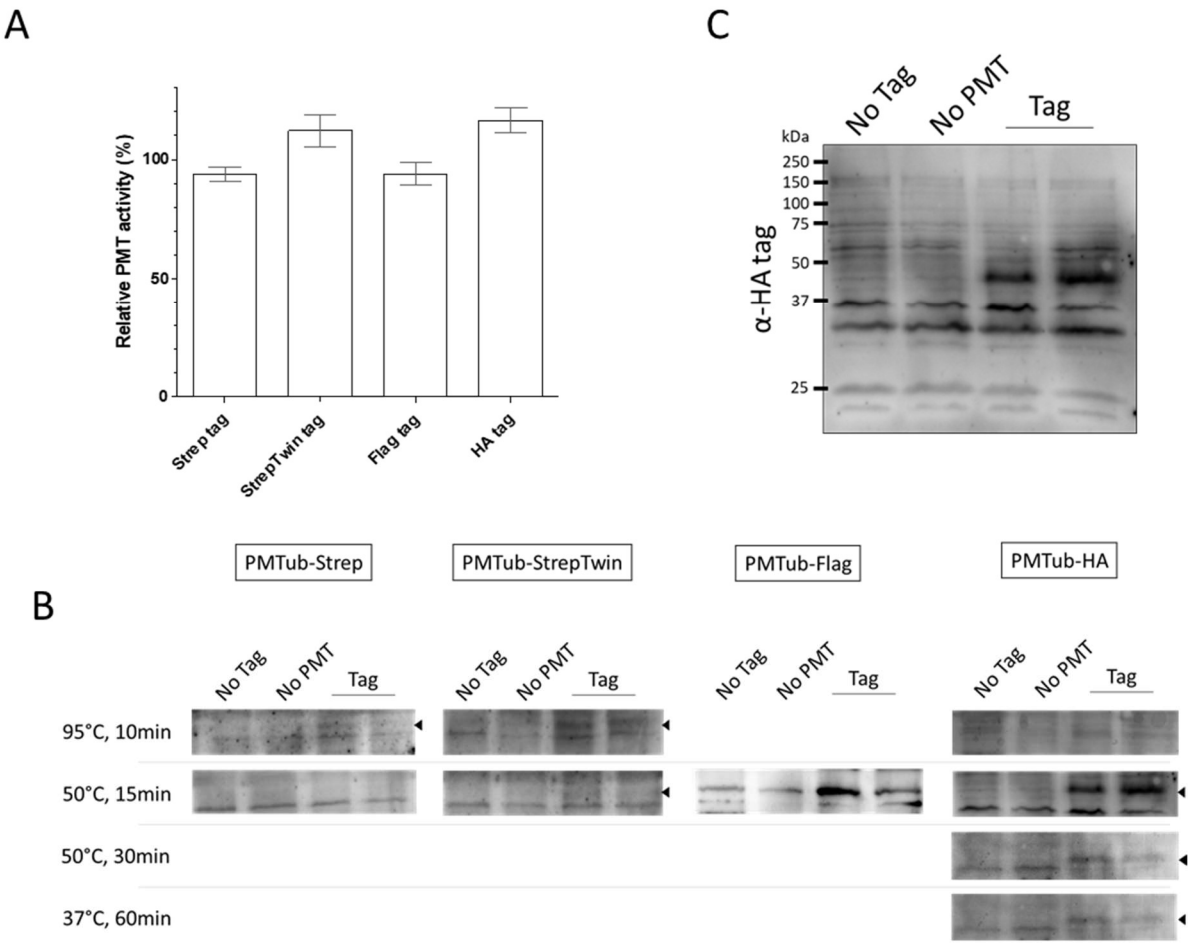


Figure 21. Contrôle de l'expression de PMTub par western blot. A) Activité relative *in vivo* de PMTub étiqueté. 0% et 100% sont définis respectivement comme la mannosylation mesurée pour la souche No PMT et la souche exprimant PMTub native. **B)** Western blot anti-tag sur la fraction contenant la paroi bactérienne pour les différentes étiquettes utilisées lorsque les échantillons sont traités selon différents protocoles dans le tampon Laemmli (température et durée). La portion représentée permet de visualiser les protéines migrant entre les marqueurs de taille 37 kD et 50 kD. **C)** Western blot anti-HA sur la fraction contenant la paroi bactérienne. Les témoins négatifs sont la fraction contenant la paroi bactérienne des souches exprimant PMTub native (sans aucun tag, No Tag) et la souche No PMT. Des duplicats biologiques de la fraction contenant la paroi bactérienne de la souche exprimant PMTub-HA sont présentés.

En conclusion, compte tenu du fait que PMTub soit la seule enzyme qui initie la mannosylation des protéines et en considérant que la mannosylation des protéines est linéaire chez les mycobactéries[52], nous avons postulé que la mesure du taux de fixation de la ConA sur la mannoprotéine permet d'évaluer son taux de mannosylation et par là-même, indirectement, l'activité *O*-mannosyltransférase de PMTub.

Toutefois cette approximation suppose que la quantité d'enzyme exprimée soit comparable. En effet, dans le cas d'une réduction apparente du taux de mannosylation comme observée pour la souche exprimant la forme PMTub Y444F par exemple, il est important de mesurer le niveau de production de l'enzyme mutée pour évaluer si la diminution du signal observée est due, comme attendu, à une diminution de l'activité enzymatique, ou bien à un défaut d'expression du gène.

C. Contrôle de l'expression de PMTub pleine taille à la membrane

Afin de consolider l'interprétation de notre bio-essai nous avons donc cherché à mettre en évidence l'expression de PMTub. Nous avons pour cela exploré différentes stratégies avec des résultats très inégaux.

1. Détection de l'expression de PMTub par western blot

La première stratégie mise en place pour la détection de PMTub « pleine taille » à la membrane plasmique de la bactérie a été par marquage de l'extrémité C-terminale par une étiquette détectable par des anticorps après transfert de la protéine sur membrane (méthode dite de western-blot ou immuno-blot).

Pour un premier essai, j'ai utilisé une étiquette de type « Strep-tag » constituée d'un octapeptide fusionné à l'extrémité C-terminale de PMTub. Cette étiquette, différente du « His-tag » utilisé pour étiqueter la FasC, a été choisie pour sa haute spécificité et la disponibilité d'outils de purification spécifiques des protéines ainsi marquées. Afin de vérifier si l'ajout de l'étiquette pouvait altérer l'activité de la protéine recombinante PMTub étiquetée Strep, la souche *M. smegmatis* ΔPMT a été successivement transformée par les 2 plasmides permettant l'expression respectivement de l'enzyme de fusion $PMTub^{Strep}$ et de son substrat $FasC^{6His}$. Ainsi que le montre la Figure 21A, nous avons observé pour cette souche un taux

relatif de mannosylation de la FasC proche de 100% comparé à la FasC issue de la souche exprimant la protéine PMTub native (**Figure 21A**). Cette expérience confirme l'expression hétérologue de l'enzyme PMTub^{Strep} fonctionnelle par *M. smegmatis* et démontre que l'ajout de l'étiquette Strep à l'extrémité C-terminale de PMTub n'affecte pas la capacité de l'enzyme à catalyser le transfert de mannose sur la FasC.

J'ai alors cherché à détecter par immunoblot la présence de PMTub^{Strep} dans différentes fractions subcellulaires afin de vérifier sa localisation membranaire. Pour cela j'ai utilisé un protocole de dissociation et de fractionnement de l'extrait cellulaire de la souche ΔPMT:PMTub^{Strep} précédemment décrit[75]. En bref, la lyse bactérienne est réalisée par sonication et différentes étapes de centrifugation et d'ultracentrifugation sont utilisées pour obtenir les fractions « paroi mycobactérienne », « membrane plasmique » et « cytoplasme ». Les différentes fractions ont ensuite été analysées par SDS-PAGE et révélées par western blot à l'aide d'un anticorps anti-strep tag. Malheureusement, les résultats de ce premier essai se sont avérés peu concluants : aucun signal spécifique n'a pu être détecté dans les conditions utilisées (données non montrées).

Nous avons donc voulu aller plus loin en criblant différentes étiquettes. Pour cela nous avons sélectionné plusieurs tags courts parmi ceux décrits dans la littérature, tout en vérifiant systématiquement que le marquage C-terminal de PMTub avec ces nouveaux tags n'interférait pas avec l'expression de l'enzyme sous une forme active (**Figure 21A**). Ainsi, en complément de l'étiquette Strep utilisée lors du premier essai, nous avons testé les étiquettes suivantes

- Le Twin-Strep tag qui consiste en une répétition de l'étiquette Strep censée augmenter la sensibilité de la reconnaissance par anticorps anti-Strep ;
- Le Flag tag: autre octapeptide d'environ 1kDa dont l'utilisation a été décrite pour détecter la mannosyltransférase membranaire PimE de *M. smegmatis* [138]
- Le tag HA nonapeptide(issu de l'hémagglutinine) décrit pour détecter les PMT de levure (utilisé en répétition en 3 ou 6 copies)[53,131].

Pour détecter l'expression à la membrane plasmique de PMTub fusionnée à ces différentes étiquettes, nous avons utilisé le même protocole que précédemment décrit, tiré

de la littérature[75]. Malheureusement, une fois encore aucune de ces constructions, à l'exception du tag Twin-Strep ne s'est révélée satisfaisante pour détecter PMTub dans les différentes fractions subcellulaires analysées. Dans le cas de la fusion PMTub-TwinStrep, un signal spécifique très faible a pu être détecté, mais uniquement dans la fraction censée correspondre à la paroi bactérienne et non pas dans la fraction correspondant à la membrane plasmique comme attendu (**Figure 21B**).

Afin d'améliorer le signal obtenu avec l'étiquette Twin-Strep et de vérifier ce résultat surprenant, nous avons essayé d'optimiser la méthode de solubilisation de PMTub. En effet PMTub étant une protéine membranaire polytopique très hydrophobe, l'extraction, la solubilisation, la migration sur SDS-PAGE et le transfert en western blot peuvent constituer des étapes limitantes au cours desquelles la protéine peut être perdue. Ainsi, pour éviter une agrégation possible fréquemment observée pour ce type de protéines, nous avons modifié le protocole de dénaturation de la protéine avant migration sur SDS-PAGE. Basé sur ce qui est décrit pour la détection des *Sc*PMT par western blot[139], la dénaturation des échantillons dans le tampon Laemmli a été réalisée non pas à 100°C comme habituellement, mais à plus faible température (50 °C) afin d'éviter une exposition trop rapide des domaines transmembranaires hydrophobes pouvant conduire à une agrégation de ces protéines polytopiques. Ce simple changement de protocole a grandement amélioré le signal de la bande spécifique dans la fraction contenant la paroi bactérienne (**Figure 21B**).

Des bandes spécifiques des protéines fusionnées aux étiquettes Strep, Twin-Strep et HA, ont pu ainsi être détectées, mais toujours dans les fractions censées correspondre aux constituants de la paroi bactérienne (**Figure 21C**). De plus, les bandes observées présentent des migrations électrophorétiques correspondant à des composés des poids moléculaires apparents de l'ordre 40 kDa qui semblent incompatibles avec la masse attendue de 55 kDa de PMTub. Toutefois, ce comportement n'est pas totalement inhabituel et a déjà été décrit dans le cas les protéines hautement hydrophobes, qui ont tendance à lier une grande quantité de SDS, conduisant à une migration aberrante[140].

Finalement ces défauts de migration et de localisation nous ont incités à vérifier si les bandes spécifiques détectées correspondaient réellement à PMTub.

Nous avons choisi pour cela de nous concentrer sur la construction comportant l'étiquette HA pour laquelle nous avons observé le signal le plus intense. Pour confirmer que

la bande spécifique correspondait bien à PMTub, nous avons effectué une analyse ESI-MS/MS après digestion tryptique *in-gel* des protéines et extraction des peptides. Deux bandes de gel de la fraction « paroi bactérienne », correspondant à des composées migrant respectivement autour poids moléculaire observé en immunoblot (30 à 50 kD) et du poids moléculaire attendu de PMTub (50 à 70 kD) ont été analysées en parallèle. Un seul peptide pouvant correspondre à PMTub a pu être détecté, dans la bande « poids moléculaire observé » uniquement. Toutefois, le faible score d'identification de ce peptide et le fait qu'il ne soit répertorié dans aucune des banques de données protéomiques de *M. tuberculosis*[141–143] rendent cette identification relativement fragile. Par conséquent, l'analyse par spectrométrie de masse n'exclut pas que la bande observée dans les western blot et présentant une migration aberrante puisse correspondre à PMTub, mais sans pour autant le prouver de manière indiscutable.

Finalement, quelle que soit l'étiquette testée, l'ensemble de ces observations révèle de façon inattendue la difficulté de détecter la protéine étiquetée par western-blot et remet en question la validité de l'utilisation de cette approche pour mettre en évidence l'expression de PMTub.

Plusieurs autres problèmes rencontrés au cours de ces essais peuvent également être notés. Tout d'abord, comme on peut le voir sur les différents western blots **Figure 21C**, de nombreux signaux non spécifiques sont également détectés. Aussi, au cours des expérimentations nous avons été confrontés à une forte instabilité du signal. En effet, une forte fluctuation de l'intensité du signal d'une expérimentation à l'autre et une dégradation de ce signal après chaque décongélation de l'échantillon laisse présager d'une faible stabilité de PMTub dans le tampon utilisé et une forte variabilité de la solubilisation de PMTub d'une expérience à l'autre.

Pour résoudre ces problèmes, nous avons essayé d'améliorer le signal spécifique par extension de l'étiquette HA en une étiquette triple HA, comme cela a été décrit dans le cas des PMT de levure[53,131]. Par ailleurs, dans le but d'optimiser l'extraction et la stabilité du PMTub, nous avons évalué différentes méthodes de lyse et tenté d'utiliser des tampons contenant du détergent (SDS) et/ou un agent dénaturant (urée) pour solubiliser les composants de la fraction membranaire. Nous avons également pris soin d'évité de congeler-

décongeler l'échantillon en effectuant l'ensemble des étapes du protocole en continu dans la même journée.

Malheureusement, malgré de multiples tentatives, nous ne sommes jamais parvenus à obtenir par cette approche, des résultats incontestables et reproductibles permettant de conclure à la présence de l'enzyme PMTub à la membrane. En revanche, quelle que soit l'étiquette fusionnée de PMTub, nous avons pu vérifier le maintien de l'activité enzymatique confirmant l'expression de l'enzyme fonctionnelle.

Devant ces résultats incohérents, nous nous sommes interrogés sur la pertinence et l'incompatibilité éventuelle de la localisation de l'étiquette à l'extrémité carboxy-terminale de la protéine.

Pour vérifier ce point, une souche exprimant PMTub marquée par une étiquette histidine a son extrémité N-terminale a été construite. L'analyse de cette souche a été effectuée comme précédemment et n'a pas permis non plus de mettre en évidence de bande spécifique à la condition étiquetée (données non montrées). La difficulté de détection de PMTub n'est donc pas liée au positionnement du marquage, mais semble inhérente à la protéine et à la méthode utilisée.

En conclusion, PMTub est une protéine hautement hydrophobe que nous soupçonnons être très faiblement exprimée dans la cellule malgré l'utilisation d'un promoteur constitutif fort. Les résultats obtenus suggèrent que l'enzyme est difficile à extraire et à maintenir en quantité suffisante pour être détectée en western blot, quel que soit le positionnement du marquage. En effet, la forte hydrophobicité de la protéine peut entraver chacune des étapes du protocole conduisant à un signal faible, voire pas de signal, lors de la révélation. Il apparait donc clairement de ces observations et des résultats négatifs ou incohérents obtenus jusqu'ici que l'approche western blot ne semble pas des plus adaptée pour détecter l'expression de PMTub *ex-cellulo*.

Devant cet échec, nous avons donc décidé de privilégier une approche phénotypique pour détecter de l'expression de PMTub qui permet de s'affranchir de toutes les difficultés de

préparation et de manipulation de l'échantillon propre aux méthodes de détection moléculaire par western-blot.

2. Détection de l'expression de PMTub sur bactéries

Pour détecter l'expression de PMTub dans son intégralité au niveau de la membrane plasmique, nous avons décidé de fusionner des protéines détectables *in-situ* sur la bactérie à l'extrémité C-terminale de PMTub.

a) Protéine fluorescente

Notre premier choix s'est porté sur l'utilisation d'une protéine fluorescente qui devrait permettre de confirmer la localisation de l'enzyme au niveau de la membrane plasmique par microscopie à fluorescence, mais aussi de quantifier l'expression de la protéine de fusion par spectrofluorimétrie.

Le principal défi de cette approche est la difficulté de trouver une protéine fluorescente fonctionnelle dans le périplasme bactérien. En effet, la littérature sur les PMT[88,131] et la modélisation 3D de PMTub (voir partie2.A, p.101) indique une topologie membranaire de type trans avec les extrémités N- et C-terminales localisés de part et d'autre de la membrane dans des compartiments cellulaires opposés. Dans le cas de la protéine bactérienne PMTub, les extrémités N- et C-terminales sont prédites pour être respectivement localisées dans le cytoplasme et dans le périplasme extracellulaire. Or, le périplasme constitue un environnement oxydant favorisant la formation de ponts disulfures susceptibles d'altérer la conformation fonctionnelle et par là même la fluorescence de protéines comme la GFP[144]. La détection de l'extrémité C-terminale de PMTub par fusion avec une protéine fluorescente constitue donc un certain défi puisque celle-ci va se retrouver dans un environnement périplasmique défavorable à sa fluorescence.

Des études d'optimisation ont permis de mettre au point des protéines fluorescentes adaptées à ce type de milieu[145]. Cette adaptation résulte en partie par l'absence de cystéines dans la séquence de ces protéines ce qui permet d'éviter la formation de ponts disulfures, et d'obtenir un repliement correct et fonctionnel dans l'environnement oxydant du périplasme.

Ainsi, il a été démontré que mCherry, dont la séquence est dépourvue de cystéines, est fluorescente dans le périplasme de *E. coli*[146].

Bien que celle-ci n'ait encore jamais été utilisée dans le contexte périplasmique chez la mycobactérie, nous avons choisi de l'utiliser pour tenter de détecter et de localiser *in-situ* la protéine de fusion PMTub-mCherry par imagerie de super-résolution en collaboration avec la plateforme TRI du réseau d'imagerie de la Génopole de Toulouse (TRI-GENOTOUL).

Pour comparer l'impact de l'environnement subcellulaire sur l'expression et l'intensité de fluorescence de la protéine mCherry, nous avons réalisé deux constructions en fusion à l'extrémité N-terminale (mCherry-Nter-PMTub) et C-terminale (PMTub-Cter-mCherry) de PMTub respectivement localisées coté cytoplasmique et périplasmique.

(1) Optimisation des conditions d'observation

Nous avons tout d'abord voulu mettre au point les techniques d'observation par microscopie de super-résolution. Pour cela, différents protocoles de préparation des bactéries ont été testés sur une souche de *M. smegmatis* exprimant la construction avec mCherry à l'extrémité N-terminale de PMTub. La préparation des échantillons peut avoir un effet sur l'image obtenue en microscopie de super-résolution. Nous avons donc comparé la visualisation sur bactéries fixées avec du paraformaldéhyde ou sur bactéries vivantes immobilisées par pad d'agarose. La fixation chimique mène à la formation de « hotspot » de fluorescence absents des images sur bactéries vivantes immobilisées (**Figure 22A**). Nous avons donc décidé de favoriser l'expérimentation sur bactéries vivantes pour l'ensemble des observations suivantes.

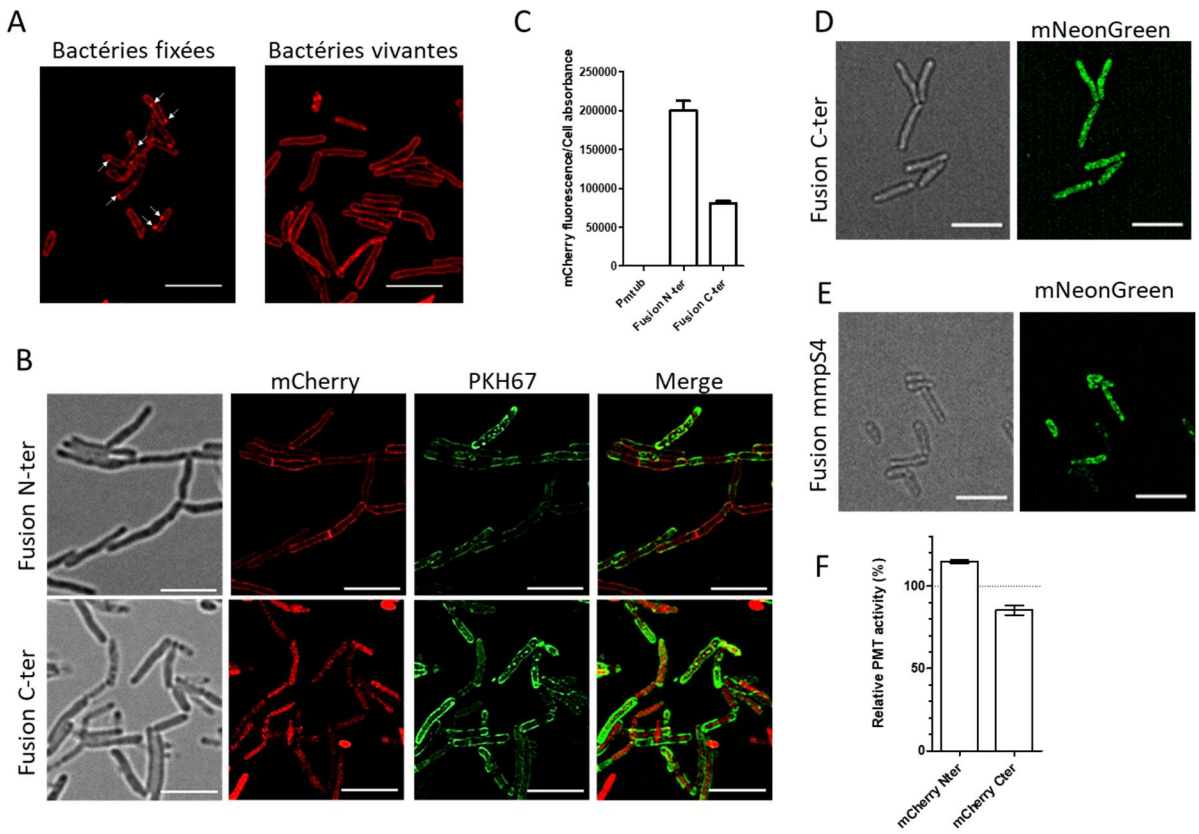


Figure 22. Visualisation de PMTub par microscopie à fluorescence. A) Image de microscopie à illumination structurée sur bactéries exprimant PMTub en fusion avec mCherry à l'extrémité N-terminale, fixées par le paraformaldéhyde (image de gauche) ou vivantes (image de droite). Les flèches blanches montrent les « hotspot » de fluorescence artefactuels. Barre d'échelle = 5µm. **B)** Image de microscopie à illumination structurée sur bactéries vivantes. Les bactéries expriment la protéine de fusion PMTub-mCherry avec mCherry à l'extrémité C-terminale (périplasmique) ou N-terminal (cytoplasmique) de PMTub. Les bactéries sont colorées au PKH (marqueur de membrane, vert). **C)** Mesure d'intensité de fluorescence par spectrofluorimétrie de bactéries exprimant la protéine de fusion PMTub-mCherry avec mCherry à l'extrémité C-terminale (périplasmique) ou N-terminal (cytoplasmique) de PMTub, normalisée par la quantité de bactéries. **D)** Image de microscopie à illumination structurée sur bactéries vivantes. Les bactéries expriment la protéine de fusion PMTub-mNeonGreen avec mNeonGreen à l'extrémité C-terminale. Barre d'échelle = 5µm. **E)** Image de microscopie à illumination structurée sur bactéries vivantes exprimant la protéine de fusion mmpS4-mNeonGreen. Barre d'échelle = 5µm**. F)** Activité relative *in vivo* de PMTub en fusion avec mCherry à son extrémité N- ou C-terminale. 0% et 100% sont définis respectivement comme la mannosylation mesurée pour la souche No PMT et la souche exprimant PMTub native. Les résultats de triplicats techniques de duplicats biologiques sont montrés.

(2) Détection et localisation de PMTub-mCherry

Pour préciser plus avant la localisation sur les bactéries de la fluorescence de la fusion mCherry-Nter-PMTub, un co-marquage des membranes bactériennes a été réalisé en parallèle par la sonde fluorescente PKH67, sonde hydrophobe qui marque spécifiquement les structures membranaires. Les images de microscopie haute résolution obtenues montrent nettement une colocalisation des deux sondes mCherry (rouge) et PKH67 (vert) démontrant formellement pour la première fois la localisation membranaire de PMTub (**Figure 22B, panels du hauts**). Il est à noter que l'on observe une distribution homogène de PMTub sur l'ensemble de la membrane bactérienne dans les conditions de l'expérience, c'est-à-dire lorsque PMTub est exprimée sous contrôle d'un promoteur constitutif fort.

Dans un deuxième temps, nous avons voulu comparer les images de fluorescence obtenues lorsque mCherry est localisée aux différentes extrémités de PMTub. Contrairement à ce que nous avons pu observer dans le cas de la fusion mCherry-Nter-PMTub, la fusion mCherry à l'extrémité C-terminale de PMTub se traduit par un marquage diffus distribué sur la bactérie entière (**Figure 22B, panels du bas**).

Ces images de super-résolution mettent en évidence une différence significative selon que la protéine fluorescente est fusionnée à l'extrémité N- ou C-terminale de la protéine PMTub .

Une des raisons pouvant expliquer ces observations serait que dans ces conditions expérimentales, la fluorescence détectée pour la fusion PMTub-Cter-mCherry serait trop faible et insuffisante pour déterminer avec une résolution suffisante la localisation de la protéine fluorescente ; on atteindrait dans ce cas les limites de détection l'appareil d'imagerie. Pour vérifier cette hypothèse, nous avons comparé par spectrofluorimétrie la fluorescence des bactéries. Comme le montre la figure 2C, la quantification de la fluorescence intrinsèque des bactéries confirme la présence d'une différence significative de fluorescence de près d'un facteur 3 entre les bactéries exprimant la construction mCherry-Nter-PMTub de celles exprimant PMTub-Cter-mCherry, qui pourrait donc être à l'origine des différences observées en microscopie haute résolution(**Figure 22B et C**).

Enfin comme ultime contrôle nous avons voulu vérifier si ces différences d'intensité de fluorescence intrinsèque des bactéries pouvaient résulter de différences d'expression des protéines de fusion selon que mCherry soit positionnée en amont ou en aval de la séquence

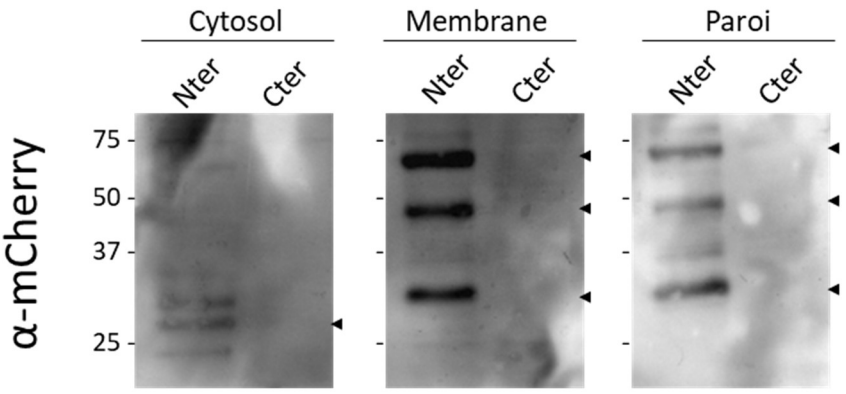


Figure 23. Détection de PMTub-mCherry par western blot. Western blot anti-mCherry après fractionnement sur des souches de *M. smegmatis* exprimant la protéine de fusion PMTub-mCherry avec mCherry située à l'extrémité C-terminale (périplasmique) ou N-terminale (cytoplasmique) de PMTub.

de PMTub. En effet on peut raisonnablement imaginer que la fusion de mCherry à l'extrémité C-terminale de PMTub puisse impacter le repliement de la protéine de fusion entrainant une localisation aberrante et une dégradation partielle par les systèmes cellulaires de contrôle qualité des protéines.

Pour cela les 2 souches exprimant respectivement les protéines de fusion mCherry-Nter-PMTub et PMTub-Cter-mCherry ont été transformées pour exprimer également la protéine cible FasC dont nous avons mesuré le taux de mannosylation relatif après 7 jours de culture. Les résultats obtenus montrent que les 2 constructions présentent des niveaux d'activités de mannosylation de la FasC comparables à celui de la protéine PMTub native (**Figure 22F**). Ces données indiquent clairement que, quelle que soit la configuration de la fusion, la présence de mCherry ne semble pas affecter l'expression ni la localisation des protéines de fusion fonctionnelles à la membrane.

Il apparait donc que les différences d'intensité de fluorescence observées que ce soit en microscopie haute résolution ou en spectrofluorimétrie ne peuvent pas être imputées à des différences d'expression des protéines fusion, mais bien liées aux différences de localisation subcellulaire de la protéine fluorescente dans le cytoplasme ou le périplasme selon que mCherry soit fusionnée en amont ou en aval de la séquence de PMTub.

Cependant, il peut être noté que la fusion à l'extrémité C-terminale entraine une légère baisse de la mannosylation témoignant d'un possible problème de repliement à la membrane d'une partie de la protéine de fusion. Il ne peut donc pas être totalement exclu que cette configuration puisse conduire à une topologie aberrante se traduisant par une dégradation protéolytique d'une fraction des protéines à l'origine de la présence de forme tronquée fluorescente de la protéine de fusion dans le compartiment cytoplasmique. Ceci pourrait expliquer la localisation cytoplasmique apparente de la fluorescence observée par microscopie avec la construction PMTub-Cter-mCherry.

Pour tester cette hypothèse nous avons réalisé des western blot avec un anticorps anti-mCherry sur les constructions exprimant mCherry en fusion avec PMTub. Le western blot a été réalisé après fractionnement subcellulaire des extraits protéiques pour déterminer la présence de mCherry à la membrane ou dans le cytoplasme. Les résultats du western blot ne permettent pas de détecter mCherry dans les fractions cytoplasmiques pour la fusion C-terminale, écartant l'hypothèse d'un clivage et d'un relargage cytoplasmique de mCherry pour

Tableau 4. Liste de peptides de mCherry-Nter-PMTub détectés par spectrométrie de masse. Les peptides dichargés détectés sont rapporté ici. Ils permettent une couverture de séquence de la position 80 à 550 pour chaque bande sauf la bande à 82kD (taille de la protéine de fusion entière) qui a une couverture de 80 à 297. En rouge : peptides de mCherry ; en vert : peptides de PMTub.

sequence	psm_score	calculated_mass	peptide_length	start	end	Bande
HPADIPDYLK	45,09	1167,59241	10	80	89	32
HPADIPDYLK	76,91	1167,59241	10	80	89	32
HPADIPDYLK	76,02	1167,59241	10	80	89	45
HPADIPDYLK	55,4	1167,59241	10	80	89	45
HPADIPDYLK	73,01	1167,59241	10	80	89	62
HPADIPDYLK	55,55	1167,59241	10	80	89	82
LSFPEGFK	21,92	923,475266	8	90	97	32
LSFPEGFK	23,27	923,475266	8	90	97	45
LSFPEGFK	35,88	923,475266	8	90	97	62
LSFPEGFK	19,12	923,475266	8	90	97	82
GTNFPSDGPVMQK	47,55	1376,63947	13	131	143	32
GTNFPSDGPVMQK	50,98	1376,63947	13	131	143	32
GTNFPSDGPVMQK	41,45	1376,63947	13	131	143	45
GTNFPSDGPVMQK	52,35	1376,63947	13	131	143	62
GTNFPSDGPVMQK	19,11	1376,63947	13	131	143	82
TMGWEASSER	56,13	1152,48695	10	145	154	32
TMGWEASSER	46,37	1152,48695	10	145	154	45
TMGWEASSER	36,67	1152,48695	10	145	154	62
MYPEDGALK	25,52	1022,47426	9	155	163	32
MYPEDGALKGEIK	48,85	1449,71733	13	155	167	32
MYPEDGALKGEIK	24,94	1449,71733	13	155	167	45
MYPEDGALKGEIK	45,1	1449,71733	13	155	167	62
LKDGGHYDAEVK	52,09	1330,65172	12	172	183	32
LKDGGHYDAEVK	50,58	1330,65172	12	172	183	45
KPVQLPGAYNVNIK	74,99	1539,87729	14	190	203	32
KPVQLPGAYNVNIK	58,87	1539,87729	14	190	203	62
LDITSHNEDYTIVEQYER	82,21	2224,02838	18	204	221	32
LDITSHNEDYTIVEQYER	100,65	2224,02838	18	204	221	45
LDITSHNEDYTIVEQYER	87,34	2224,02838	18	204	221	62
HSTGGMDELYK	27,43	1236,54448	11	226	236	32
HSTGGMDELYK	46,18	1236,54448	11	226	236	45
HSTGGMDELYK	39,11	1236,54448	11	226	236	62
GSSGSVPVVSPGPLVPVADFGPLDR	66,9	2405,25913	25	237	261	32
GSSGSVPVVSPGPLVPVADFGPLDR	67,84	2405,25913	25	237	261	45
GSSGSVPVVSPGPLVPVADFGPLDR	78,65	2405,25913	25	237	261	62
GSSGSVPVVSPGPLVPVADFGPLDR	54,26	2405,25913	25	237	261	82
FLNLGSLTDAGTPIFDEK	99,77	1936,97821	18	280	297	32
FLNLGSLTDAGTPIFDEK	113,47	1936,97821	18	280	297	45
FLNLGSLTDAGTPIFDEK	85,86	1936,97821	18	280	297	62
FLNLGSLTDAGTPIFDEK	64,07	1936,97821	18	280	297	82
MHIALLAGR	18,74	980,558929	9	422	430	32
MHIALLAGR	16,03	980,558929	9	422	430	45
MHIALLAGR	41,54	980,558929	9	422	430	62
SAATVWGPR	65,02	943,487565	9	431	439	62
PWLGTVR	20,26	827,465373	7	490	496	45
PWLGTVR	21,33	827,465373	7	490	496	62
NSVVPLPDAVR	26,05	1165,64552	11	540	550	32
NSVVPLPDAVR	22,16	1165,64552	11	540	550	45
NSVVPLPDAVR	45,71	1165,64552	11	540	550	62

la construction. Toujours dans le cas de la fusion PMTub-Cter-mCherry, la protéine mCherry n'est pas non plus détectée dans les fractions correspondant aux composés membranaires ou pariétaux. En revanche, plusieurs bandes spécifiques peuvent être détectées pour la fusion à mCherry-Nter-PMTub dans les fractions dites « non solubles » des membrane et parois (**Figure 23**). Plus précisément, trois bandes migrant à 32, 45 et 62kD sont détectées. Aucune de ces tailles ne correspondant à la taille attendue de 82kD (55kDa pour PMTub + 27kDa pour mCherry) pour la protéine recombinante mCherry-Nter-PMTub.

Pour vérifier que ces bandes correspondent bien à mCherry et déterminer si PMTub est toujours en fusion avec la protéine fluorescente nous avons effectué une analyse par spectrométrie de masse. L'analyse a été effectuée sur les trois bandes de gels détectées par western blot à 32, 45 et 62kDa ainsi que sur une bande de gel correspondant à la taille attendue pour la protéine de fusion entière (82kDa). Cette analyse a permis de confirmer que les bandes observées correspondaient en effet à la protéine de fusion mCherry-Nter-PMTub. De nombreux peptides ont pu être identifiés avec un haut score en spectrométrie de masse. Toutefois la couverture de séquence obtenue pour ces 3 bandes bien que sensiblement différentes ne permet pas de confirmer que ces bandes pourraient correspondre à des formes tronquées de taille différente de la protéine de fusion (**Tableau 4**). Pour la zone correspondant à la taille attendue de la protéine de fusion entière, des peptides correspondants à mCherry-PMTub sont aussi identifiés, mais en un nombre plus restreint et avec un score plus faible (**Tableau 4**).

Bien que l'identification précise des séquences protéiques à l'origine de ces différentes bandes reste à faire, ces expériences démontrent de façon inattendue et intéressante que l'ajout d'une protéine fluorescente à l'extrémité N-terminale de PMTub permet sa détection en western-blot après fractionnement. Ainsi la fusion de mCherry, protéine soluble de 27 kDa à PMTub, protéine membranaire hydrophobe de 55kDa, aide peut être à la solubilisation de cette dernière. Il serait donc possible que la fusion à l'extrémité N-terminale d'une grosse protéine soluble permette de stabiliser PMTub pendant son extraction. Ceci permet d'envisager une possible purification de PMTub ouvrant la porte à des études d'interactomique par co-immunoprécipitations, qui permettraient de mieux définir l'environnement cellulaire et les partenaires moléculaires de PMTub. Ces résultats sont encourageants pour la mise en place de nouvelles stratégies d'étude de PMTub dans le futur.

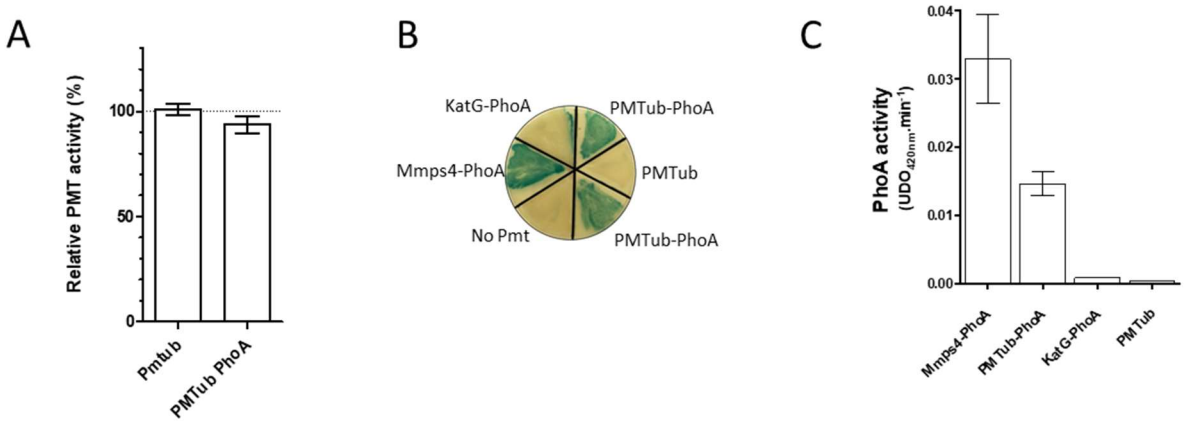


Figure 24. Contrôle de l'expression de PMTub en utilisant la phosphatase alcaline. A) Activité relative *in vivo* de PMTub en fusion avec PhoA. 0% et 100% sont définis respectivement comme la mannosylation mesurée pour la souche No PMT et la souche exprimant PMTub native. Les résultats de triplicats techniques de duplicats biologiques sont présentés. **B)** Etalement des différentes souches sur boite contenant du BCIP permettant de détecter l'activité PhoA par coloration des souches exprimant une PhoA fonctionnelle en bleu. **C)** Activité phosphatase alcaline de la bactérie *M. smegmatis* « No PMT » complémentée avec Mmps4-PhoA (contrôle périplasmique), PMTub-PhoA, KatG-PhoA (contrôle cytoplasmique) ou PMTub non marqué. La formation du produit de la réaction catalysée par PhoA est suivie par l'augmentation de l'absorbance à 420 nm au cours du temps, normalisée par la quantité de bactéries.

(3) Recherche de marqueurs fluorescents de l'expression « pleine taille » de PMTub.

Comme nous le verrons par la suite, il sera important au cours de notre projet de pouvoir contrôler l'expression « pleine taille » de PMTub c'est dire de la protéine entière fonctionnelle jusqu'à son extrémité C-terminale localisée dans le périplasme. Dans ce cadre, le faible rendement de fluorescence de mCherry dans le périplasme s'étant avéré insuffisant pour vérifier ce point, nous nous sommes tournés vers d'autres protéines fluorescentes qui ont été démontrées comme fonctionnelles dans le périplasme de *E.coli*[146,147]. Nous avons ainsi testé les protéines Superfolder mTurquoise2ox (SfTq2ox) et mNeonGreen en fusion à l'extrémité C-terminale de PMTub. Malheureusement, l'intensité de fluorescence de ces 2 protéines s'est révélée encore moindre que celle de mCherry, allant jusqu'à être indétectable dans le cas de SfTq2ox (**Figure 22D**).

Pour déterminer si la faible fluorescence détectée dans le périplasme est due à un impact de l'environnement périplasmique sur ces protéines fluorescentes seules ou si la fusion à PMTub a aussi un rôle, nous avons fusionné mNeonGreen à l'extrémité périplasmique d'une autre protéine membranaire mycobactérienne : mmpS4[148]. Le signal détecté pour cette construction est meilleur, en microscopie super résolution, que ce qui avait pu être mesuré pour la fusion avec PMTub (**Figure 22D et E**). Il semblerait donc que l'environnement périplasmique impacte l'intensité de fluorescence de ces protéines, mais que la fusion à l'extrémité C-terminale de PMTub a aussi un effet.

En conclusion, l'utilisation de protéines de fusion PMTub-protéine fluorescente permet de détecter l'expression de l'enzyme, mais nous n'avons pas pu trouver de condition permettant d'utiliser cette approche pour le contrôle de l'expression de la protéine pleine taille. Nous nous sommes donc tournés vers une autre stratégie pour détecter l'expression de PMTub pleine taille à la membrane et avec une topologie correcte. Cette nouvelle approche repose sur la fusion traductionnelle de la phosphatase alcaline (PhoA) à l'extrémité C-terminale de PMTub.

b) Phosphatase alcaline

PhoA est une enzyme active uniquement lorsqu'elle est exprimée dans le périplasme et détectable par des tests phénotypiques[149,150]. PhoA catalyse le clivage de fonctions phosphodiester. Cette enzyme est capable de catalyser cette réaction seulement après son export dans le périplasme[151]. En effet, pour adopter sa conformation finale et être fonctionnelle, la formation de deux ponts disulfures est nécessaire, ceux-ci étant formés dans le périplasme. Alors que la capacité des protéines à former des ponts disulfures dans le périplasme était une limitation à l'utilisation de protéines fluorescentes dans ce compartiment bactérien, c'est ici la caractéristique qui permet d'utiliser la phosphatase alcaline comme protéine rapportrice.

L'utilisation de la phosphatase alcaline comme gène rapporteur chez les mycobactéries est possible du fait de l'absence d'homologue chez ces bactéries : jusqu'à présent, aucun gène codant pour une telle enzyme et aucune activité phosphatase alcaline n'ont pu être identifiés chez *M. smegmatis*, notre organisme modèle. De plus, le test d'activité PhoA peut être quantitatif, et une telle approche est déjà utilisée chez les mycobactéries pour quantifier l'expression de protéines dans le périplasme[152]. En plus de l'approche quantitative, une lecture phénotypique blanc/bleu permet aussi d'apprécier l'expression périplasmique de PhoA sur boite de culture.

La fusion de PhoA à l'extrémité C-terminale de PMTub et la détection d'une activité PhoA devrait pouvoir nous renseigner sur l'expression de PMTub avec une topologie de membrane appropriée confirmant la localisation de l'extrémité C-terminale dans le périplasme.

Dans un premier temps nous avons vérifié que la fusion de PhoA à l'extrémité C-terminale de PMTub n'affectait pas l'activité de PMTub. Comme illustré par le graphe A de la Figure 24 l'activité protéine *O*-mannosyltransférase de la protéine de fusion est sensiblement identique à celle de PMTub native (**Figure 24A**). Des constructions similaires ont été réalisées avec deux protéines contrôle : une protéine mycobactérienne cytoplasmique (KatG) et une protéine membranaire avec une extrémité C-terminale périplasmique (mmpS4)[148].

Pour vérifier l'expression périplasmique ou cytoplasmique de la phosphatase alcaline fusionnée aux protéines contrôles mmpS4 et KatG respectivement et pour vérifier la détection d'une activité PhoA après fusion à l'extrémité C-terminale de PMTub, nous avons utilisé une

lecture qualitative sur milieu de culture solide (**Figure 24B**). Dans cette expérience, le milieu de culture est supplémenté en BCIP (5-Bromo-4-chloro-3-indolyl phosphate) qui est oxydé en le colorant bleu 5,5'-dibromo-4,4'-dichloro-indigo par la phosphatase alcaline. Une activité PhoA se traduit par une coloration bleue des colonies sur boites. Les différentes constructions sont exprimées dans la souche *M. smegmatis* ΔPMT et étalées sur milieu complémenté en BCIP. La souche contrôle « No PMT » correspondant à *M. smegmatis* ΔPMT transformée par un plasmide vide contrôle, permet de définir qu'il n'y a pas d'activité PhoA basale détectable dans notre organisme modèle. Après plusieurs jours de culture, une coloration des colonies en bleu peut être observée pour le contrôle positif mmpS4-PhoA et pour la fusion de PhoA à l'extrémité C-terminale de PMTub. Au contraire, les souches exprimant KatG-PhoA et PMTub native forment des colonies blanches. Ces résultats permettent dans un premier temps de confirmer l'utilisation de nos deux témoins positif et négatif pour l'activité PhoA, mais ils permettent aussi de confirmer la localisation périplasmique de l'extrémité C-terminale de PMTub en condition physiologique puisque cette protéine de fusion est active pour le transfert de mannose sur des protéines (**Figure 24A**).

L'utilisation d'un autre substrat, le *p*Npp (phosphate para-nitrophénylphosphate), permet le suivi de la réaction catalysée par PhoA en milieu liquide. Le *p*Npp est transformé en un dérivé jaune par PhoA qui peut être suivi par spectrophotométrie. Des cinétiques enzymatiques peuvent donc être réalisées permettant de mesurer le niveau d'expression de PhoA, et par extension, le niveau d'expression de PMTub. Ce test complémentaire nous permet donc de valider l'utilisation des contrôles KatG et mmpS4 et de confirmer la localisation périplasmique de l'extrémité C-terminale de PMTub (**Figure 24C**).

De plus, ce test en milieu liquide permet une quantification relative de l'expression de PMTub à la membrane.

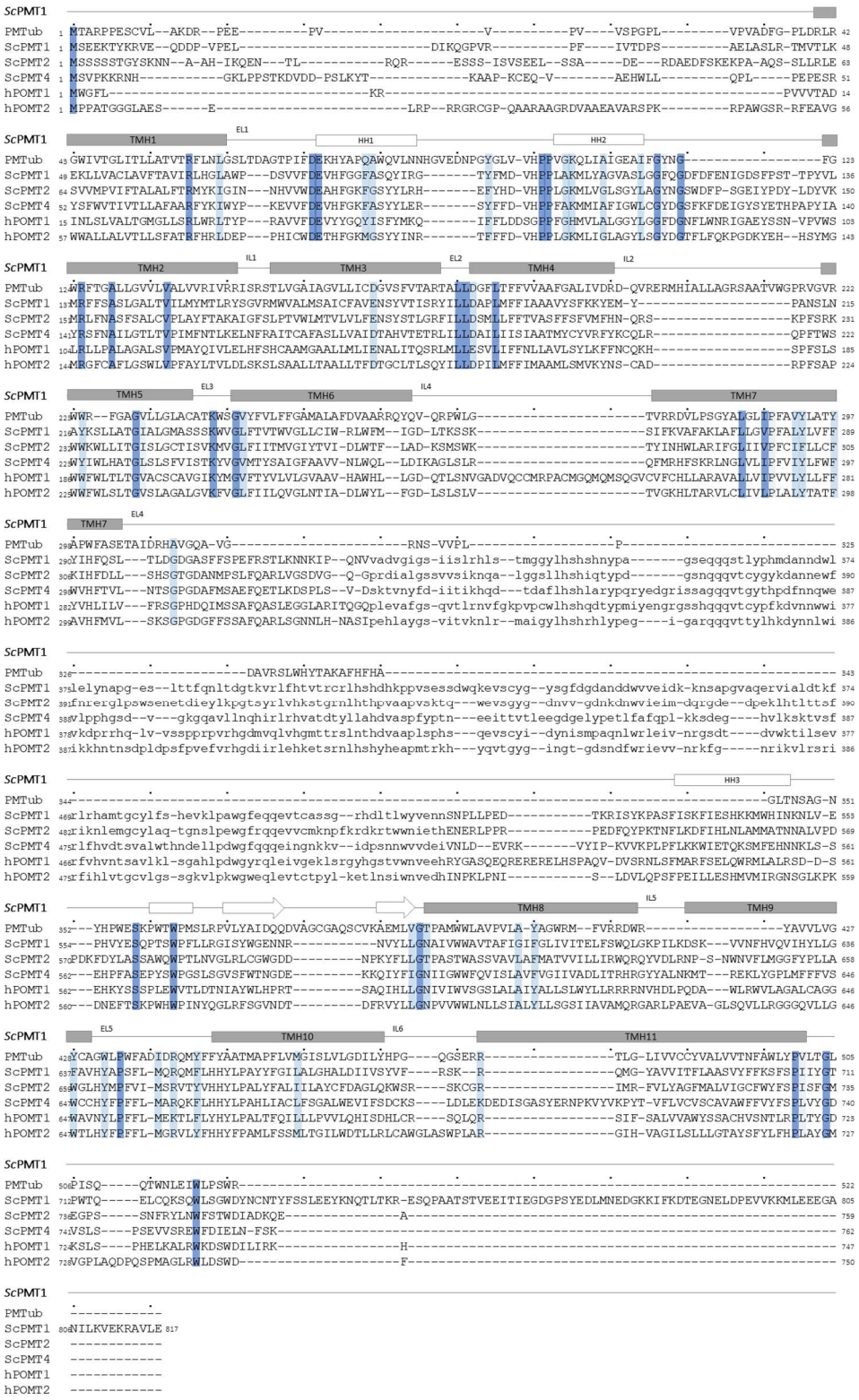


Figure 25. Alignement de séquences de PMT. Alignement de séquence des PMT de différents organismes : *M. tuberculosis* (PMTub), *S. cerevisiae* (ScPMT1, 2 et 4) et l'homme (hPOMT1 et 2). Les structures secondaires de ScPMT1 sont annotées telles qu'identifiées précédemment[88]. Les acides aminés surlignés en bleu foncé sont strictement conservés, les acides aminés surlignés en bleu clair sont conservés à 75% minimum. Les domaines MIR sont annotés par des lettres minuscules et sont définis tels que délimités dans la banque de données « ProRule »[153].

D. Conclusion

En conclusion, cette première partie de mon projet m'a permis de développer un bio-essai phénotypique fiable et robuste permettant d'évaluer l'activité enzymatique de PMTub *in-situ* sur bactérie vivante. Ceci m'a conduit à rechercher et à mettre au point des approches permettant de contrôler l'expression et le repliement correct de la protéine PMTub entière et fonctionnelle à la membrane.

Ces résultats ouvrent la voie vers la suite des travaux prévus dans le cadre de mon projet de thèse, portant sur l'étude des relations structure-fonction de PMTub et le crible d'inhibiteurs de PMTub.

De plus, la mise en place des stratégies de détection de l'expression de PMTub a permis d'apporter des éléments nouveaux pour la caractérisation de l'enzyme. Tout d'abord, la fusion avec une protéine fluorescente et son observation en microscopie de haute résolution ont confirmé formellement la localisation membranaire de PMTub jusqu'alors présumée sur la base d'observations indirectes et d'analogies au système eucaryote[55,75]. Ensuite, l'utilisation de la protéine rapportrice PhoA a permis de mettre en évidence la localisation périplasmique de l'extrémité C-terminale de PMTub. Enfin, la comparaison des résultats de détection de mCherry selon sa fusion aux différentes extrémités de l'enzyme permet quant à elle de confirmer la topologie de type trans de PMTub.

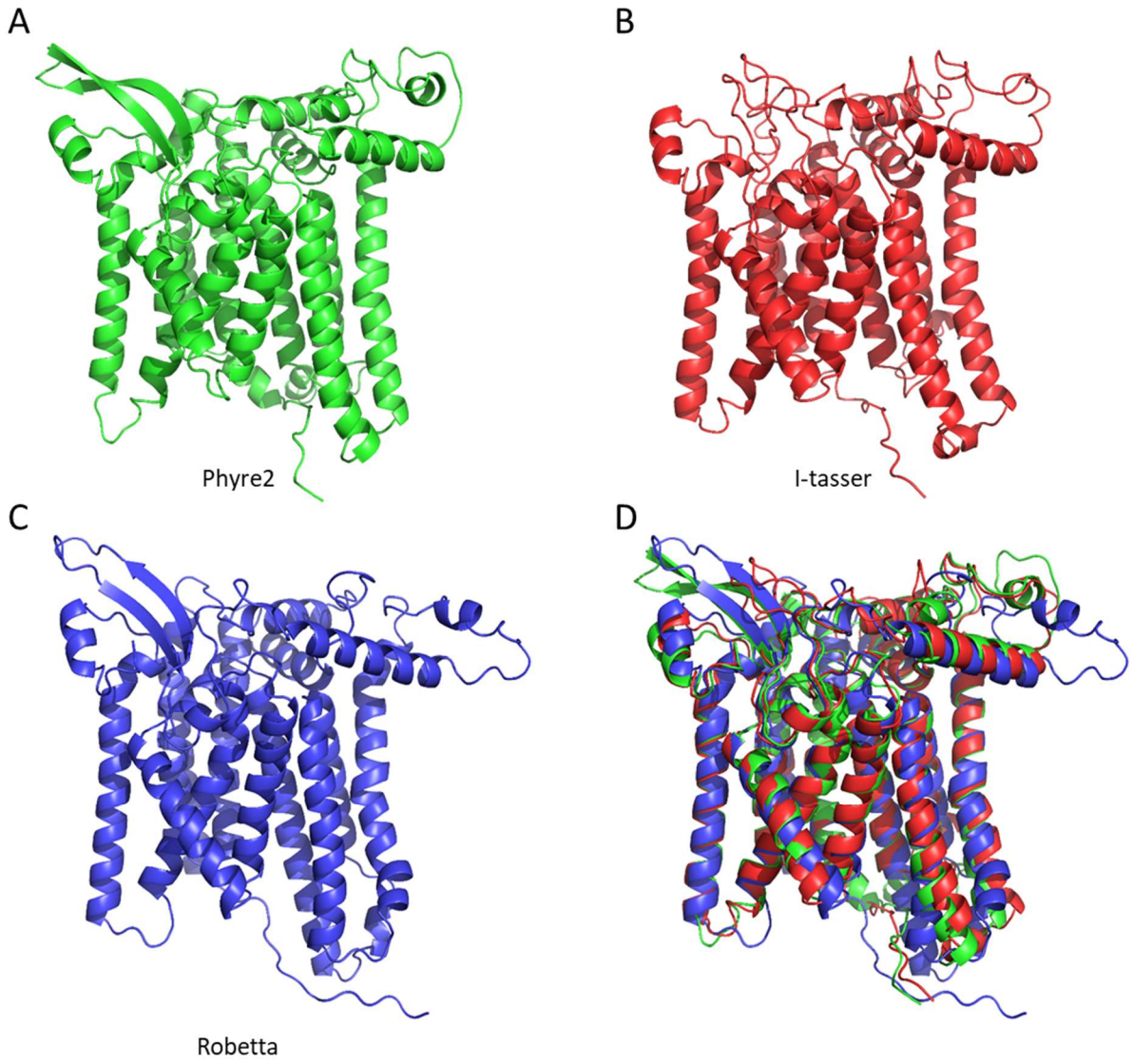


Figure 26. Modélisation 3D par homologie de PMTub. A) Modèle 3D de PMTub obtenu par Phyre2[154]. **B)** Modèle 3D de PMTub obtenu par I-tasser[155–157]. **C)** Modèle 3D de PMTub obtenu par Robetta[158,159]. **D)** Superposition des modèles 3D obtenus par les différents logiciels de modélisation par homologie en conservant les couleurs du reste de la figure.

Partie 2 : Étude des relations structure-fonction de PMTub.

La littérature offre peu de données sur la relation structure/fonction de PMTub. Seul le motif DE a été étudié précédemment, permettant de mettre en avant son caractère essentiel à la fonction de PMTub[55]. Le test phénotypique d'activité développé va permettre d'approfondir nos connaissances sur les relations structure/fonction de PMTub. En effet, dans ce test il est possible d'exprimer des versions mutées de l'enzyme et de mesurer quantitativement l'impact de la mutation sur le transfert de mannose sur la fascicline catalysé par PMTub. La construction d'un modèle 3D de PMTub est également essentielle pour nous permettre d'identifier des résidus dont l'étude fonctionnelle serait intéressante dans le cadre d'une meilleure compréhension des relations structure/fonction de PMTub.

A. Élaboration du modèle 3D de PMTub

Basé sur les précédentes études de la topologie de *Sc*Pmt1[131] et sur le profil hydropathique de PMTub, Vanderven *et al.* ont proposé, en 2005, un modèle topologique de PMTub à 11 domaines transmembranaires[55]. En 2019, la structure 3D de *Sc*PMT1 et *Sc*PMT2 en dimère a été résolue par cryomicroscopie électronique révélant effectivement une topologie à 11 segments transmembranaires pour ces enzymes[88]. Cette architecture correspond aux 11 domaines transmembranaires du cœur structural commun aux GT-C, orné d'un domaine soluble comportant les domaines MIR caractéristiques des PMT eucaryotes. Toutefois, comme signalé en introduction, à l'instar de toutes les PMT bactériennes expérimentalement caractérisées à ce jour, PMTub est dépourvue de domaine MIR[55,59]. En l'absence de ces domaines MIR, PMTub correspondrait donc à une forme fonctionnelle minimale de GT-C dont la structure serait constituée pratiquement exclusivement du cœur structural commun.

1. Méthode

Basée sur ces considérations, la modélisation 3D de PMTub par homologies nous a paru réalisable. La modélisation 3D de protéines est un outil puissant permettant de générer un modèle de structure 3D d'une protéine d'intérêt, ici PMTub, à partir de sa séquence en

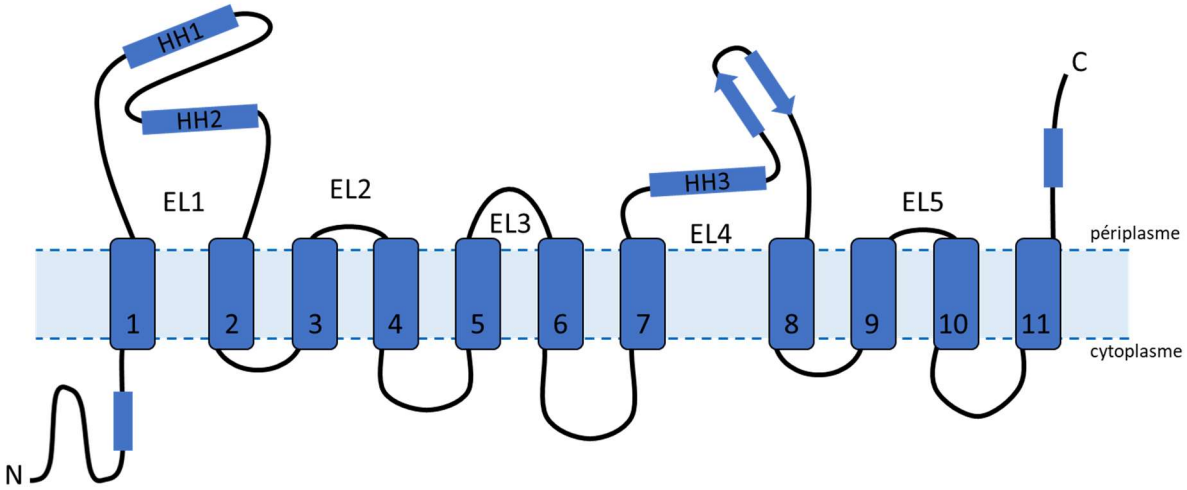


Figure 27. Topologie prédite de PMTub. La topologie prédite de PMTub décrit 11 hélices transmembranaires (TMH), 3 hélices horizontales périplasmiques (HH) et 2 brins bêta. Des hélices alpha sont aussi décrites aux extrémités N terminale et C terminale pour certains des modèles 3D.

acides aminés, et de structures 3D de protéines homologues. Il est généralement considéré qu'une identité de séquence minimale de 20% entre la protéine d'intérêt et les protéines homologues servant de « patron » est nécessaire et suffisante pour obtenir un modèle 3D vraisemblable.

Dans le cas de PMTub, les taux moyens d'identité de séquence respectivement de 27% et 23% avec *Sc*PMT1 et *Sc*PMT2 sont largement acceptables pour espérer générer un modèle 3D fiable. Il est cependant important de noter que les alignements de séquence révèlent des disparités de ces taux moyens avec des zones de très faible identité voire des zones non conservées entre les PMT de levures et PMTub, qui ne peuvent donc pas être modélisées par homologie (**Figure 25**) et nécessitent la mise en œuvre d'approches complémentaires de modélisation *ab initio*, aussi appelée modélisation *de novo*.

Ces approches de prédiction *ab initio* sont moins fiables que la modélisation par homologie, mais permettent de compléter le modèle sur les zones non conservées de la protéine. Dans le cas de PMTub, le modèle peut également être consolidé en utilisant les similarités structurales des GT-C autour de leur socle structural commun.

Plusieurs logiciels de prédiction par homologie de structure 3D de protéines existent et nous avons décidé d'utiliser trois d'entre eux et de comparer les résultats obtenus pour obtenir un modèle 3D fiable de PMTub.

Les trois logiciels de prédictions par homologie utilisés, Phyre2[154], I-tasser [155–157] et Robetta [158,159], se basent sur des méthodologies similaires avec chacun leurs propres algorithmes. La superposition des modèles obtenus pour chacun de ces logiciels permet de mettre en évidence des zones pour lesquelles les prédictions de structures sont similaires et des zones pour lesquelles elles varient (**Figure 26**).

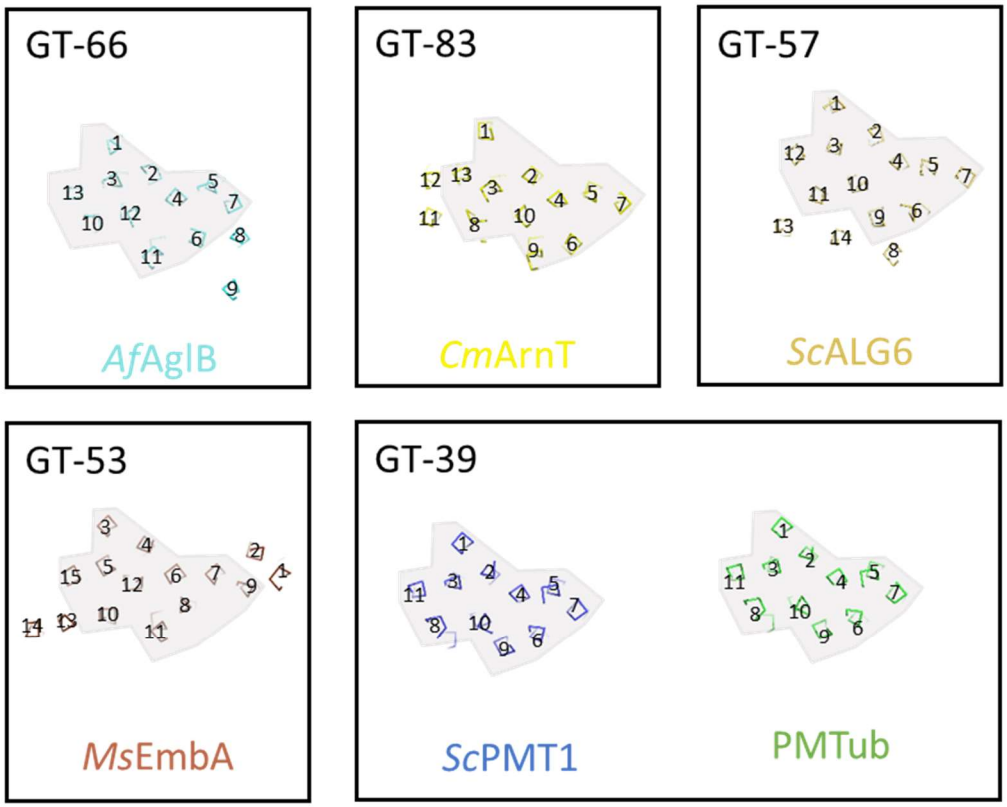


Figure 28. Comparaison du modèle de PMTub au socle commun des GT-Cs. Représentation des GT-Cs des 5 familles CAZy pour lesquelles des structures ont été résolues : GT-53 (*Ms*EmbA[121]), GT-57 (*Sc*ALG6[123]), GT-39 (*Sc*PMT1[88]), GT-66 (*Af*AglB[124]) et GT-83 (*Cm*ArnT[128]), selon une coupe au niveau de la membrane vue de dessus. Représentation similaire pour le modèle 3D de PMTub obtenu par Phyre2. Le module GT-C conservé de 11 domaines transmembranaires est grisé. Les hélices transmembranaires sont numérotées.

2. Modèle topologique de PMTub

De manière cohérente avec ce qui avait été prédit par le passé[55] et avec les structures de *Sc*PMT1-et *Sc*PMT2[88] prioritairement utilisées par les 3 logiciels, les 3 modèles obtenus décrivent une topologie à 11 domaines transmembranaires pour PMTub. Comme attendu, alors que les PMT de levure présentent un important domaine soluble proéminent correspondant aux domaines MIR, les modèles de PMTub sont plus compacts et semblent se réduire à une structure principalement membranaire reposant sur les 11 segments transmembranaires hydrophobes du socle commun des GT-C. On retrouve également à la surface de la protéine présumée au contact de l'espace périplasmique, deux hélices α extramembranaires horizontales HH1 et HH2 localisées dans la première boucle externe (EL1 pour External Loop 1) ainsi qu'une 3^{e} hélice α horizontale (HH3) et deux brins bêta dans la 4^{e} boucle extramembranaire EL4. Enfin les extrémités N- et C-terminales sont prédites pour adopter des conformations en hélices α (**Figure 26 et Figure 27**).

Si l'on compare le modèle de PMTub au socle commun des GT-C, on peut voir que PMTub suit une distribution à la membrane de ses domaines transmembranaires similaire à celui du socle commun des GT-C (**Figure 28**). Les modèles de PMTub décrivent aussi les hélices HH1 et HH2 communes à toutes les GT-C. Il y aurait donc une forte convergence de structure entre PMTub et le socle commun de GT-C. Seule la boucle externe 4 (EL4) semble se distinguer du socle commun par la présence de structures non conservées telles qu'une hélice horizontale (HH3), néanmoins présente sur plusieurs structures de GT-C, et un feuillet β (défini plus en détail en partie 2.E, p.129 et Figure 38).

3. Évaluation de la validité du modèle

Pour comparer et évaluer la fiabilité des modèles fournis par les 3 logiciels utilisés nous avons généré les diagrammes de Ramachandran correspondants. Ces diagrammes permettent de déterminer les conformations énergétiquement favorables d'une protéine en fonction de la valeur des angles dièdres ϕ et ψ des liaisons peptidiques du squelette carboné. Ramachandran *et al.*[160] ont en effet pu établir que les combinaisons énergétiquement admissibles d'angles ϕ et ψ décrivent sur le diagramme 3 zones « favorables » (en rouge sur la **Figure 29**) correspondant aux conformations naturellement adoptées par la majorité des acides aminés.

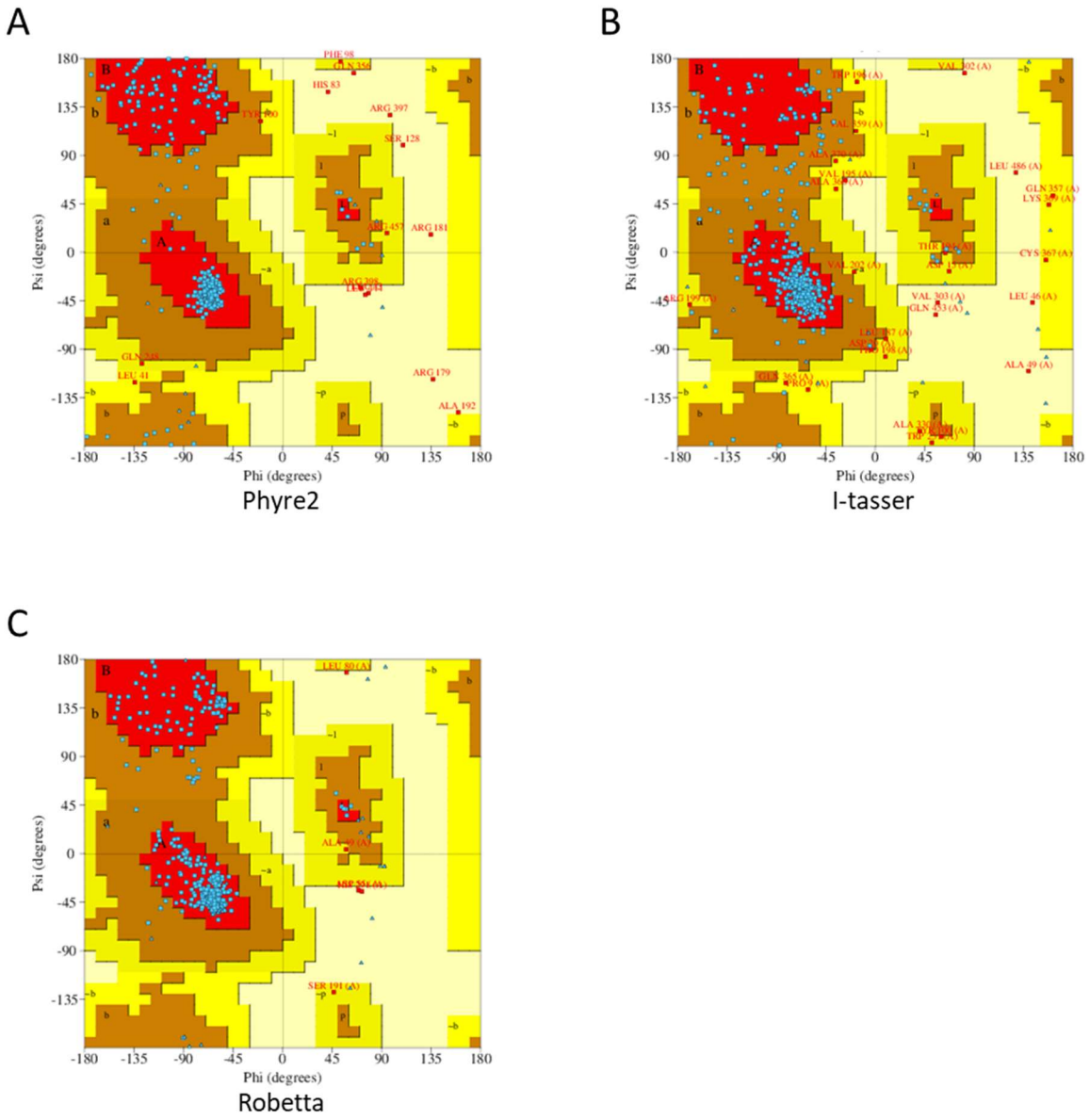


Figure 29. Diagramme de Ramachandran des différents modèles de PMTub. A) Modèle 3D de PMTub obtenu par Phyre2[154]. **B)** Modèle 3D de PMTub obtenu par I-tasser[155–157]. **C)** Modèle 3D de PMTub obtenu par Robetta[158,159].

Pour le modèle obtenu par Phyre2, on compte 98% des acides aminés dans ces espaces « favorables » (**Figure 29A**). Pour I-tasser, seuls 95% des acides aminés adoptent une configuration favorable, avec 23 acides aminés montrant un positionnement questionnable (**Figure 29B**). La modélisation par Robetta fourni le meilleur score avec 99% des acides aminés présentant des géométries de liaison du squelette carboné énergétiquement favorable (**Figure 29C**).

Par la suite, nous sommes donc principalement appuyés sur les modèles Phyre2 ou Robetta qui nous paraissent les plus fiables selon ces critères.

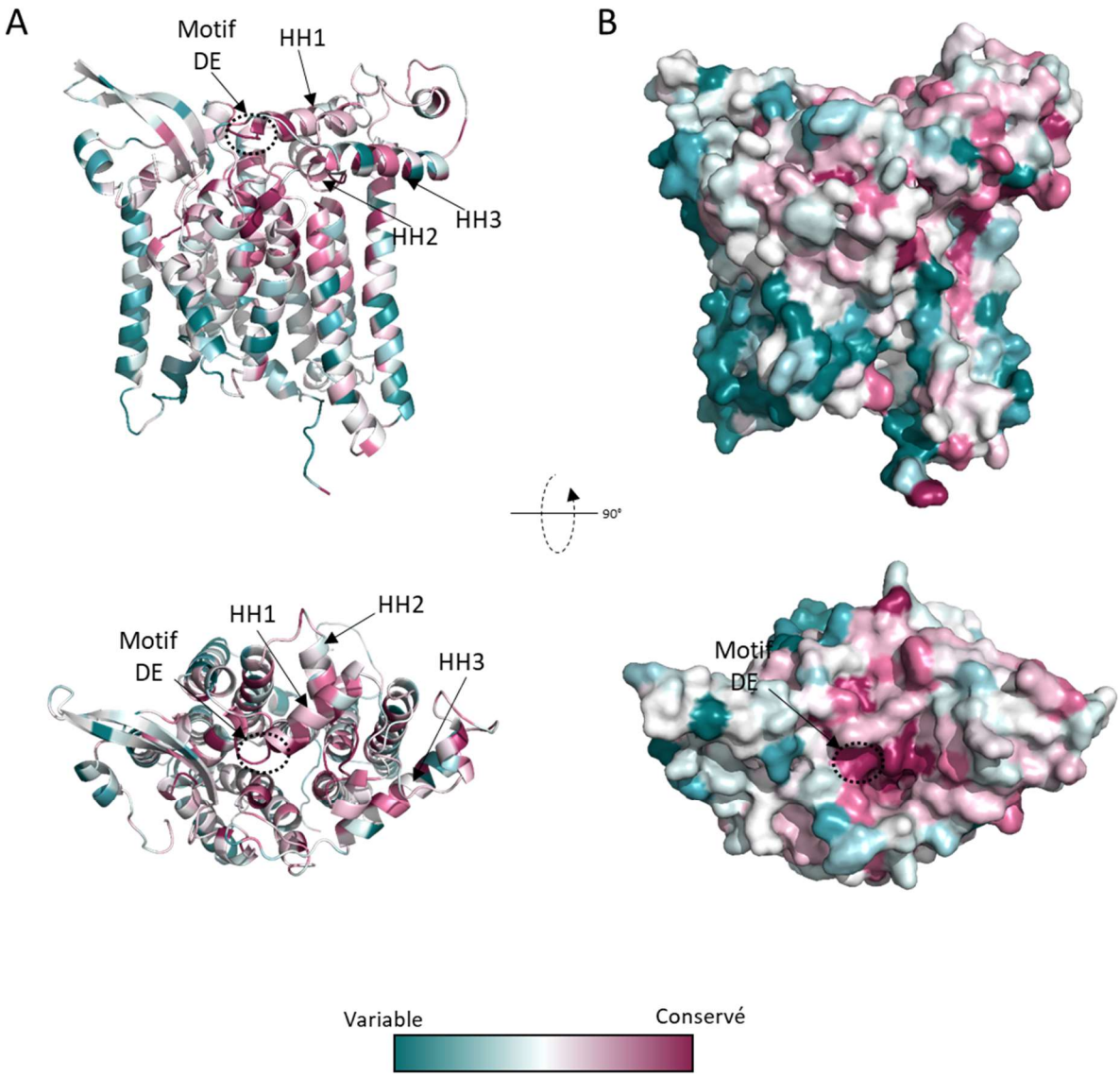
4. Relation structure-fonction : Hypothèses de travail

a) Identification d'une poche catalytique conservée

À l'heure actuelle, une seule étude fonctionnelle a été réalisée sur PMTub qui a permis de mettre en évidence que le motif DE invariant dans EL1 de toutes les PMT est essentiel à l'activité de PMTub[55]. Toutefois chez la levure, outre le motif DE acide, d'autres acides aminés composant le site actif de *Sc*PMT1 et *Sc*PMT2 ont été structurellement identifiés dans les boucles externes[88].

La résolution des structures de *Sc*PMT1 et *Sc*PMT2 a permis de mettre en évidence que l'extrémité glycannique polaire du substrat donneur ainsi que le substrat accepteur interagissent avec des structures protéiques de la face intra-luminale du complexe, correspondant à une localisation périplasmique chez la mycobactérie. Cette observation concorde avec les modèles 3D de PMTub qui font état d'un positionnement du motif invariant DE sur la face présumée périplasmique de la protéine. En outre une analyse du taux de conservation de séquence au sein de la famille des PMT révèle la présence d'une zone autour de la dyade catalytique DE présentant une forte conservation des acides aminés comme illustré par la Figure 30 (**Figure** 30**A**, zone en violet).

Cette zone de forte conservation autour du motif DE forme une sorte d'entonnoir sur le modèle 3D de PMTub. (**Figure** 30**B**). La résolution de la structure du dimère *Sc*PMT1-*Sc*PMT2 de levure a révélé des poches similaires à proximité des motifs DE dans lesquelles les substrats donneur et accepteur dans ont été respectivement détectées (**Figure** 30**C et D**). Cette poche


A
Motif DE
HH1
HH2
HH3
B
90°
HH1
HH2
Motif DE
HH3
Motif DE
Variable
Conservé

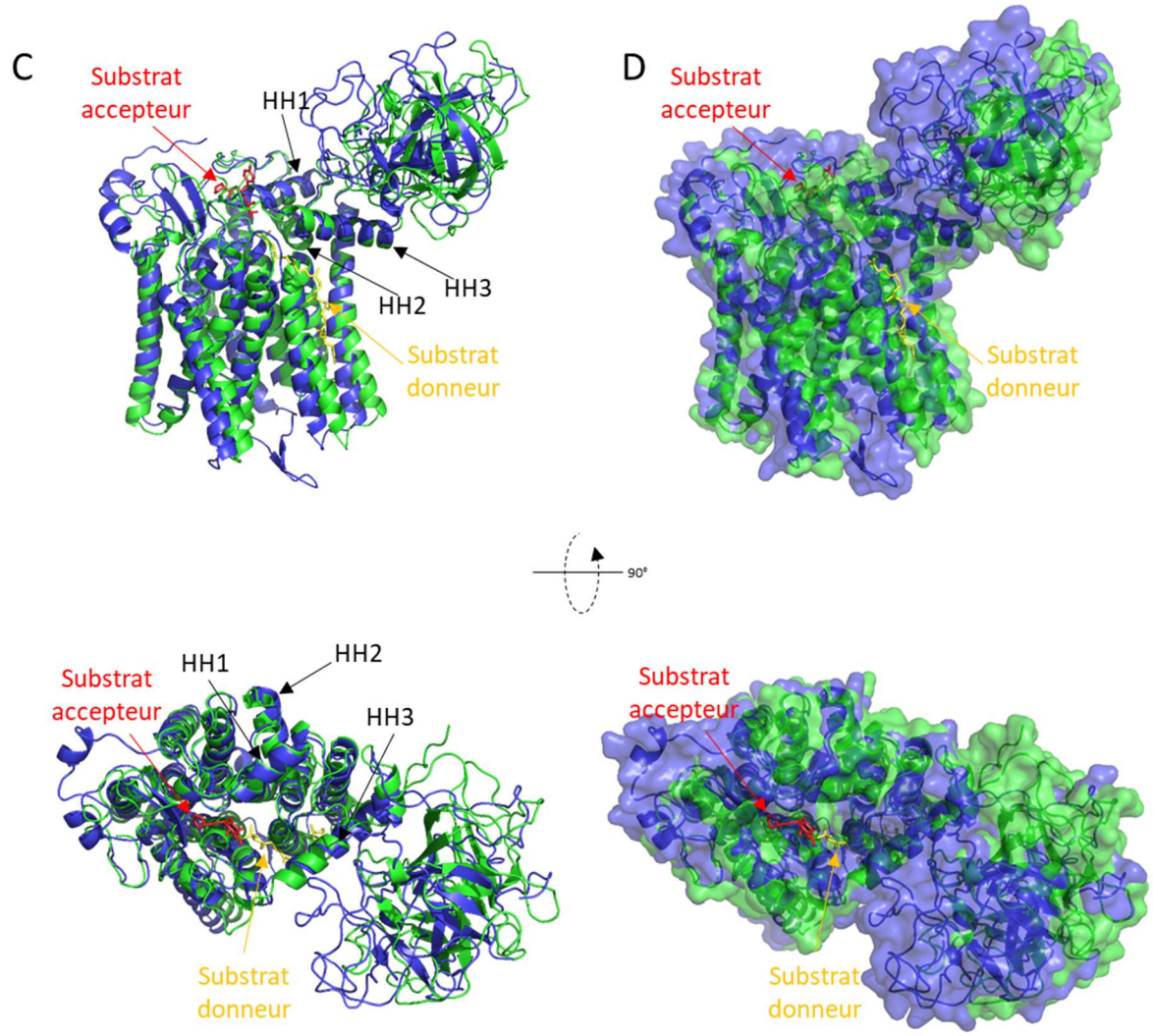


Figure 30. Conservation de séquence et de localisation 3D de la poche catalytique entre PMTub et les PMT eucaryotes. A) Visualisation en ruban du modèle de PMTub coloré selon le taux de conservation de séquence de PMTub par rapport à des PMT de divers organismes (voir liste Figure 7) vue de face (panel haut) et vue du dessus (panel bas). Sont annotées la localisation des 3 hélices horizontales (HH) ainsi que celle du motif DE. **B)** Visualisation surface du modèle de PMTub selon les mêmes codes couleur que Figure **30**A**. C)** Superposition des représentations en ruban de ScPMT1 et ScPMT2 vues de face (panel haut) et du dessus (panel bas). Est annotée la localisation des 3 hélices horizontales (HH). Le substrat accepteur est représenté en rouge et le substrat donneur en jaune. **D)** Superposition des représentations surface de ScPMT1 et ScPMT2 selon les mêmes codes couleur que Figure **30**C.

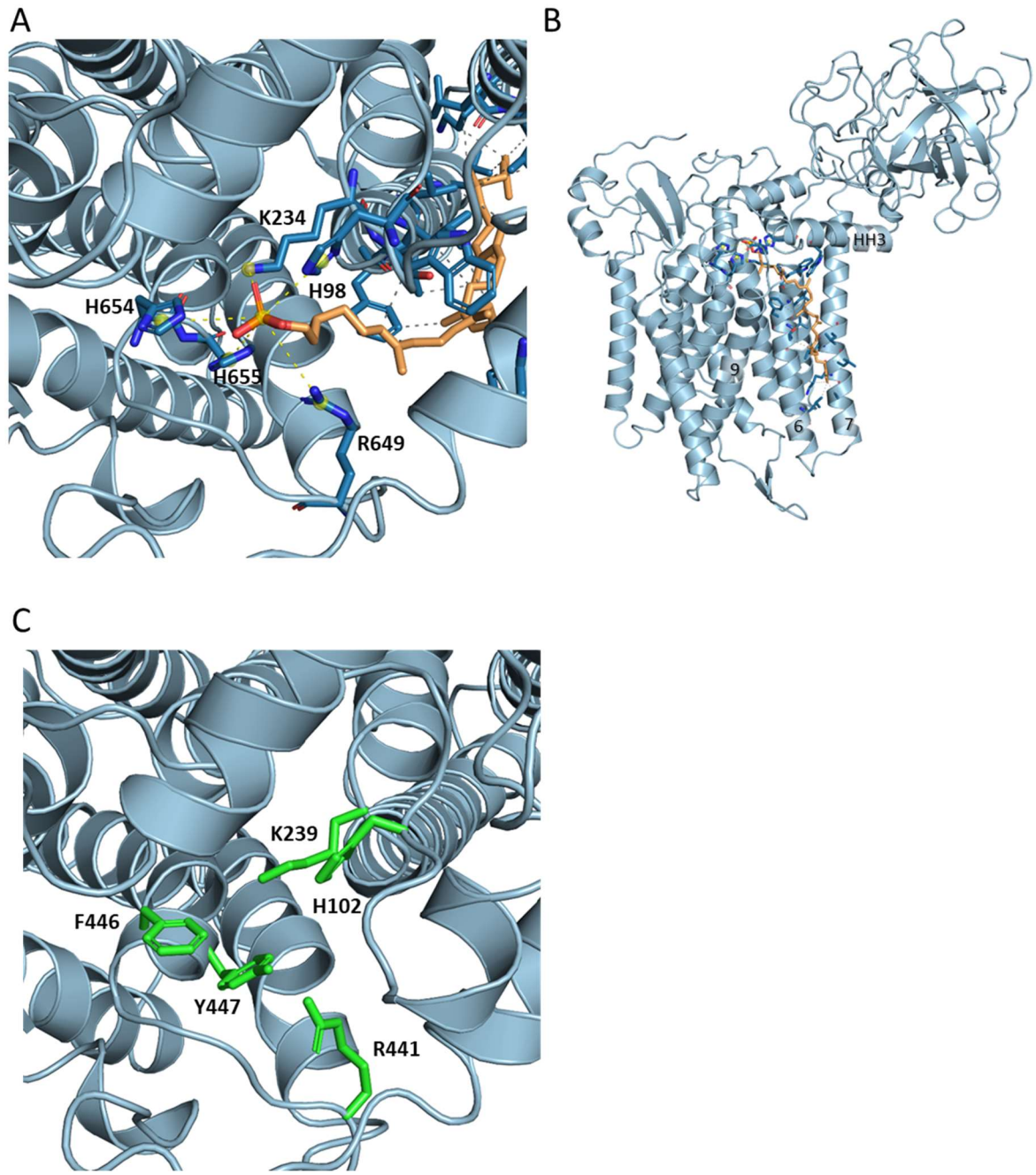


Figure 31. Interaction avec le substrat donneur. A) Acides aminés stabilisant le phosphate du substrat donneur dans la structure de ScPMT1[88], mis en évidence par l'outil PLIP[161]. Les pointillés jaunes représentent des liaisons ioniques et les pointillés gris représentent des interactions hydrophobes. Les acides aminés participant à la stabilisation du phosphate sont numérotés. **B)** Acides aminés stabilisant le substrat donneur dans la structure de ScPMT1, mis en évidence par l'outil PLIP[161]. Les segments transmembranaires 6, 7 et 9 sont numérotés et HH3 est annotée. **C)** Acides aminés stabilisant probablement le phosphate du substrat donneur dans le modèle de PMTub de phyre2 identifiés par homologie avec ScPMT1.

dans PMTub pourrait donc probablement, par analogie à ce qui a été observé dans le complexe *Sc*PMT1-*Sc*PMT2, loger les substrats donneur et accepteur.

La présence du motif DE essentiel à la catalyse, la proximité des deux substrats, et la forte conservation des acides aminés dans cette zone structurale suggèrent fortement que cette poche pourrait être le lieu où se déroule la réaction de transfert de mannose. Nous nous sommes donc intéressés dans un premier temps à cette zone conservée dans le but de confirmer expérimentalement et de mieux définir les rôles éventuels de certains de ces acides aminés conservés dans les mécanismes de reconnaissance des substrats donneur et accepteur du mannose et ainsi que sur la réaction chimique catalysée par ces enzymes.

b) Interaction avec le substrat donneur

Chez la levure, on peut identifier à une extrémité de la poche structurale fortement conservée, une zone de forte densité en acides aminés chargés positivement où vient se loger le phosphate du lipide-P-mannose. En effet, dans la structure de *Sc*Pmt1, le phosphate du substrat donneur est stabilisé par H98, K234, R649, H654 et H655[88] (**Figure 31A**). Une configuration similaire est observée pour *Sc*PMT2. Cette zone chargée positivement débouche sur une crevasse formée par les domaines transmembranaires 6, 7 et 9, et l'hélice horizontale HH3, où la partie lipidique du donneur est stabilisée par une série d'acides aminés hydrophobes (**Figure 31B**).

Compte tenu de l'homologie des substrats donneurs dans les GT-C, leurs sites de liaison aux phospholipides présentent de grandes similitudes. Un sillon hydrophobe est toujours décrit comme faisant partie de l'interface avec le lipide et qui se termine par une poche chargée positivement responsable de la stabilisation du phosphate. Dans AglB, ce site de liaison pour le groupe phosphate est composé de H81, H162 et R426[124]. Lorsque les histidines sont mutées en alanine, une perte d'activité partielle est observée, alors que la substitution en un acide aminé chargé négativement conduit à une perte complète d'activité, conséquence de la perte d'interaction avec le groupement phosphate du substrat donneur[124]. La mutation de l'arginine R426 en alanine entraine également une perte totale de l'activité[124].

Sur le modèle 3D de PMTub, une poche chargée positivement est aussi identifiée à une position identique et pourrait donc participer à un mécanisme de stabilisation similaire. Elle

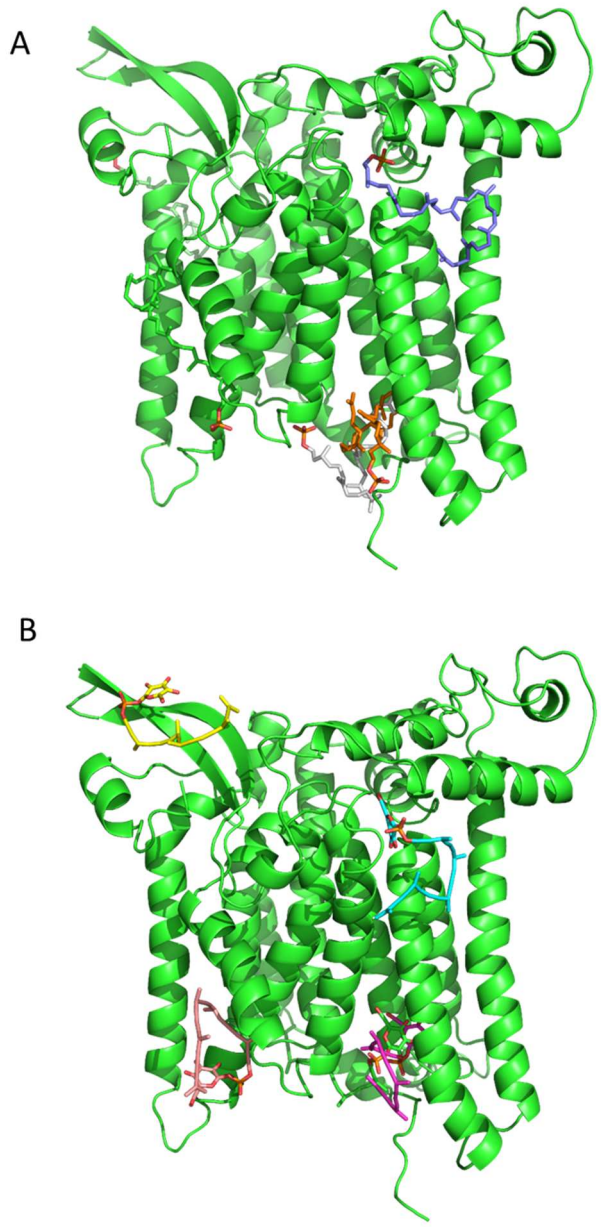


Figure 32. Docking du substrat donneur dans le modèle structural de PMTub. A) Docking de Dol-P sur le modèle de PMTub généré par Phyre2 à l'aide de la méthode de docking CB-Dock[154,162]. Le modèle de PMTub est représenté en ruban vert. Les 5 essais de docking sont représentés en bâton avec la partie lipidique colorée différemment et le phosphate en orange. Le Dol-P représenté en violet adopte un positionnement du phosphate similaire à ce qui est observé chez la levure. **B)** Docking de Dol-P-Man sur le modèle de PMTub généré par Phyre2 à l'aide de la méthode de docking CB-Dock[154,162]. Le modèle de PMTub est représenté en ruban vert. Les 5 essais de docking sont représentés en bâton avec la partie lipidique et le mannose colorés différemment et le phosphate en orange.

se compose de H102, K239, R441, F446 et Y447 qui pourrait correspondre aux acides aminés H98, K234, R649, H654 et H655 définissant cette même poche dans les PMT de levures (**Figure 31C et Figure 25**). Du fait de leur propriété chimique, seul H102, K239 et R441 pourraient avoir un rôle dans la stabilisation du groupement phosphate par interaction ionique. Pour les acides aminés F446 et Y447 un rôle différent peut être envisagé : il a été démontré que les acides aminés aromatiques peuvent avoir un rôle dans l'interaction avec des glucides[163,164]. F446 et Y447 pourraient donc interagir avec le mannose du substrat donneur par stacking.

Afin de vérifier la vraisemblance de cette hypothèse, j'ai réalisé un amarrage moléculaire (docking) d'un lipide-P dans le modèle 3D de PMTub en me servant de la suite logiciel CB-Dock (Cavity-detection guided blind docking)[162,165] . Pour cela, un dolichol-phosphate similaire à celui obtenu dans la structure de *Sc*PMT1 a été utilisé. Les positionnements du substrat résultant du docking correspondent principalement à des zones hydrophobes de PMTub, notamment au niveau des domaines transmembranaires qui interagissent avec les lipides de la membrane. Cependant, un des résultats du docking vient placer le lipide-phosphate de manière similaire à ce qui est observé dans la structure des PMT de levure (**Figure 32A**, essai de docking coloré en violet). Afin de vérifier le rôle possible des acides aminés conservés de la poche chargée positivement dans la stabilisation de cet analogue tronqué du substrat donneur, j'ai analysé les interactions potentielles du groupement phosphate de la molécule de Dol-P après docking sur la protéine.

Ainsi lorsque le lipide adopte par docking un positionnement similaire à celui observé dans le complexe *Sc*PMT1-*Sc*PMT2, le groupement phosphate du substrat donneur se trouverait stabilisé par H102, K239 et R441 chez PMTub. Par conséquent, si l'on se réfère à ce qui est observé par mutagenèse chez AglB[124], on peut prévoir une perte d'activité totale de PMTub après mutation de l'arginine R441 et une perte partielle après mutation de l'histidine H102.

Du fait de l'absence de données sur l'interaction entre le mannose et les *Sc*PMT, il est plus difficile de prédire la zone d'interaction du sucre chez PMTub. Des essais de docking d'un Dol-P-Man sur le modèle de PMTub ont conclu à une zone d'interaction possible. Néanmoins, ces résultats semblent peu cohérents avec un transfert du sucre puisque le carbone anomérique C1 du mannose n'est pas accessible à la dyade catalytique DE (**Figure 32B**, essai

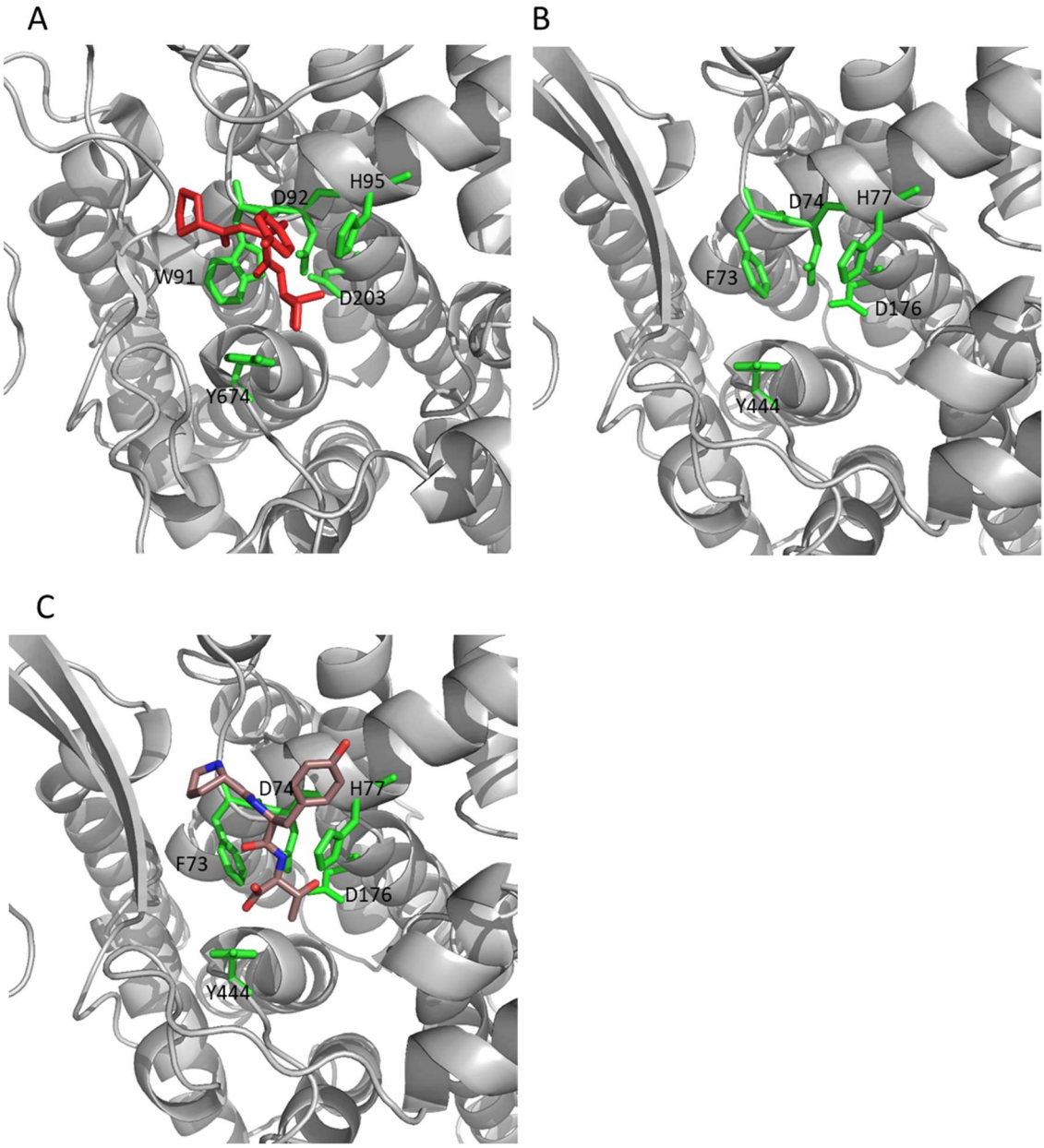


Figure 33. Interaction avec le substrat accepteur. A) Vue du site d'interaction du substrat peptidique accepteur tel que défini structuralement pour ScPMT2[88]. Le peptide substrat est représenté en rouge et les acides aminés du site d'interaction identifiés structuralement en vert. **B)** Vue similaire sur le modèle 3D de PMTub, les acides aminés correspondant au site d'interaction prédit de ScPMT2 sont représentés en vert. **C)** Positionnement d'un peptide substrat accepteur dans le modèle de PMTub au niveau du site d'interaction prédit par le serveur « flexpepdoc »[166,167]**.**

de docking en cyan). Cependant, comme décrit dans le cas d'une autre GT-C : *Sc*ALG6 [123], il est possible que la structure des PMT soit dynamique et que la présence ou l'absence de mannose entraine des changements conformationnels subtils. En effet, dans le cas de ScALG6 Bloch *et al.*[123] ont récemment émis l'hypothèse que la présence du substrat donneur portant un sucre engendre un changement de configuration de l'enzyme permettant de bloquer le substrat donneur dans le site actif jusqu'au transfert du sucre. Cette étude décrit aussi une spécificité de substrat forte pour le sucre : ALG6 utilise un dolichol-P-Glucose comme substrat donneur et des études de docking ont permis de révéler qu'un dolichol-P-Mannose ne pouvait pas se loger dans la cavité du substrat donneur. Basé sur ces observations on peut raisonnablement envisager un comportement similaire pour les PMT. Des analyses supplémentaires de docking autorisant une plus grande flexibilité du modèle pourraient être d'une grande utilité pour définir plus précisément l'interface de reconnaissance du mannose par les PMT.

En conclusion, nous avons pu identifier les acides aminés H102, K239 et R441 de PMTub comme ayant un rôle probable dans l'interaction avec le phosphate du substrat donneur. Basés sur l'étude antérieure d'AglB[124], nous proposons que seul R441 ait un rôle essentiel dans cette interaction et que H102 et R239 aient des rôles accessoires. Nous avons donc étudié le rôle de R441 dans l'activité de PMTub par mutagenèse (voir p.123).

c) Interaction avec le substrat accepteur

Sur la structure de *Sc*PMT2, qui a été résolue avec un peptide accepteur, plusieurs acides aminés ont été structuralement identifiés dans un voisinage proche suggérant qu'ils pourraient participer à l'interaction et à la stabilisation du groupement hydroxyle de l'acide aminé accepteur (**Figure 33A**).

Comme l'ensemble des GT-C, les PMT portent une variation plus ou moins dégénérée du motif DxD à l'extrémité de l'hélice HH1. Chez les PMT, on retrouve donc à cet endroit un motif DE conservé. Les deux acides aminés de ce motif ont pu être démontrés comme essentiels à l'activité des PMT chez la levure avec une perte totale d'activité lors de la substitution de ce motif par deux alanines[54]. L'étude des structures de *Sc*PMT1 et *Sc*PMT2 a permis d'éclaircir le rôle de ce motif essentiel comme expliqué précédemment (voir p.37 et 63). L'acide aspartique est idéalement positionné pour activer l'hydroxyle de l'acide aminé accepteur et l'acide glutamique forme un pont salin avec une lysine conservée pour maintenir

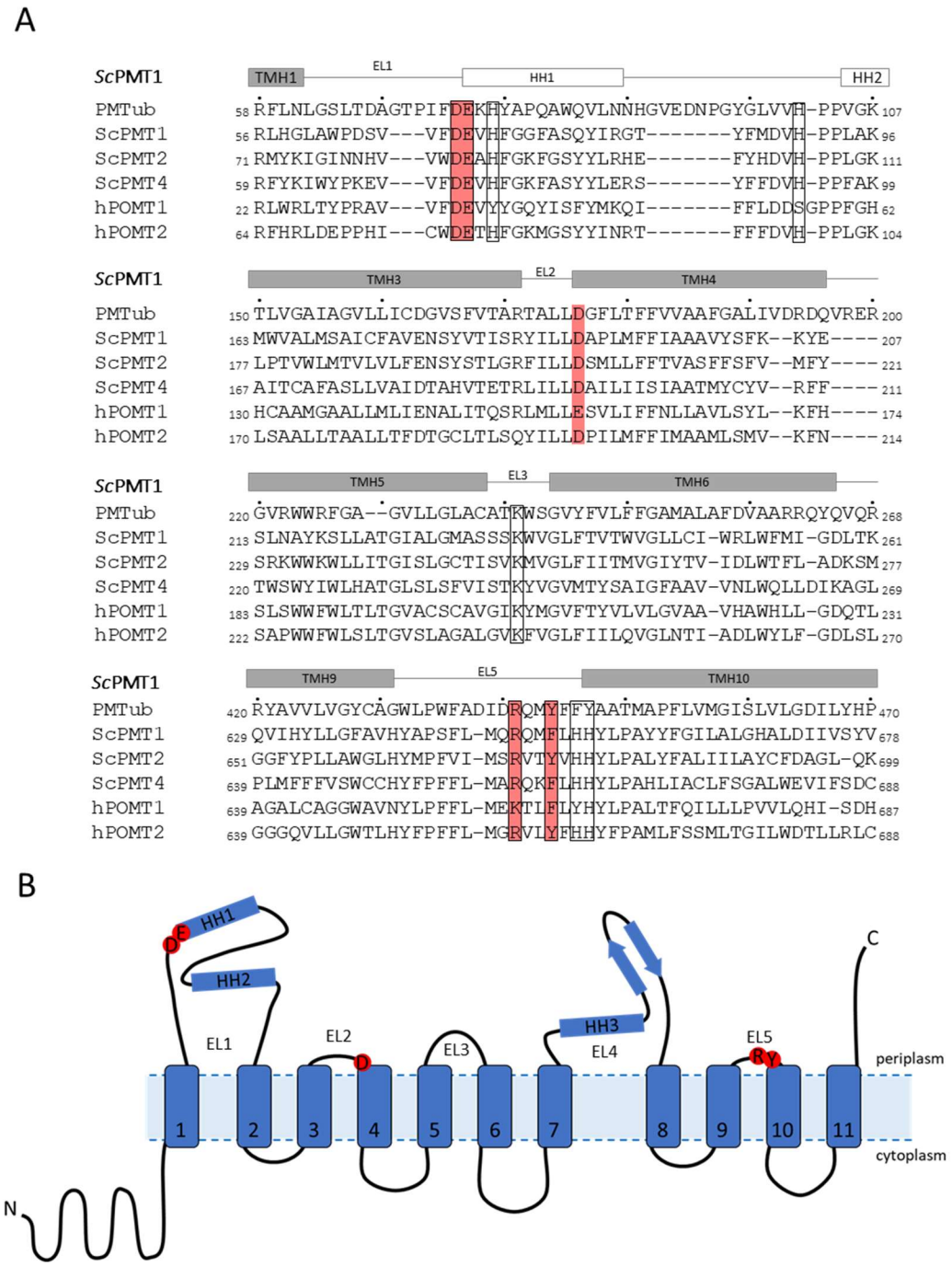


Figure 34. Site actif prédit pour PMTub A) L'alignement des boucles externes (EL) 1 à 5 de PMTub (EL) avec d'autres membres de la famille des PMT met en évidence les acides aminés conservés. Les résidus impliqués dans le site actif de *Sc*PMT1 et *Sc*PMT2 sont encadrés en noir, tels qu'identifiés précédemment[88]. En rouge sont représentés les acides aminés qui ont été mutés au cours de cette étude. ScPMT : *S. cerevisiae* PMT ; hPOMT : POMT humain. **B)** La topologie prédite de PMTub décrit 11 hélices transmembranaires (TMH), 3 hélices horizontales périplasmiques (HH) et 2 brins bêta. D, E, R et Y indiquent les acides aminés conservés dans les boucles externes périplasmiques (external loops ou EL) 1,2 et 5 et dont le rôle a été étudié durant la thèse

l'intégrité structurale du site actif en orientant la chaine latérale de l'acide aspartique **(Figure 13B**). Un autre acide aspartique est localisé à proximité de la fonction hydroxyle du substrat accepteur sur la structure de *Sc*PMT2 : D203 (**Figure 33A**). On retrouve toujours un acide aminé acide à cette position dans les PMT et chez PMTub il correspond à D176. Cet acide aminé pourrait aussi participer à la stabilisation de la fonction hydroxyle du substrat accepteur.

En plus de ces acides aspartiques, plusieurs autres acides aminés semblent être impliqués dans la stabilisation de la fonction hydroxyle. Chez *Sc*PMT2, on note ainsi W91, H95 et Y674 (correspondant à F76, H80 et F652 chez *Sc*PMT1) (**Figure 33A**)[88]. Ces acides aminés ne sont pas strictement conservés chez toutes les PMT, mais on retrouve toujours des acides aminés aromatiques à ces positions (**Figure 34A**). Cette observation résulte uniquement de la structure des PMT de levures, mais ces positions n'ont jamais été étudiées fonctionnellement.

Il est intéressant de noter que l'on retrouve également dans PMTub, des acides aminés équivalents : F73, H77 et Y444 (**Figure 33B**). Le serveur « flexpepdoc » a été utilisé pour modéliser l'interaction entre le modèle 3D de PMTub et un peptide « PYT » tel que celui visualisé dans la structure de *Sc*PMT2[166,167]. Le docking du peptide mène, *in silico*, à une interaction entre la chaine latérale de D74 (du motif DE) et le groupement hydroxyle de la thréonine du peptide substrat cohérente avec la fonction catalytique prédite de ce résidu (**Figure 33C**). Autour de la fonction hydroxyle viennent aussi se positionner les autres acides aminés identifiés par comparaison avec les *Sc*PMT. Cette simulation conforte donc l'identification de D74, F93, H77 et Y444 comme site d'interaction avec la fonction hydroxyle de l'acide aminé cible du substrat accepteur.

Hormis ces interactions impliquant la fonction hydroxyle de l'acide aminé cible, il n'a pas pu être mis en évidence de quelconque conservation d'acides aminés de PMTub pouvant suggérer une similarité du mécanisme de reconnaissance des autres acides aminés du peptide cible par les PMT. Ce constat est en accord avec l'observation qu'il n'a pas pu être défini de séquence canonique en acide aminé pour les sites de mannosylation. Alors que de telles séquences ont pu être définies pour les sites de N-glycosylation (N-X-T/S)[168] et de C-glycosylation (W-X-X-W)[169], aucune séquence canonique de reconnaissance des mannoprotéines par les PMT n'a pu être identifiée à ce jour. Il peut cependant être observé que les sites de *O*-glycosylation mycobactériens se caractérisent par une répétition de serines

ou thréonines avec une richesse en proline et alanine de part et d'autre[38]. De plus, il est décrit que les différentes isoformes de PMT de levures ne glycosylent pas toutes les mêmes protéines, sous-entendant des mécanismes de reconnaissance propre à chaque enzyme[92]. L'étude des mécanismes de reconnaissance des mannoprotéines par les PMT est présentée plus en détail dans la Partie 2.C de ce document (page 129).

5. Conclusion et perspectives de l'étude par modélisation

Construit à partir des structures 3D de *Sc*PMT1 et *Sc*PMT2, et affiné à l'aide de structures de GT-C, le modèle de PMTub permet d'énoncer plusieurs hypothèses quant aux relations structure/fonction de cette enzyme. L'étude du modèle 3D de PMTub permet en effet de mettre évidence que cette enzyme semble être constituée principalement, d'un point de vue structural, du socle commun des GT-C, décoré de structures solubles caractéristiques, principalement localisées dans la boucle EL4.

La comparaison du modèle de PMTub aux structures de *Sc*PMT1 et *Sc*PMT2 a permis de proposer des sites d'interaction pour les substrats donneurs et le substrat accepteur de PMTub. Ces sites se composent d'acides aminés hautement conservés parmi les PMT (**Figure 34A**) qui se répartissent sur les différentes boucles périplasmiques de PMTub (**Figure 34B**).

Le mécanisme catalytique des PMT reste peu étudié. La résolution de la structure de ScPMT1 et ScPMT2 a permis d'identifier structuralement des résidus d'intérêt, mais hormis le motif DE invariant, ceux-ci n'ont jamais été caractérisés d'un point de vue fonctionnel. De façon générale, l'étude fonctionnelle des acides aminés identifiés dans la structure PMTub comme conservés parmi les PMT, permettrait d'apporter une meilleure compréhension de la réaction catalysée par cette enzyme, et plus largement par toutes les PMT.

Afin d'étudier le rôle de ces différents acides aminés dans l'activité catalytique de PMTub, en fonction leur position et de la nature biochimique de leur chaîne latérale, l'approche conventionnelle par mutagenèse dirigée a été choisie en première intention par l'équipe. Des mutations ponctuelles ont donc été réalisées sur les acides aminés d'intérêt, correspondant aux différentes positions identifiées par l'étude du modèle 3D de PMTub, dans le but d'évaluer leur impact sur l'activité de l'enzyme en mettant en œuvre le test d'activité préalablement développé (Partie 1, p.68).

B. Caractérisation du site actif de PMTub

1. Mutagenèse

Basés sur les observations que nous avons réalisées sur les structures des *Sc*PMT et le modèle 3D de PMTub, ainsi que sur les alignements de séquences des PMT avec celles des GT-C de structure 3D résolue, plusieurs acides aminés nous sont apparus intéressants à étudier d'un point de vue fonctionnel. Les mutants de PMTub correspondants ont donc été construits. En particulier, nous avons identifié et prédit quatre acides aminés comme essentiels pour la catalyse de la réaction de transfert de mannose : D74, D176, R441 et Y444 (**Figure 31C, Figure 33C et Figure 34**). Des mutations dites « destructrices » et « compensatrices » ont été effectuées permettant, respectivement, de détruire ou compenser la fonction chimique de la chaine latérale de l'acide aminé muté.

- D74

Dans la littérature, les mutations de D74 en alanine ou asparagine ont déjà été rapportées comme abolissant totalement l'activité de PMTub[55]. Nous avons pu confirmer ces résultats en utilisant le test phénotypique d'activité de PMTub développé au début de ma thèse (**Figure 35A**). Cet acide aspartique prend part au motif DE des PMT qui est un dérivé du motif DxD commun à toutes les GT-C. Son rôle est communément admis comme étant celui d'une base dans un mécanisme SN_2 qui viendrait ici activer le groupement hydroxyle du substrat accepteur[133]. C'est en effet ce qui est décrit pour l'ensemble des GT-C de structure connue pour lesquelles on trouve toujours un acide aspartique (faisant partie d'un motif DxD) impliqué dans le mécanisme catalytique avec un rôle de base. Cet acide aspartique se présente toujours en amont de l'hélice horizontale HH1.

La capacité à compenser la perte d'activité par un autre acide aminé acide a aussi été mesurée par substitution du D74 en un acide glutamique (**Figure 35A**). Cette substitution n'a pas permis de restaurer significativement l'activité de PMTub, démontrant le rôle crucial de ce résidu dans l'activité PMTub. L'acide glutamique a une chaine latérale plus longue qu'un acide aspartique ce qui pourrait expliquer le défaut de restauration de l'activité : le groupement acide ne serait plus positionné correctement pour stabiliser la fonction hydroxyle

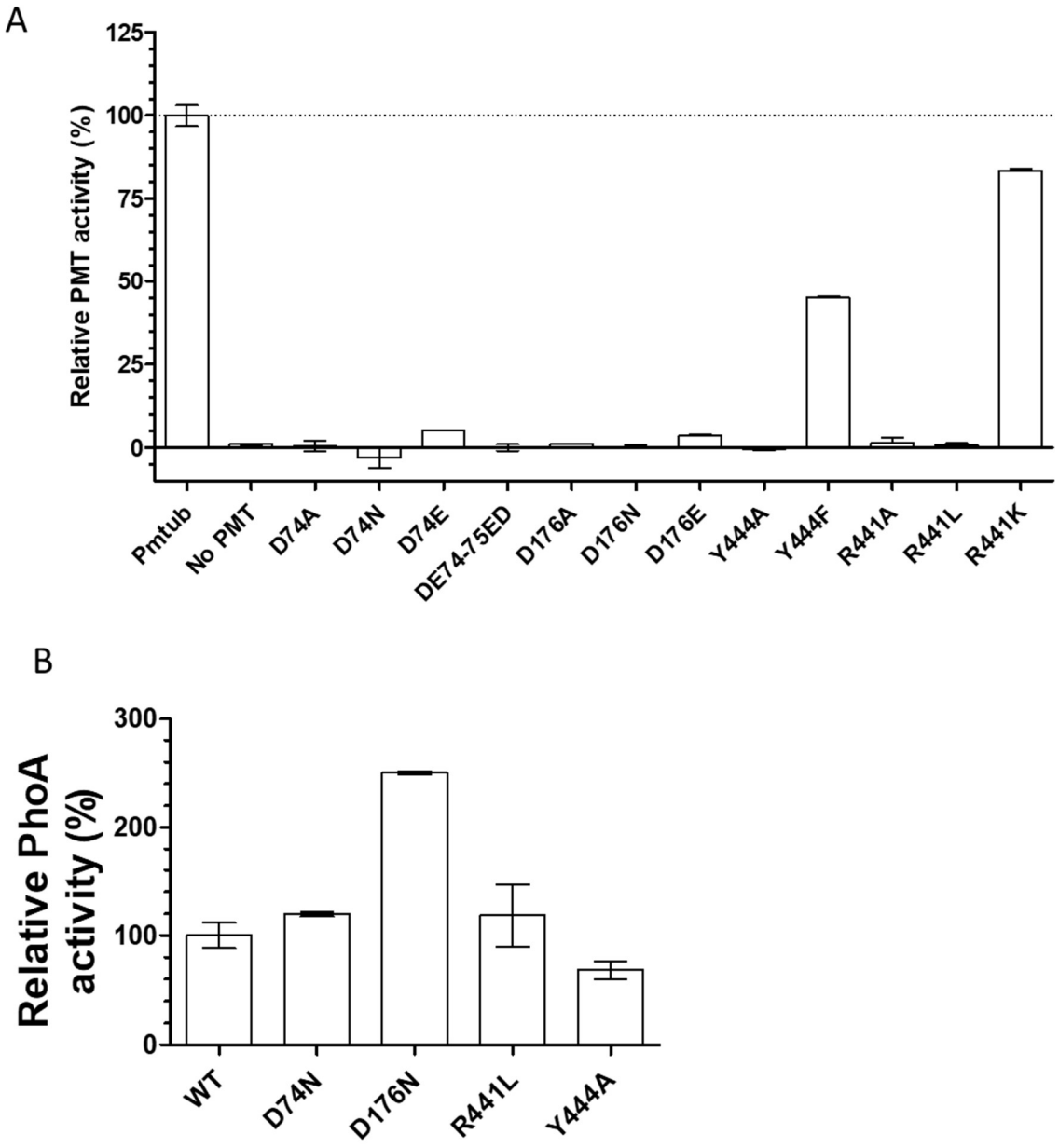


Figure 35. Etude du site actif prédit de PMTub par mutagénèse dirigée. A) Résultats du test d'activité phénotypique pour des mutants du site actif prédit de PMTub. 0% et 100% sont définis comme les résultats pour la souche « No PMT » et pour la souche exprimant PMTub sauvage respectivement. Les résultats présentés sont des triplicats techniques de duplicats biologiques. **B)** L'activité PhoA est mesurée en relatif à l'activité de la souche exprimant PMTub-PhoA (défini comme 100% arbitrairement). Les résultats présentés sont des triplicats techniques de duplicats biologiques.

du substrat accepteur. Les deux acides aminés acides du motif DE ont aussi été intervertis pour déterminer si leur rôle pouvait se compenser. Le mutant DE74-75ED mène à une perte totale de l'activité de PMTub, réfutant donc notre hypothèse (**Figure 35A**). Ces résultats confirment les observations d'études précédentes et permettent de mettre en exergue le rôle fondamental de ce motif acide dans l'activité des PMT.

Cette perte d'activité n'est pas liée à un problème d'expression de PMTub à la membrane (**Figure 35B**). En effet, la mutation de D74 sur la protéine de fusion PMTub-PhoA n'affecte pas l'activité phosphatase alcaline mesurée.

- D176

La mutagenèse du second acide aspartique identifié, D176, mène à un profil similaire à celui observé pour D74 (**Figure 35A**). La position D176 n'avait jamais été étudiée auparavant ni pour PMTub, ni pour aucune PMT. Dans la structure des PMT de levure, l'acide aminé équivalent se positionne à proximité de l'acide aspartique du motif DE, et il participe à la formation de la poche conservée où se positionnent les deux substrats. La perte totale d'activité pour les mutants D176A, D176N et D176E montre la grande importance de cette position dans la catalyse puisque le niveau d'expression n'est quant à lui pas affecté (**Figure 35B**). Un rôle d'acide dans une réaction SN_2 peut être proposé pour ce résidu, où D176 viendrait compléter celui de D74. On aurait donc une dyade catalytique composée de 2 aspartates, qui avec la sérine ou la thréonine acceptrice formeraient le trio enzymatique classique acide-base-nucléophile (correspondant à D176-D74-Ser/Thr respectivement). Cependant, chez la levure ces aspartates sont assez éloignés du peptide substrat, remettant en cause l'hypothèse d'un rôle direct dans la catalyse[88].

Dans d'autres GT-C, un aspartate essentiel est aussi retrouvé à cette position et celui-ci a pour rôle la coordination de l'ion métallique de ces enzymes. Par exemple, PglB contient un ion qui a pour rôle de maintenir l'orientation des chaînes latérales des acides aminés du site actif[116,170]. De façon notable, pour les PMT, il n'a jamais été formellement démontré qu'un ion était essentiel à la catalyse. Pour *h*POMT par exemple, l'activité mannosyltransférase ne semble pas requérir la présence d'un ion divalent[97]. Il est donc difficile de conclure sur le potentiel rôle de coordinateur d'ion de D176. De plus, au vu des observations réalisées sur les

PMT de levure dont la structure a été résolue[88], notre hypothèse serait que l'orientation de D74 dans le site actif soit maintenue par un pont salin formé entre E75 et R125 pour PMTub. Si l'orientation de tous les éléments du site actif est maintenue par des ponts salins, la fonction d'un ion stabilisateur parait alors redondante et par conséquent non essentielle.

D176 pourrait en revanche également former un pont salin afin de maintenir l'intégrité du site actif, à l'instar de E75. En accord avec cette hypothèse, il est possible que la chaine latérale de D176 soit stabilisée par celle de la lysine voisine dans l'espace, et conservée dans les PMT (**Figure 34**). Par analogie avec les PMT de levure[88], K239 se situerait à proximité de la zone d'interaction avec le phosphate du substrat donneur et participerait donc probablement à la stabilisation de celui-ci au travers de la charge positive de sa fonction biochimique. Notablement, dans les PMT de levures, les chaines latérales des acides aminés équivalents à D176 et K239 se situent à une distance suffisante pour former un pont salin : 2,8Å pour ScPMT1 et 2,4 Å pour ScPMT2[88] , étayant donc cette hypothèse.

En conclusion, D176 est un acide aminé essentiel à l'activité de PMTub, mais à défaut de structure 3D résolue, son rôle précis n'est pas élucidé à ce jour. Nos observations et comparaisons avec la littérature nous ont amenés à formuler 3 hypothèses distinctes quant au rôle de D176.

Hypothèse 1 : Ce résidu pourrait jouer un rôle direct dans la catalyse en tant qu'acide ou « donneur de proton », rappelant le mécanisme réactionnel décrit communément pour les β-glycosidases. Cependant l'éloignement de l'acide aspartique équivalent, dans la structure du complexe ScPmt2/peptide chez la levure[88], vis-à-vis du substrat accepteur, tend à réfuter cette première hypothèse.

Hypothèse 2 : Alternativement, la chaîne latérale de D176 pourrait contribuer à la stabilisation du groupement phosphate du substrat donneur, indirectement, *via* la coordination d'un ion métallique chargé positivement (mais dont l'essentialité est plutôt réfutée dans les PMT)[97].

Hypothèse 3 : Le deuxième rôle indirect évoqué, de stabilisation du groupement phosphate du substrat donneur, pourrait plutôt s'exercer *via* la stabilisation, par pont salin, de la lysine voisine dans l'espace (K239, dont la fonction biochimique est également chargée positivement). La conservation de cette dernière parmi les PMT tendrait à étayer plutôt le

dernier de ces 3 scénarios (**Figure 34**), cependant une étude précédente a montré que cet acide aminé ne semble pas indispensable à l'activité de ScPMT1[53] remettant aussi en question cette troisième hypothèse.

- Y444

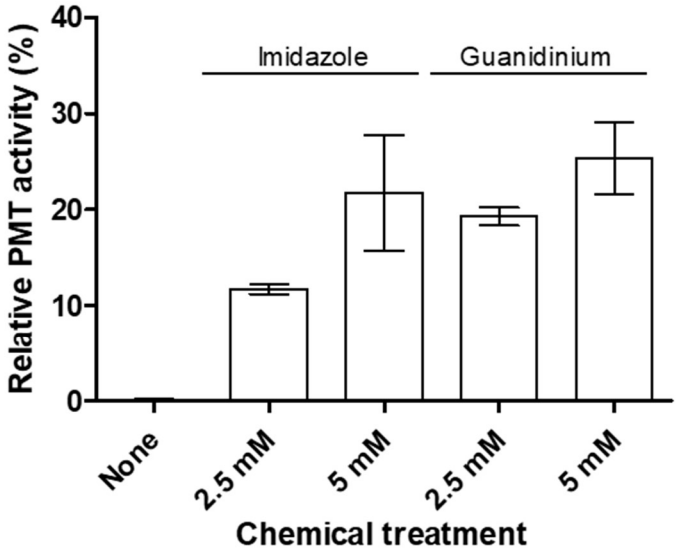
Y444 est conservé dans ScPMT2 où il est prédit comme impliqué dans l'activité enzymatique[88]. Dans la structure de ScPMT1, l'équivalent à Y444, F652, est prédit comme formant le site de liaison à Ser/Thr accepteur, avec F76, D77 et H80[88]. Selon nos résultats, la mutation de Y444 en alanine entraine une perte totale d'activité, et sa mutation en phénylalanine réduit l'activité de moitié (**Figure 35A**). Comme pour les mutants précédents, un maintien d'expression est observé (**Figure 35B**). Le caractère aromatique de cet acide aminé semble donc important, mais pas suffisant pour expliquer son rôle dans l'activité de l'enzyme. La fonction hydroxyle de cette tyrosine peut peut-être expliquer la différence avec la phénylalanine, en facilitant l'interaction avec Ser/Thr accepteur.

- R441

Dans *Sc*PMT1, pour lequel la structure a été co-résolue avec le phospholipide (partie lipidique du substrat donneur), H98, K234, R649, H654 et H655 sont situés à proximité immédiate du phosphate du donneur et pourraient participer à un mécanisme de stabilisation de celui-ci. Dans PMTub, ces résidus correspondent à H102, K239 (mentionnée dans la paragraphe précédent), R441, F446 et Y447. Par analogie avec les études structurales et fonctionnelles effectuées pour *Af*AglB, nous avons considéré que seul R441 pourrait jouer un rôle essentiel dans la stabilisation du phosphate, les autres acides aminés de ce quintuor ayant seulement un rôle accessoire[124].

Nous avons donc étudié seulement R441 par mutagenèse. La mutation en leucine ou en alanine de cette position mène à une perte totale d'activité, qui est presque totalement restaurée pour la mutation R441K (**Figure 35A**). Ces résultats montrent le caractère essentiel d'un acide aminé chargé à cette position, cohérent avec le rôle prédit dans la stabilisation du

A



B

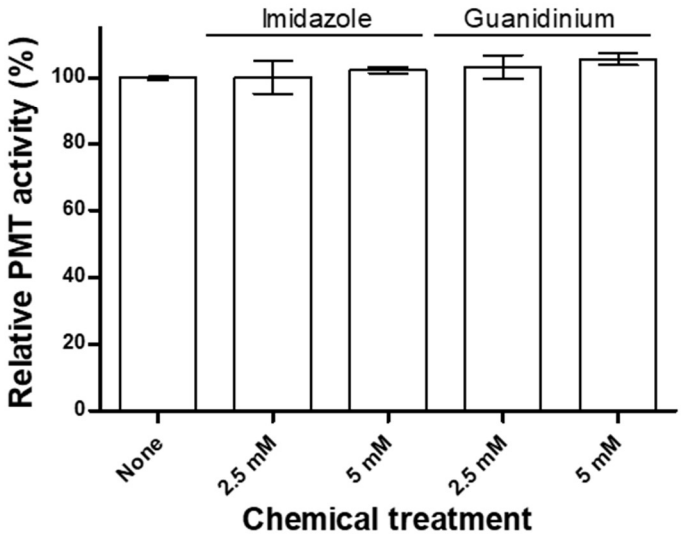


Figure 36. Résultats du « chemical rescue » de l'activité du mutant PMTub-R441A *in vivo*. A) Activité relative de l'enzyme PMTub portant la mutation R441A traitée ou non traitée, exprimée en pourcentage calculé à partir de l'activité de référence PMTub native non traitée. Les valeurs moyennes des duplicats biologiques sont présentées. **B)** Pourcentage d'activité relatif de l'enzyme PMTub native traitée ou non traitée. Les valeurs moyennes des duplicats biologiques sont présentées.

phosphate par formation d'une liaison ionique. Aucune perte d'expression de PMTub n'est notée après mutagenèse de cette position (**Figure 35B**).

Du fait de la large interface d'interaction avec la partie lipidique du substrat, des mutations ponctuelles individuelles ne sont pas attendues comme empêchant l'interaction avec le substrat lipidique donneur et cette zone n'a donc pas été étudié par la méthodologie que nous avons mise en place.

2. Chemical rescue

Une autre méthode possible de confirmation de l'implication de ces résidus dans l'activité catalytique par la fonction chimique de leur chaine latérale est la mise en place d'une expérience de « chemical rescue ». L'objectif de l'expérience de « chemical rescue » est de confirmer que le rôle de D74, D176 et R441 est dû aux propriétés chimiques de leur chaine latérale. Cette expérience consiste à essayer de restaurer les propriétés chimiques de ces trois acides aminés mutés en alanine par de petites molécules exogènes, supplémentées dans le milieu de culture. La molécule remplira la cavité laissée vide par la taille réduite de la chaine latérale de l'alanine par rapport aux acides aminés d'origine et devrait remplacer les fonctions chimiques supprimées par la mutation en alanine. Une activité résiduelle récupérée est une confirmation du rôle indispensable de la fonction biochimique du résidu muté.

Les rôles prédits des acides aminés D74, D176 et R441 vont guider la sélection des molécules exogènes utilisées. Pour les aspartates, les molécules exogènes rapportées dans la littérature pour compenser la fonction chimique acide sont l'acétate de sodium[171] et l'azide de sodium[172]. Pour l'arginine, les molécules couramment utilisées sont l'ion guanidinium[173] et l'imidazole[174].

Pour la mutation R441A, une restauration de l'activité entre les enzymes traitées et non traitées a été observée pour les deux composés chimiques employés (**Figure 36A**) sans que le traitement n'affecte l'activité de l'enzyme native (**Figure 36B**). Cette restauration confirme le caractère essentiel de la fonction basique de R441 dans l'activité de l'enzyme. Pour les mutations D74A et D176A, la supplémentation par les molécules exogènes n'a eu aucun effet détectable sur l'activité de PMTub (données non montrées).

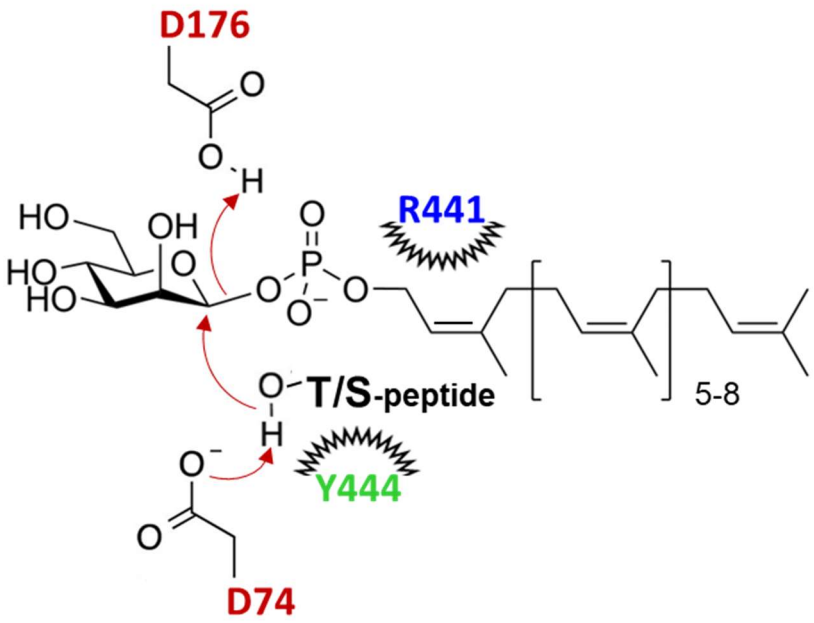


Figure 37. Proposition d'un mécanisme catalytique pour PMTub. Cette proposition se base sur l'ensemble des résultats récoltés lors de ce travail. Le transfert du sucre avec inversion de l'anomérie serait réalisé selon un mécanisme où D74 aurait le rôle de base générale et D176 aurait le rôle d'acide général, et Y444 et R441 stabiliserait les substrats accepteur et donneur respectivement.

Il est à noter que c'est la première fois qu'une expérience de « chemical rescue » est réalisée avec succès sur une glycosyltransférase et sur des bactéries vivantes. Cette expérience est également la première preuve que les molécules exogènes peuvent atteindre PMTub *in vivo* et affecter son activité. Cette démonstration est d'une importance majeure pour valider la faisabilité d'une inhibition de l'enzyme par traitement avec une molécule exogène.

3. Conclusion et proposition d'un mécanisme catalytique

En conclusion, cette étude permet de proposer un modèle de réaction catalysée par les PMT. Plusieurs positions conservées dans les PMT et localisées près du site actif putatif sur les structures de *Sc*PMT1 et *Sc*PMT2 ont pu être identifiées.

En déduction de nos observations et des données de la littérature des GT-C, nous proposons, pour la première fois, un mécanisme catalytique détaillé des Protéines *O*-Mannosyltransférases (**Figure 37**). Dans ce mécanisme catalytique, R441, situé sur EL5, stabiliserait le phosphate du substrat donneur consolidant sa capacité de groupe partant. Ceci favoriserait l'attaque nucléophile de l'hydroxyle de l'accepteur sérine ou thréonine sur le carbone anomérique de l'unité mannose. Le D74 agirait comme une « base générale » capturant le proton libéré lors de l'attaque nucléophile, et le D176 agirait soit comme un « acide général », libérant un proton pour le groupe partant des phospholipides, soit comme stabilisateur de la lysine voisine, qui elle-même concourrait à la stabilisation du phosphate avec R441. Le mécanisme proposé ici est en accord avec l'inversion de la configuration anomérique du sucre du substrat donneur, lors du transfert du mannosyle depuis le polyprénol-P-β-mannose, vers l'hydroxyle de la Ser/Thr de la protéine acceptrice.

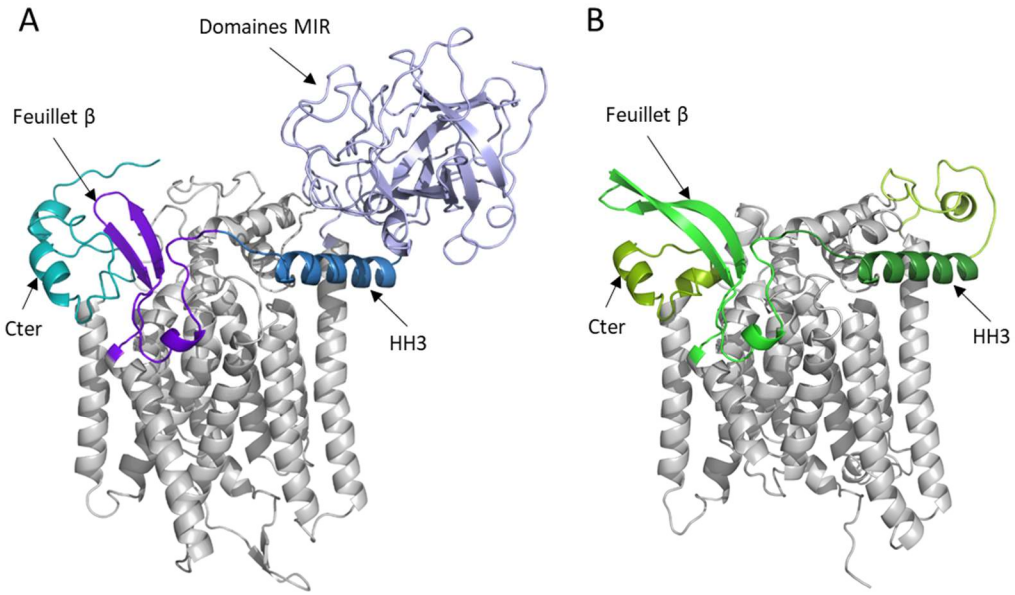


Figure 38. Visualisation de la boucle EL4 et de l'extrémité C-terminale des PMT. A) Structure 3D de *Sc*PMT1[88]. **B)** Modèle 3D de PMTub. Les différentes structures composant la boucle EL4 sont colorées, de l'extrémité C-terminale à N-terminale de la boucle (de droite à gauche sur la figure) : les domaines MIR (violet clair sur A) ou boucle non structurée chez la bactérie (vert clair sur B), HH3 (bleu sur A ou vert d'eau sur B), la petite hélice α et le feuillet β (violet sur A ou vert foncé sur B). L'extrémité C terminale de l'enzyme est aussi colorée (cyan sur A ou jaune sur B).

C. Identification de domaines fonctionnels

Deux principaux domaines d'intérêt, se distinguant des autres GT-C, peuvent être identifiés dans la structure des PMT : la boucle EL4 et l'extrémité C-terminale.

1. Étude structurale des GT-C et comparaison au modèle 3D de PMTub

a) Boucle « EL4 »

La quatrième boucle périplasmique, ou EL4 (pour External Loop 4 en anglais) correspond à la partie soluble de PMTub délimitée par les domaines transmembranaires 7 et 8. Son équivalent chez la levure se situe dans le lumen du réticulum endoplasmique et est par conséquent parfois nommé LL4 dans la littérature (Lumenal Loop 4).

Chez la levure, la boucle EL4 (ou LL4) est la plus longue des boucles solubles. Son étude structurale permet de mettre en évidence différentes structures secondaires. On note ainsi, en partant de son extrémité N-terminale, la présence des domaines MIR, suivis de l'hélice horizontale HH3, collée contre la membrane, puis d'une seconde hélice α, plus courte (5 acides aminés), elle aussi en contact avec la membrane, et finalement d'un feuillet β (**Figure 38A**).

L'EL4 des PMT eucaryote présente 3 domaines MIR (Mannosyl-transferase Inositol-triphosphate Ryanodine receptor). Ces domaines sont essentiels à l'activité enzymatique puisque leur délétion entraine une perte de la *O*-mannosylation des protéines chez *S. cerevisiae* [53]. De plus, la délétion du domaine MIR de *Sc*PMT1 s'accompagne d'une perte de la dimérisation avec *Sc*PMT2[53]. La résolution récente de la structure du complexe *Sc*PMT1-*Sc*PMT2 a confirmé que les domaines MIR constituent un point d'interaction entre les deux enzymes et a permis de définir ceux-ci comme structurés en tonneau, formant ainsi un domaine globulaire soluble[88].

Au-delà de leur essentialité et de leur participation à la dimérisation, le rôle des domaines MIR est mal défini. Néanmoins, plusieurs modules canoniques d'interaction avec des sucres ont pu être identifiés sur leur séquence. Récemment, l'étude fonctionnelle et structurale des domaines MIR de *Sc*PMT2 et *Sc*PMT5 a permis d'analyser le rôle de ces modules canoniques identifiés, mais a aussi apporté une meilleure compréhension de la fonction des domaines MIR[87]. Cette étude a validé la présence de sites d'interaction avec des

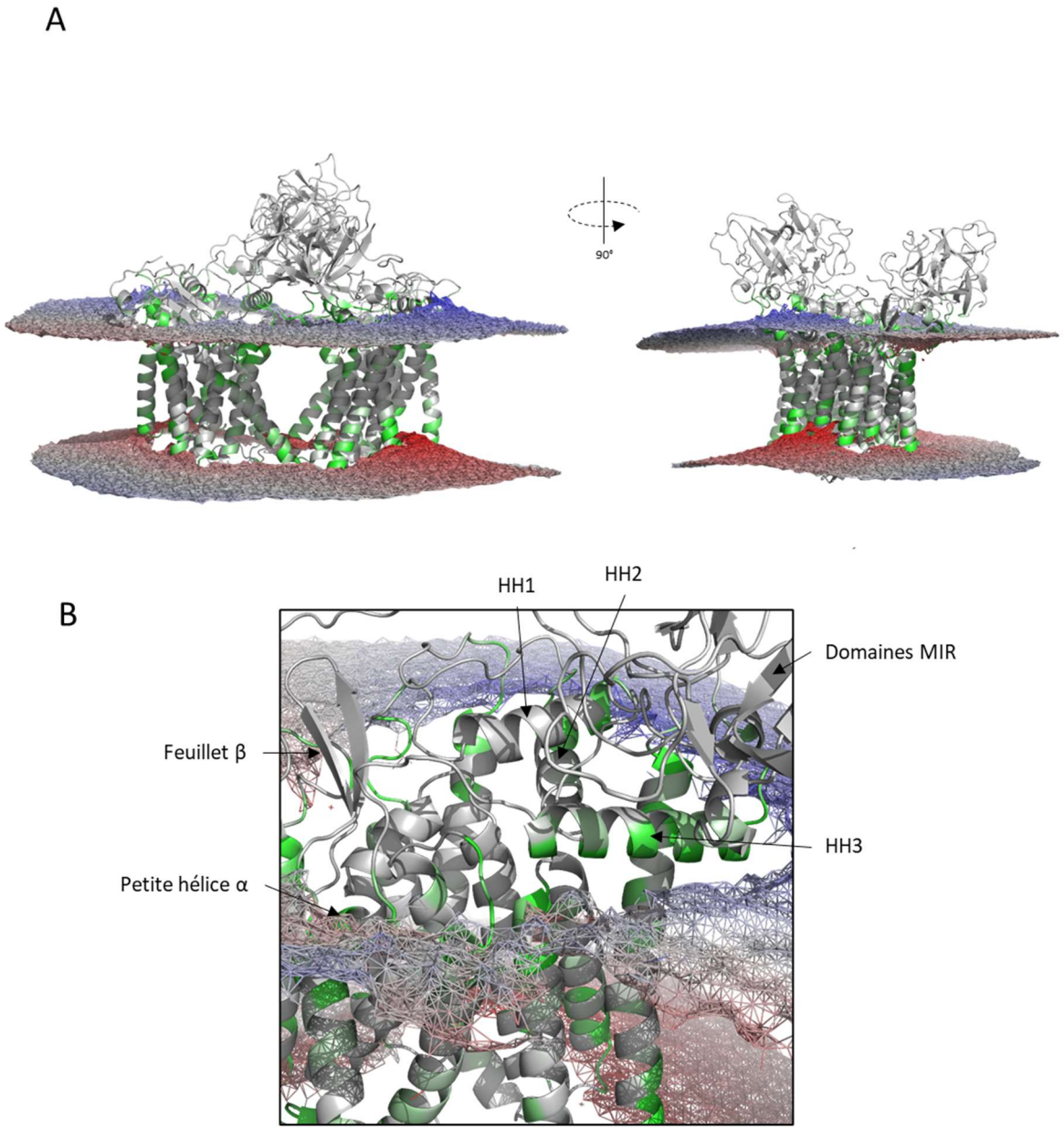


Figure 39. Distorsion de la bicouche lipidique autour de l'hétérocomplexe *Sc*PMT1-*Sc*PMT2. Les déformations sont représentées sous forme de maillage sur une échelle rouge-blanc-bleu, en rouge est représenté un amincissement de la membrane, alors qu'en bleu est représenté un épaississement. La coloration verte des protéines représente les acides aminés en contact avec les têtes polaires des lipides de la membrane simulée par MemProtMD[175]. **A)** Hétérocomplexe *Sc*PMT1-*Sc*PMT2 dans une membrane simulée [88,175] vu sous deux angles . **B)** Grossissement au niveau de HH3 de *Sc*PMT1.

sucres sur les structures des domaines MIR. Ces sites d'interaction ont été démontrés comme capables d'interagir avec du mannose et des peptides O-mannosylés[87]. En outre, il a pu être mis en évidence une différence de comportement dans l'interaction d'un peptide mannosylé avec les domaines MIR de différentes PMT. En effet, une divergence a pu être notée entre les domaines MIR de *Sc*PMT2 (sous-famille PMT2) et de *Sc*PMT5 (sous-famille PMT1). Cette variabilité de reconnaissance du peptide *O*-mannosylé pourrait sous-entendre un rôle des domaines MIR dans la spécificité des PMT de différentes sous-familles vis-à-vis de certains substrats[87].

Les données fonctionnelles acquises sur les domaines MIR ont également permis d'apporter une meilleure compréhension de la structure de l'hétérocomplexe *Sc*PMT1-*Sc*PMT2[87,88]. En effet, les auteurs émettent l'hypothèse qu'il pourrait exister une interaction dynamique entre le domaine MIR d'une PMT et le site actif de la PMT avec laquelle elle forme un dimère[87]. Dans l'héterodimère *Sc*PMT1-*Sc*PMT2, le domaine MIR de *Sc*PMT2 se localise de manière à pouvoir lier le peptide mannosylé ce qui permettrait d'organiser le site actif de *Sc*PMT1, et inversement. Le domaine MIR de chaque enzyme complémenterait ainsi le site actif de l'autre en jouant un rôle dans la régulation de la réaction[87,88].

PMTub est dépourvue de domaines MIR, ce qui rend sa boucle EL4 plus courte (**Figure 38**). L'absence de ces domaines chez les bactéries pourrait alors se traduire par des mécanismes différents de régulation. Les modélisations 3D de PMTub définissent une séquence non structurée localisée en lieu et place du domaine MIR présent chez les eucaryotes (**Figure 38B**).

Une autre structure retrouvée dans la boucle EL4 des PMT eucaryotes, et cette fois-ci prédite comme présente dans la structure de PMTub, par modélisation, est l'hélice horizontale HH3 (**Figure 38**). Cette hélice est décrite comme participant à l'interaction avec le substrat donneur dans le cas de *Sc*PMT1[88]. Elle est structurée même en l'absence du lipide donneur ce qui pourrait lui permettre de jouer un rôle dans la sélection des substrats pouvant accéder au site actif. De ce fait HH3 pourrait empêcher par exemple l'entrée d'un lipide-phosphate lié à des groupements saccharidiques plus imposants qu'un mannose.

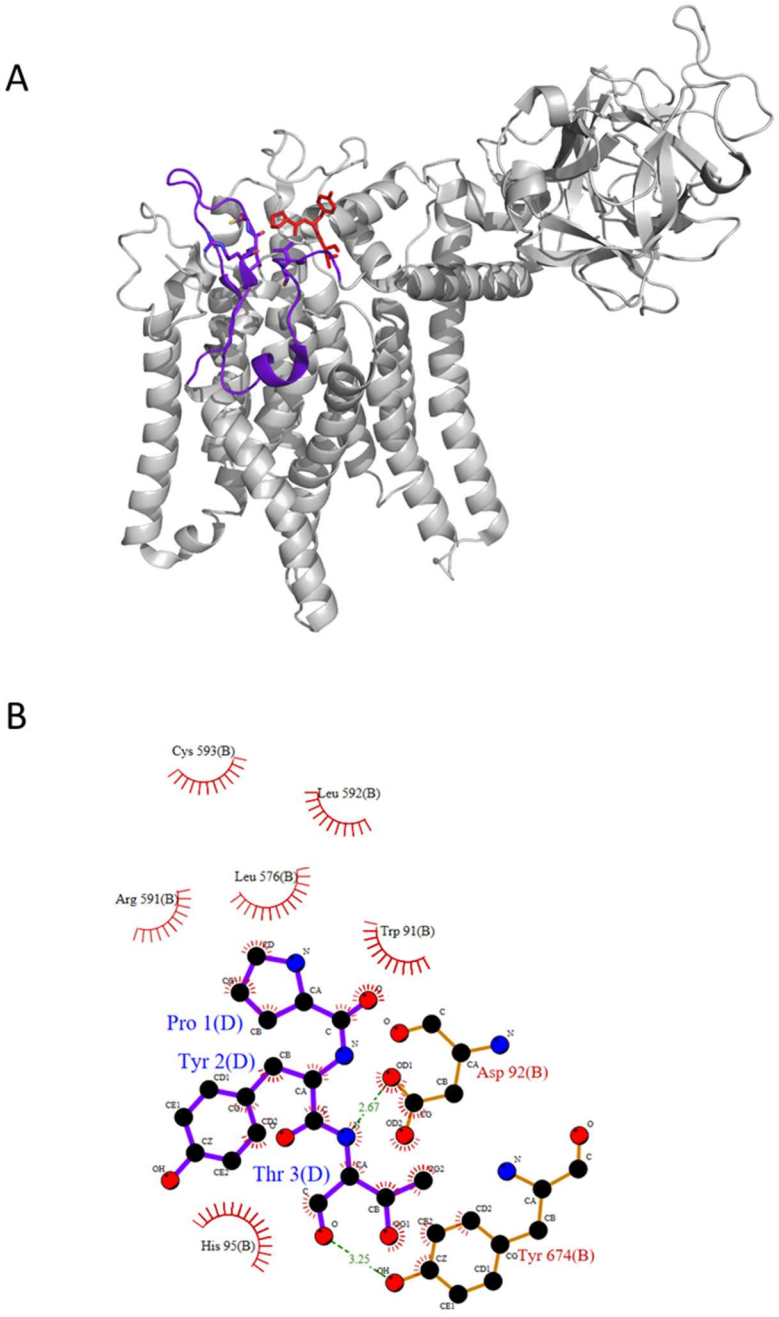


Figure 40. Positionnement du peptide accepteur dans *Sc*PMT2. A) Structure de *Sc*PMT2 résolue avec un substrat peptidique[88]. Le substrat (PYT) est coloré en rouge dans la structure. L'extrémité terminale de EL4, composée d'une petite hélice α et d'un feuillet β, est colorée en violet. Sur cette extrémité terminale, 4 acides aminés sont représentés en bâton, ils correspondent aux acides aminés interagissant avec le peptide selon l'analyse ligplot[176]. **B)** Représentation des interactions entre ScPMT2 et le peptide accepteur, par Ligplot[176]. Les interactions ioniques sont représentées par des lignes vertes pointillées et les interactions hydrophobes sont représentées par des arcs de cercle avec des pics dirigés vers l'atome avec lequel ils interagissent. L576, R591, L592 et C593 font partie de la boucle EL4. Ces acides aminés sont représentés en bâton sur la représentation 3D de la figure 2.3A et sont mis en évidence dans l'alignement de séquence de la Figure 41.

Ce type de mode d'action est déjà supposé pour d'autres GT-C, comme la glucosyltransférase de levure *Sc*ALG6[123]. En effet, pour cette enzyme une hélice horizontale similaire a été décrite[123]. Ainsi, l'hélice HH3 de *Sc*ALG6 forme un arc au-dessus du site d'insertion du substrat lipidique donneur, résultant en une entrée en forme d'entonnoir. Cette structuration en entonnoir est accentuée par une légère déformation de la surface membranaire à cet endroit. Cet entonnoir pourrait participer à la sélection du lipide-phosphate portant le bon sucre et à son maintien dans le site actif jusqu'à transfert du sucre sur le substrat accepteur[123]. Une simulation de dynamique moléculaire d'une membrane plasmique autour de l'hétérocomplexe *Sc*PMT1-*Sc*PMT2 montre aussi une légère variation dans l'épaisseur de la membrane au niveau de HH3, suggérant une possible structuration en entonnoir et donc un possible rôle de HH3 dans la sélection et le maintien du substrat donneur[88,175](**Figure 39**).

Finalement, nous observons à l'extrémité terminale de EL4 un motif composé d'une hélice α, en contact avec la membrane, suivie d'un feuillet β (**Figure 38 et Figure 39**). Notablement, le rôle de ce motif n'a jamais été étudié pour les PMT. Une hypothèse peut néanmoins être émise, basée sur les observations faites sur la structure de l'oligosaccharyltransférase bactérienne, PglB[116,177] et sur la structure de *Sc*PMT2. Tout d'abord, pour PglB, le domaine correspondant (par alignement) est décrit comme le site d'interaction du substrat protéique accepteur[116,177]. De plus, on peut noter sur la structure de *Sc*PMT2 que le peptide accepteur se situe à proximité de cette structure en feuillet β (**Figure 40A**). Une analyse de la structure 3D de *Sc*PMT2 permet, en effet, de mettre en évidence des contacts hydrophobes possibles entre des acides aminés de ce motif, et le peptide substrat (**Figure 40B**). Nous avons donc formulé l'hypothèse que ce motif, présent dans la boucle EL4 de PMTub, jouerait un rôle dans l'interaction avec le site de mannosylation sur le substrat protéique accepteur. De façon remarquable, la présence de plusieurs acides aminés fortement conservés dans cette zone de la boucle EL4 suggère un rôle dans l'activité de l'enzyme et renforce cette hypothèse (**Figure 41A**).

Tous ces éléments observables semblent donc indiquer un rôle potentiel de cette boucle EL4 dans l'interaction avec les substrats et le produit de réaction des PMT, sans que cela n'ait jamais été formellement démontré. Nous émettons les hypothèses que la structure

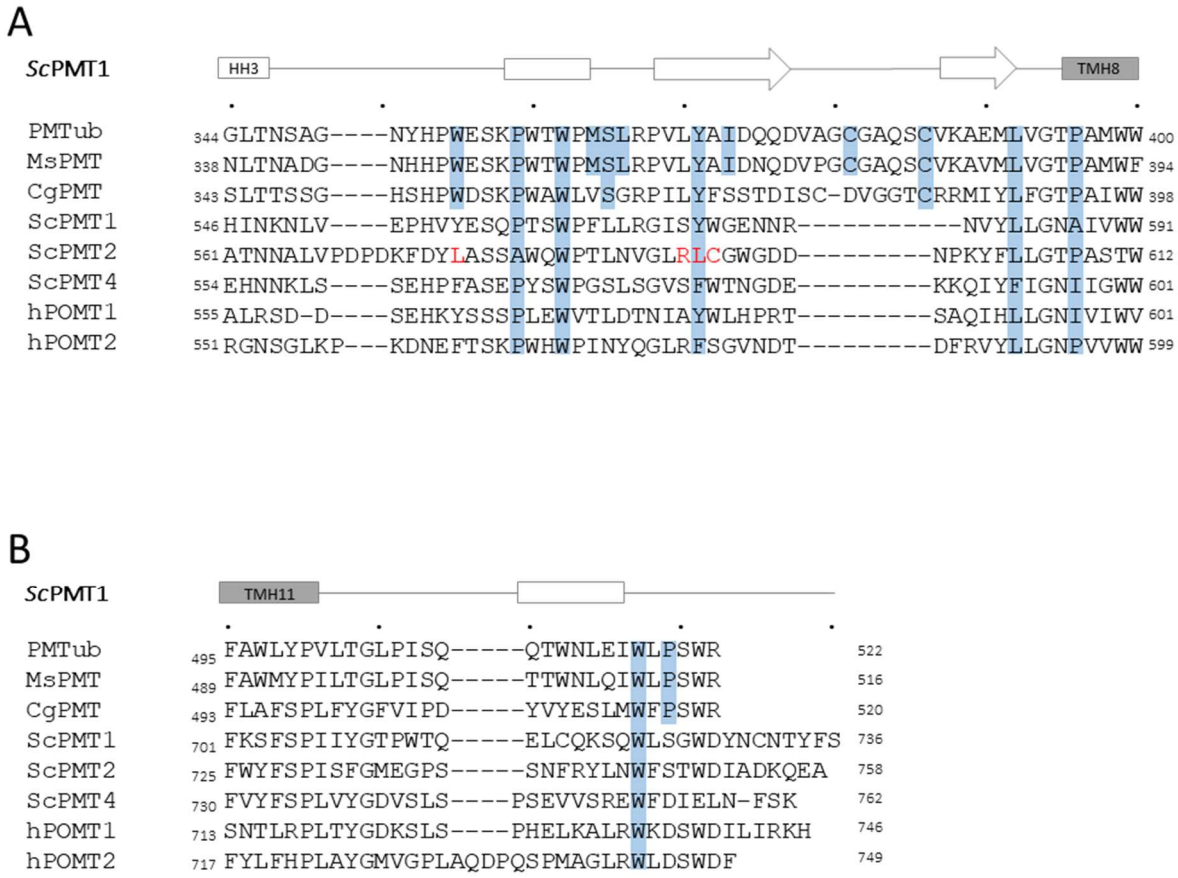


Figure 41. Alignement de séquence des domaines fonctionnels prédits de plusieurs PMT. A) Alignement de séquences de l'extrémité C-terminale de la boucle EL4 de plusieurs PMT. Les structures secondaires de ScPMT1 pour cette zone sont représentées, les rectangles représentent des hélices α et les flèches des brins β, TMH correspond à une hélice transmembranaire. En bleu sont surlignés les acides aminés qui ont été mutés au cours de cette étude. Les acides aminés représentés en rouge dans la séquence de ScPMT2 sont les acides aminés identifiés par l'étude Ligplot (voir Figure 40B). **B)** Alignements de séquence de l'extrémité C-terminale de plusieurs PMT. Les structures secondaires de *Sc*PMT1 pour cette zone sont représentées. En bleu sont surlignés les acides aminés qui ont été mutés durant cette étude.

en feuillet β pourrait participer à l'interaction avec le substrat accepteur, alors que l'hélice HH3 participerait à l'interaction avec le substrat donneur[88]. Les domaines MIR seraient quant à eux en mesure d'interagir avec le produit de réaction[87]. L'étude fonctionnelle de cette boucle nous est donc apparue primordiale pour mieux comprendre les mécanismes de la réaction catalysée par les PMT.

b) Extrémité C-terminale

L'autre portion de séquence d'intérêt est l'extrémité C-terminale. En effet, dans PglB, l'extrémité C-terminale (définissant un domaine soluble terminal protubérant) est également impliquée dans l'interaction avec le substrat protéique[116]. De plus, il existe une forte disparité de longueur de cette extrémité dans les PMT. Chez la levure, l'extrémité C-terminale peut représenter jusqu'à 100 acides aminés en fonction de l'isoforme. Chez l'homme, elle correspond à 29 et 28 acides aminés pour hPOMT1 et hPOMT2 respectivement. Pour PMTub, cette extrémité ne représente que 22 acides aminés, et se situe à proximité de la structure en feuillet β de EL4 (**Figure 38B**). La présence de plusieurs acides aminés très conservés, tels que des prolines et tryptophanes, suggère un motif structural pouvant jouer un rôle fonctionnel dans l'activité enzymatique (**Figure 41B**).

2. Caractérisation du site d'interaction prédit avec le substrat protéique

Le rôle prédit pour la boucle EL4 des PMT eucaryotes est donc multiple : interaction avec les deux substrats et le produit. Alors que la fonction des domaines MIR et de l'hélice HH3 ont déjà été abordés dans différents articles[87,88], celui de la structure en feuillet *β* n'a quant à lui jamais été examiné. Pour PMTub, les domaines MIR ne sont pas présents, de ce fait la boucle EL4 pourrait endosser moins de rôles que ses homologues eucaryotes. Néanmoins, l'hélice HH3 et la structure en feuillet β sont toutes deux présentes selon notre modèle 3D.

Notre hypothèse est que le motif en feuillet β à l'extrémité terminale de EL4 participerait à l'interaction avec le substrat accepteur, de concert avec l'extrémité C-terminale de PMTub. Sur le modèle 3D de PMTub, ces deux domaines qui se font suites structuralement

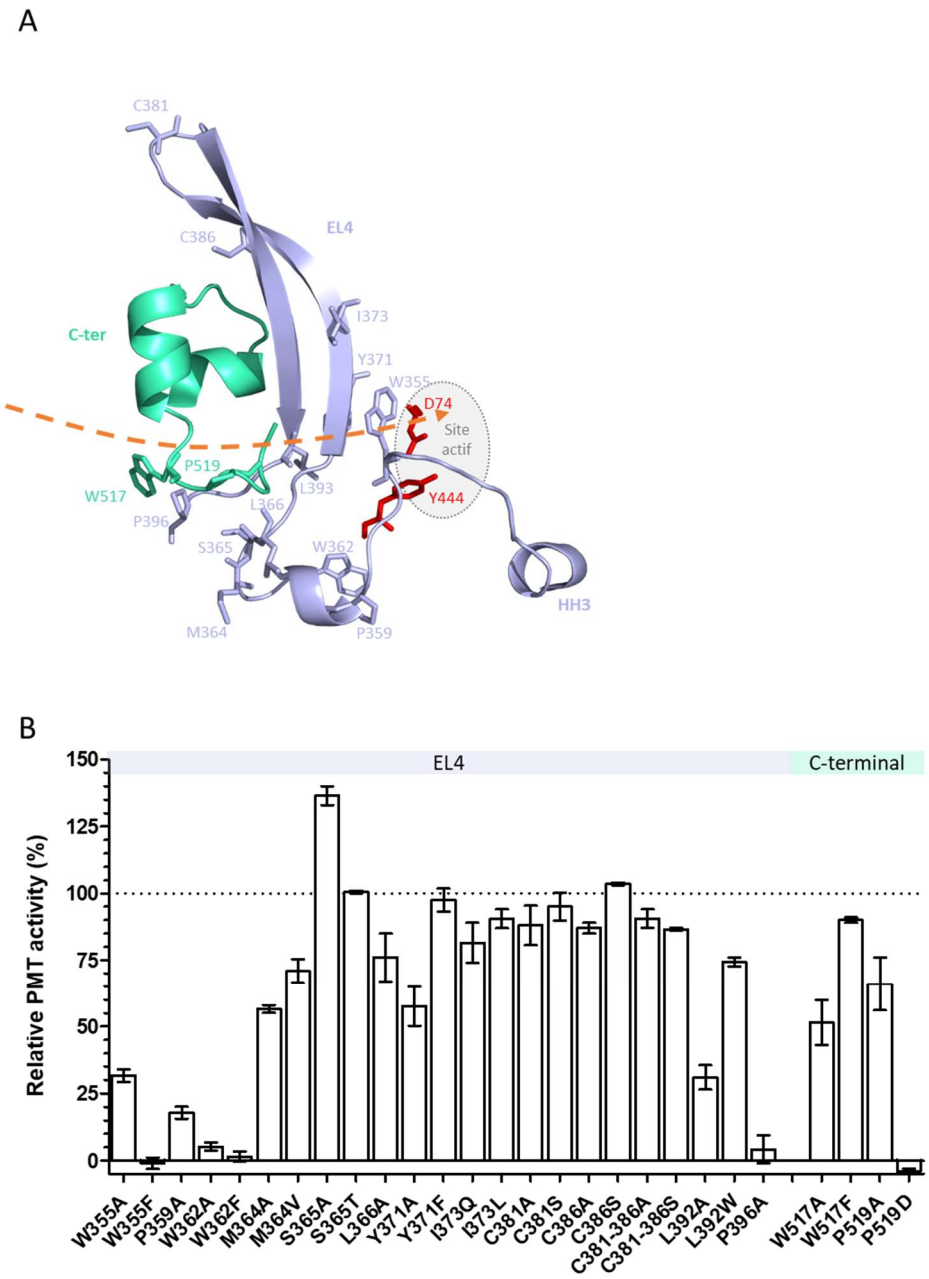


Figure 42. Implication de EL4 et de l'extrémité C-terminale de PMTub. A) Structure du modèle 3D du site d'interaction avec les substrats accepteurs de PMTub. Les chaines latérales des acides aminés mutés et évalués sont présentées et colorées. Le site actif de PMTub est représenté par un ovale gris, les acides aminés D74 et Y444 composant le site d'interaction avec la sérine/thréonine acceptrice sont représentés en rouge. La flèche en pointillé orange représente le chemin que pourrait emprunter le substrat protéique lorsqu'il interagit avec PMTub. **B)** Résultats du test d'activité phénotypique pour des mutants de PMTub. 0% et 100% sont définis comme les résultats pour la souche « No PMT » et pour la souche exprimant PMTub sauvage respectivement. Les résultats présentés sont des triplicats techniques de duplicats biologiques.

semblent former « une rampe » se terminant au niveau du site actif de PMTub (**Figure 42A**). Cette « rampe » pourrait donc constituer le site d'interaction avec le substrat protéique. Pour valider cette hypothèse, nous avons effectué une étude de ces domaines par mutagenèse dirigée dans la séquence de PMTub.

De façon notable, ces structures affichent un intérêt tout particulier puisqu'elles montrent une divergence de séquence avec les PMT eucaryotes et pourraient donc constituer, dans le futur, la cible d'inhibiteurs spécifiques de PMTub. Les positions mutées ont été choisies en fonction de leur spécificité bactérienne ou bien au contraire de leur conservation avec les PMT eucaryotes, ainsi que de leur positionnement clé sur le modèle 3D de PMTub (**Figure 41 et Figure 42**).

a) Boucle « EL4 »

- W355

La position W355 est spécifique des bactéries. Sur le modèle de PMTub, cet acide aminé est situé sur une boucle non structurée et vient se positionner face au feuillet β et à proximité du site actif (**Figure 41A et Figure 42A**). De plus, l'analyse de la structure 3D de *Sc*PMT2 permet de mettre en évidence une possible interaction hydrophobe avec le substrat de l'acide aminé équivalent (L576) (**Figure 40B et Figure 41A)**. La mutation de cette position entraine une forte diminution du taux de mannosylation de la FasC démontrant son implication dans l'activité protéine *O*-mannosyltransférase de PMTub, sans pour autant affecter l'expression de l'enzyme à la membrane plasmique (**Figure 42B** et **Figure 43**). Du fait de sa position et de son rôle essentiel, cette position pourrait avoir un rôle dans l'interaction avec le substrat protéique avant son entrée dans le site actif.

- P359 et W362

Deux acides aminés fortement conservés présents dans l'hélice α en amont du feuillet β ont aussi été analysés : P359 et W362 (**Figure 41A et Figure 42A**). La mutation ponctuelle de ces deux positions en alanine entraine une forte diminution de la mannosylation de la FasC (**Figure 42B**). La mutation de W362 en un autre acide aminé aromatique, une phénylalanine,

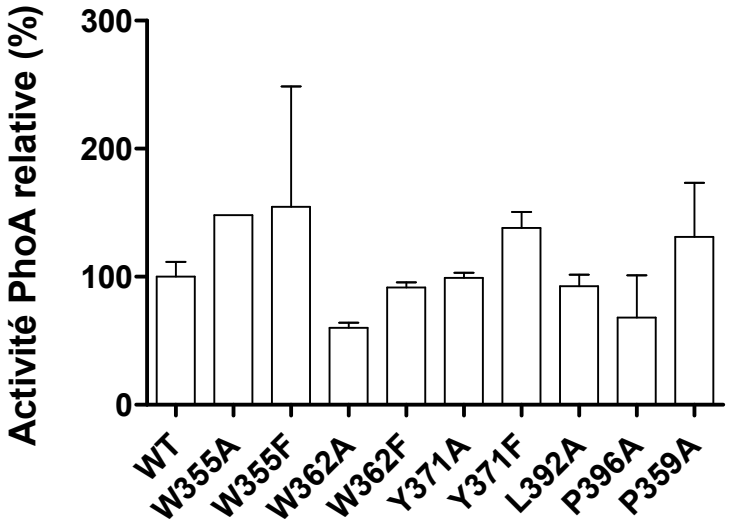


Figure 43. Activité PhoA relative mesurée pour les mutants non fonctionnels de PMTub. L'activité PhoA est mesurée en relatif à l'activité de la souche exprimant PMTub-PhoA (défini comme 100% arbitrairement). Les résultats présentés sont des triplicats techniques de duplicats biologiques.

ne permet pas de restaurer l'activité mannosyltransférase de PMTub (**Figure 42B**). Ces résultats traduisent le rôle fondamental de ce tryptophane dans l'activité de l'enzyme. Ces deux acides aminés très conservés parmi les PMT et qui délimitent l'hélice α en amont du feuillet β ont probablement un rôle dans le maintien de la conformation de cette hélice, ce qui doit permettre la cohérence structurale de la boucle EL4. Du fait du rôle central prédit de cette boucle EL4 dans l'interaction avec les différents substrats, sa déstructuration doit prévenir l'activité *O*-mannosyltransférase.

- M364, S365 et L366

Entre l'hélice α qui semble donc primordiale dans l'activité PMTub, *via* P369 et W362, et le feuillet β, se trouvent plusieurs acides aminés conservés chez les mycobactéries, et qui ne sont pas retrouvés dans les boucles EL4 des autres PMT (**Figure 41A et Figure 42A**). Du fait de la localisation de ces résidus à l'interface entre les deux structures secondaires prédites et de leur faible conservation avec les PMT eucaryotes, nous avons voulu déterminer s'ils avaient une fonction dans l'activité de PMTub. Les positions M364, S365 et L366 ont été mutées en alanines n'entrainant aucune perte significative de la mannosylation de la FasC (**Figure 42B**). De manière intéressante, on peut observer une légère hausse de l'activité mannosyltransférase lors de la mutation S365A, qui retrouve un niveau similaire à l'enzyme sauvage pour la mutation S365T. Cette tendance devra être validée ultérieurement par un plus grand nombre de réplicats permettant de réaliser les tests statistiques nécessaires à sa confirmation. Néanmoins, ces 3 acides aminés non conservés avec les eucaryotes ne sont pas essentiels à l'activité de PMTub.

- Y371

Nous avons ensuite identifié Y371 comme position intéressante. En effet, cette tyrosine se situe sur premier brin du feuillet β selon notre modèle, et cette zone du feuillet β pourrait participer à l'interaction avec le substrat accepteur selon la structure de *Sc*PMT2 (**Figure 40B et Figure 42 A**). On trouve à cette position dans la majorité des PMT une tyrosine ou une phénylalanine, nous avons donc effectué les mutations Y371A et Y371F (**Figure 41A**). La mutation en alanine entraine une diminution de moitié de l'activité de PMTub qui est

restaurée par la mutation en phénylalanine (**Figure 42B**). Ces résultats démontrent le rôle essentiel d'un acide aminé aromatique à cette position, rôle qui peut néanmoins être partiellement rempli par un acide aminé hydrophobe non aromatique.

- I373

Un second acide aminé également situé sur le premier brin du feuillet β selon notre modèle a été aussi étudié : I373. Cet acide aminé est conservé chez les mycobactéries seulement et est proche de Y371 qui participe à l'activité de PMTub (**Figure 41A et Figure 42A**). Nous avons muté cet acide aminé spécifique des mycobactéries en alanine et en glutamine. I373A et I373Q n'ont pas entrainé de variation de l'activité mesurée comparée à PMTub sauvage, ne permettant pas de mettre en évidence de rôles pour cet acide aminé spécifique des mycobactéries (**Figure 42B**).

- C381 et C386

Sur un alignement de séquence, une insertion spécifique des bactéries peut être notée à l'interface entre les deux brins β décrits chez les PMT de levure (**Figure 41A**). Cette insertion compte une dizaine d'acides aminés, dont deux cystéines (C381 et C386) (**Figure 41A et Figure 42A**). La présence de ces deux cystéines à proximité du site prédit d'interaction avec le substrat accepteur nous a menés à formuler l'hypothèse d'un possible rôle de celles-ci dans l'interaction avec certains substrats protéiques, à l'instar de ce qui est observé pour le processus de N-glycosylation chez les eucaryotes. En effet, la N-glycosylation des protéines chez les eucaryotes est réalisée par le complexe oligosacharryltransférase (OST). Dans ce complexe, on trouve la protéine OST6 (ou son paralogue OST3) qui a été défini comme possédant un motif de deux cystéines jouant un rôle dans l'interaction avec certains substrats protéiques du complexe OST[178]. Il a été montré que la mutation de ce motif de deux cystéines dans OST6 affectait la glycosylation de protéines portant des cystéines proches de leur site de glycosylation. Il a donc été proposé que OST6 forme des ponts disulfures transitoires avec certains substrats pour améliorer leur efficacité de glycosylation[178]. L'étude du mannoprotéome mycobactérien montre qu'il est possible de trouver des cystéines à proximité des sites de mannosylation[38], un mécanisme similaire peut donc être envisagé.

La mutagenèse ponctuelle de ces deux cystéines en alanine ou en sérine sur la séquence de PMTub n'a pas entrainé de perte de mannosylation de la FasC (**Figure 42B**). Pour éliminer la possibilité que la perte d'une des cystéines soit compensée par l'autre, nous avons aussi effectué les doubles mutations C381-386A et C381-386S. Ces doubles mutations n'entrainent pas, non plus, de perte de mannosylation de la FasC (**Figure 42B**). Il semblerait donc que ces deux acides aminés spécifiques des bactéries n'aient pas de rôles dans la mannosylation de la FasC (qui ne possède pas de cystéines à proximité de ses sites de mannosylation) par PMTub. Il n'est cependant pas à exclure que ces positions jouent un rôle dans la reconnaissance d'autres mannoprotéines.

- L392

Sur le second brin du feuillet β selon notre modèle, face à Y371, se trouve une leucine fortement conservée dans les PMT (**Figure 41A et Figure 42A**). Cette leucine correspond à L392 dans la séquence de PMTub. Nous avons effectué les mutations L392A et L392W. La mutation en alanine entraine une diminution forte de l'activité mesurée, et la mutation en tryptophane impacte peu l'activité, rappelant les résultats obtenus pour Y371. Il semblerait donc que cette position soit importante pour l'activité PMTub (**Figure 42B**).

- P396

Finalement, nous avons étudié la proline P396 qui marque la fin de la boucle EL4 et le début de l'hélice transmembranaire (**Figure 41A et Figure 42A**). P396 est alignée dans l'espace avec les positions W362 et P359, que nous avons démontrées comme essentielles à l'activité et pourrait, comme ces deux positions, participer au maintien structural de la boucle EL4 (**Figure 42A**). La mutation P396A entraine une perte totale de l'activité mesurée pour PMTub sans affecter son expression membranaire (**Figure 42B**). Cette proline joue donc aussi un rôle central dans l'activité de PMTub.

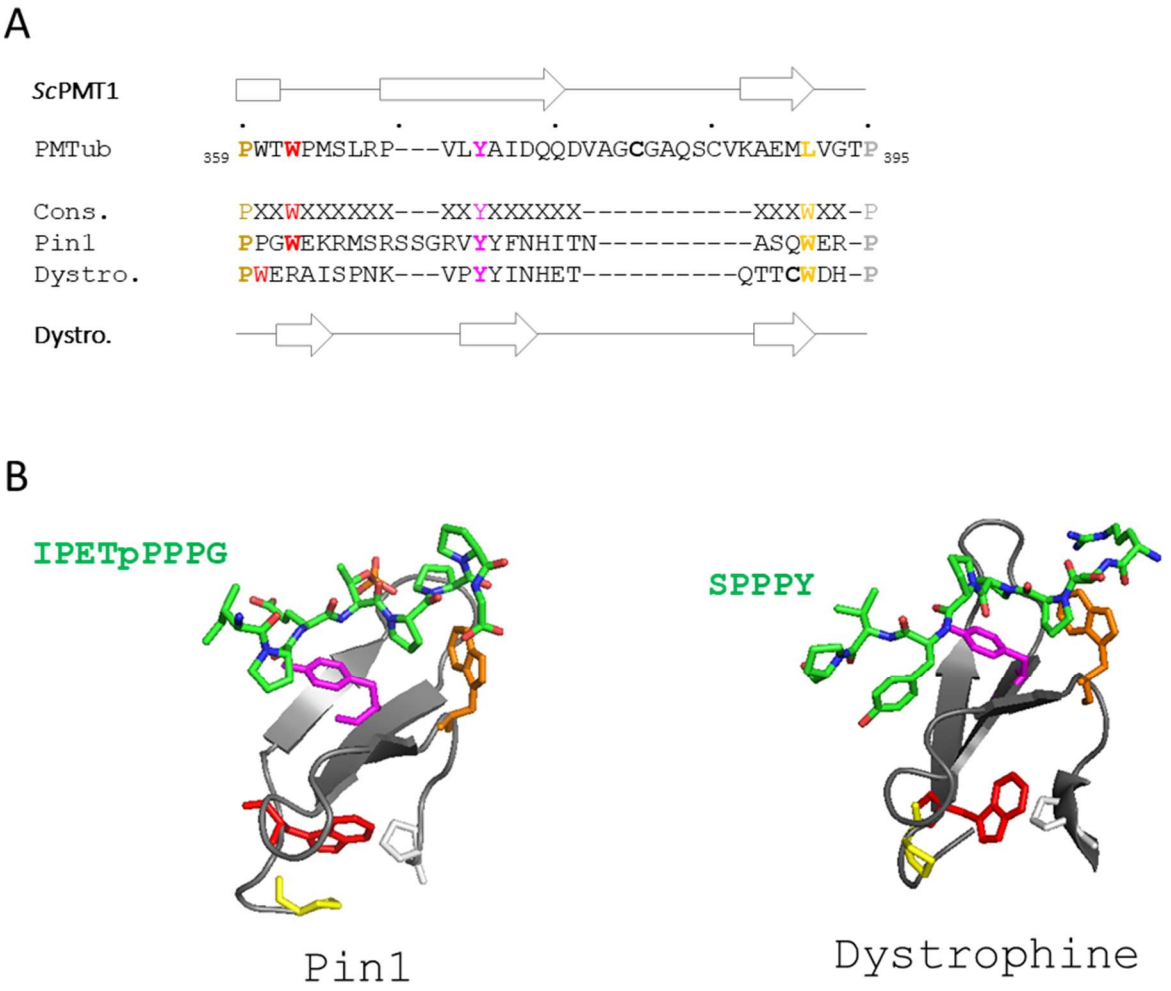


Figure 44. EL4 comme « pseudo-domaine WW ». A) Alignement de séquence entre EL4 et les domaines WW de la dystrophine (Dystro.)[179] et Pin1[180]. La séquence consensus (Cons.) des domaines WW est représenté avec chacun des 5 acides aminés colorés différemment[181]. Les structures secondaires de *Sc*PMT1 et du domaine WW de la dystrophine sont représentées, les rectangles représentent des hélices α et les flèches des brins β. **B)** Structure 3D des domaines WW de la dystrophine[179] et de Pin1[180] en interaction avec leur substrat riche en proline. Les 5 acides aminés consensus sont mis en évidence et colorés comme dans les alignements de séquence. Le peptide interactant est représenté en vert et sa séquence est annoté dans la même couleur.

a) Extrémité C-terminale

Concernant l'extrémité C-terminale, nous avons étudié une proline et un tryptophane que l'on suppose participer au site d'interaction avec le substrat protéique (**Figure 42A**). Le tryptophane W517 est conservé chez les PMT alors que la proline P519 est spécifique des mycobactéries (**Figure 41B**).

- W517

Pour W517, la mutation en alanine divise par deux l'activité et la mutation en phénylalanine n'affecte pas l'activité (**Figure 42B**). Ainsi, il semble que ce tryptophane, bien que non essentiel, participe à l'activité par ses propriétés aromatiques.

- P519

En raison de sa conservation dans les mycobactéries, nous avons muté P519 en alanine et acide aspartique (comme trouvé dans les PMT humaines à la position correspondante, **Figure 41B**). Fait intéressant, la mutation en alanine affecte peu l'activité de PMTub, mais la mutation en acide aspartique supprime totalement la *O*-mannosylation de la FasC (**Figure 42B**). Nous avons émis l'hypothèse que cet effet soit dû à l'intégration d'un acide aminé chargé (l'acide aspartique) perturbant la structure de l'enzyme ou l'interaction avec le substrat protéique. Des investigations supplémentaires seront donc nécessaires pour conclure sur le rôle de l'extrémité C-terminale dans l'activité de PMTub et pour déterminer comment l'introduction d'une charge (par une mutation ou un inhibiteur putatif) pourrait altérer l'activité de l'enzyme.

3. EL4 comme "pseudo-WW domain"

a) Similarité avec les domaines WW

Par l'étude de mutagenèse, nous avons donc pu mettre en évidence un rôle de EL4 dans l'activité de PMTub avec notamment plusieurs positions pour lesquelles une mutation a un effet sur l'activité : W355, P359, W362, Y371, P396 et L392. Une caractéristique intéressante de l'extrémité C-terminale de EL4 de PMTub est sa similarité de séquence avec

un domaine WW (30% d'identité et 52% d'homologie avec le domaine WW de la dystrophine, **Figure 44**) et que les acides aminés que nous avons identifiés comme importants pour le rôle de EL4 sont similaires aux acides aminés consensus des domaines WW.

Les domaines WW (famille pfam : PF00397) sont les plus petits modules protéiques dédiés à l'interaction protéine-protéine. On les retrouve dans plus de 15000 protéines eucaryotes et ils n'ont jamais été identifiés chez la bactérie. Des études précédentes ont permis de définir cinq acides aminés consensus des domaines WW : un couple proline-tryptophane espacé de deux acides aminés à l'extrémité N-terminale (PxxW) du domaine ; un couple tryptophane-proline espacé de deux acides aminés à l'extrémité C-terminale (WxxP) du domaine ; et une tyrosine centrale (**Figure 44A**)[181]. Ces acides aminés ne sont cependant pas tous strictement nécessaires : certains domaines WW ne possèdent pas la proline initiale et d'autres le tryptophane terminal[182].

Ces différents acides aminés consensus ont tous un rôle défini dans la structuration et le rôle de ces domaines. En effet, les domaines WW suivent tous le même modèle structural : ce sont des modules en feuillet avec trois brins β antiparallèles formant un plan légèrement courbé dans lequel vient se loger leur interactant protéique. Les deux prolines délimitent le domaine et s'agencent autour du tryptophane de l'extrémité N-terminale permettant de stabiliser la structure en feuillet β du domaine (**Figure 44B**). Les deux autres acides aminés consensus, la tyrosine et le tryptophane terminal, ont pour rôle d'interagir avec le substrat. Les domaines WW interagissent avec des séquences riches en proline[181]. La tyrosine et le tryptophane consensus vont donc former une « pince à prolines » autour des prolines de l'interactant (**Figure 44B**).

Sur notre modèle 3D de PMTub, l'extrémité terminale de EL4 n'est pas prédite comme structurée comme un feuillet β antiparallèle à 3 brins, mais comme deux brins β antiparallèles et une hélice α située légèrement en amont (position équivalente au premier brin des motifs WW) (**Figure 42A**). Néanmoins, EL4 possède quatre des cinq acides aminés consensus des domaines WW (P359, W362, Y371 et P394) (**Figure 44A**). Ces caractéristiques peuvent laisser penser que la boucle EL4 pourrait donc agir comme un « pseudo domaine WW ».

Quatre des cinq acides aminés consensus sont retrouvés dans EL4 et notre étude a permis de leur attribuer un rôle important, voire essentiel, dans l'activité de PMTub (**Figure 42B et Figure 44**). Parmi ces acides aminés, on retrouve le couple proline-tryptophane espacé de deux acides aminés à l'extrémité N-terminale (P359 et W362) ainsi que la proline terminale (P396) qui, pour rappel, jouent un rôle dans la stabilisation structurale des domaines WW, à la manière d'un socle hydrophobe. Sur le modèle 3D de PMTub, ces trois acides aminés se retrouvent alignés au niveau de la surface de la membrane plasmique, à la base du feuillet β (**Figure 42A**). Le caractère essentiel de ces positions pourrait donc également être expliqué par un rôle dans la stabilisation du feuillet β, par l'ancrage de celui-ci à la membrane.

Enfin, nos résultats montrent également la contribution des résidus Y371 et L392 qui correspondent respectivement à la tyrosine centrale et au tryptophane terminal consensus des domaines WW, participant à la « pince à prolines ». Ces deux acides aminés semblent impliqués au travers du caractère hydrophobe de leur chaine latérale, puisque leur mutation en résidus aromatiques entraine une restauration totale d'activité (**Figure 42B**). Ces résultats suggèrent qu'ils pourraient contribuer à la séquestration du substrat protéique accepteur au cours de la catalyse, probablement par stacking avec des prolines, ou par interaction hydrophobe, puisque les sites de mannosylation des protéines mycobactériennes sont des thréonines ou des sérines entourées d'une zone riche en proline et alanine[38].

Cette extrémité terminale de la boucle EL4 adopterait donc une structuration légèrement différente des domaines WW, mais pourrait fonctionner selon un mécanisme similaire à ceux-ci, à la différence supplémentaire que l'un des tryptophanes consensus des domaines WW est une leucine dans la boucle EL4 de PMTub. Dans les domaines WW, les prolines du peptide substrat sont retenues par stacking de la « pince à prolines », formée par les cycles aromatiques de la tyrosine et du tryptophane consensus. Dans EL4 ce mécanisme serait donc un peu différent si l'on considère la leucine L392 comme participant à la « pince à prolines ». Une autre possibilité est que le stacking des prolines se produise entre la tyrosine Y371 et le tryptophane W355 voisin dans l'expace. Nous avons démontré que ce tryptophane était impliqué dans l'activité de PMTub, par notre étude par mutagenèse, et celui-ci est localisé en face de la tyrosine Y371 dans le modèle 3D de PMTub (**Figure 42**).

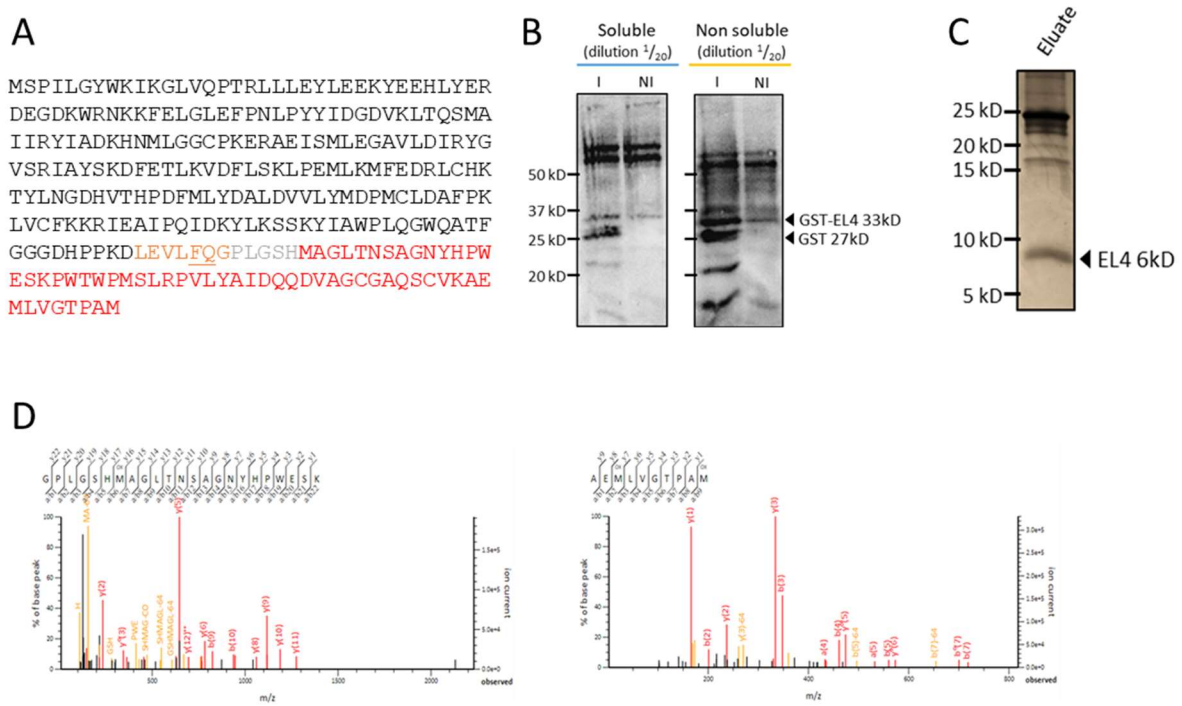


Figure 45. Production de la boucle EL4 isolée dans *E. coli*. A) Séquence de la protéine de fusion GST-EL4. La séquence en noir correspond au tag GST, en orange la séquence reconnue par la protéase spécifique (PreScission), en rouge la séquence d'une partie de EL4. Le site de clivage spécifique est localisé entre les deux acides aminés soulignés. **B)** Western blot anti-GST de la production de la protéine de fusion GST-EL4, réalisé sur les fractions solubles et non solubles obtenues après lyse et centrifugation de la culture de *E. coli* avec induction (I) ou sans induction (NI) par de l'IPTG pendant 3h à 16°C. **C)** Coloration au nitrate d'argent après purification et clivage par la protéase spécifique de EL4. **D)** Analyse par spectrométrie de masse de EL4 après digestion sur gel par la trypsine, permettant de détecter les 23 premiers acides aminés (spectre de gauche) ainsi que les 10 derniers (spectre de droite).

Le feuillet β formerait ainsi « une rampe », permettant d'accueillir le substrat protéique et de le retenir de sorte à orienter le site de mannosylation vers la poche catalytique identifiée dans la partie 2.B de ce travail (voir p.119 à 127). Du fait de la position de l'extrémité C-terminale de PMTub sur le modèle 3D, de son rôle dans d'autres GT-C et de nos résultats de mutagenèse, on peut imaginer que cette extrémité participe aussi à la rétention du substrat dans la structure de PMTub (**Figure 42**).

L'ensemble de ces observations renforcent l'hypothèse que EL4 participerait à l'interaction avec le substrat protéique à proximité du site de mannosylation et qu'une partie de la boucle EL4 pourrait se comporter comme un « pseudo domaine WW ». Pour étudier ce comportement avec plus de précision et pour apporter des éléments de réponse à l'hypothèse formulée pour l'interaction entre PMTub et ses substrats, nous avons mis en place plusieurs stratégies expérimentales se basant sur des systèmes d'interaction *in vitro* ou des stratégies innovantes utilisant des outils de biologie synthétique.

b) Étude in vitro du "pseudo-WW domain" EL4

Nous avons cherché à suivre l'interaction entre EL4 et le substrat protéique à l'échelle moléculaire en utilisant un test d'interaction *in vitro*. En raison de la taille et de la topologie membranaire de PMTub, ces tests ne peuvent pas être effectués sur l'enzyme entière. Une caractéristique des domaines WW est qu'ils sont des modules de repliement autonomes, c'est-à-dire que lorsqu'ils sont produits isolés en solution, ils sont capables de se replier correctement et d'interagir avec leurs substrats riches en proline[182]. En raison de l'homologie de séquence élevée, nous nous attendions à ce que l'extrémité terminale de EL4 agisse également tel un module de repliement autonome. Une partie de cette boucle pourrait alors être produite, isolée en solution, et ses capacités à interagir avec le substrat protéique étudiées par la suite.

(1) Production du domaine EL4.

La séquence à produite fait une taille de 56 acides aminés, correspondant à la structure décrite précédemment, prédite pour être arrangée en deux brins bêta précédés par une petite hélice α, et contenant l'ensemble des acides aminés identifiés comme impliqués dans l'activité

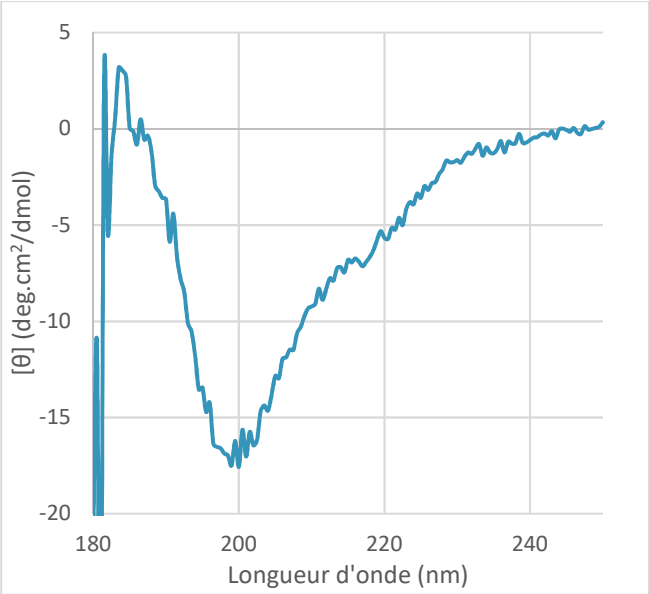


Figure 46. Spectre de dichroïsme circulaire du peptide EL4$_{346\text{-}394}$. Les résultats sont exprimés par l'ellipticité moyenne des résidus [Θ] (en deg.cm^2/dmol) en fonction de la longueur d'onde.

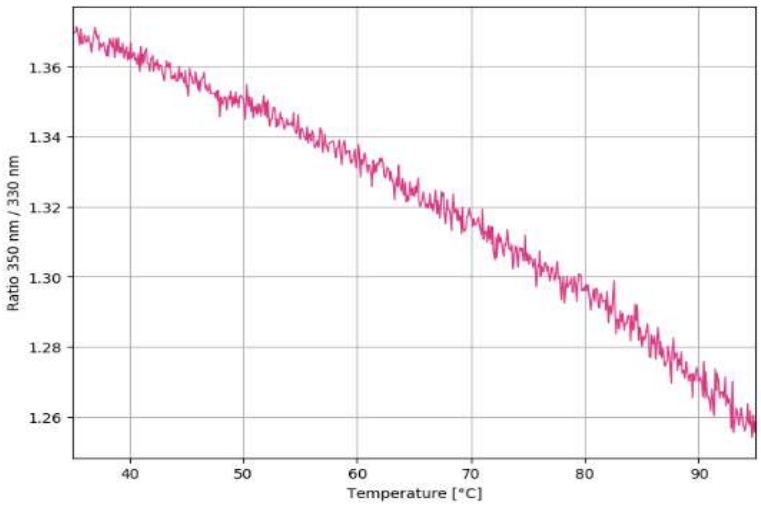


Figure 47. Suivi de la stabilité thermique du peptide EL4$_{346\text{-}394}$ par fluorescence. La fluorescence est suivie par le ratio 350nm/330nm dans un tampon phosphate 20mM, pH 7.0, 25 °C, à une concentration protéique de 20 µM. Le peptide EL4 contient 3 tryptophane permettant la mise en place de cette approche.

de PMTub par l'étude de mutagenèse précédente. Cette séquence a donc été produite avec un marquage GST et un site de clivage protéasique spécifique (**Figure 45A**), permettant la purification de l'extrémité C-terminale de EL4 seule, après clivage. La production dans *E. coli* a montré un très faible rendement avec une forte propension de la protéine de fusion à précipiter dans la fraction non soluble (**Figure 45B**). Le rendement de production de l'extrémité terminale de EL4 soluble ne s'est pas amélioré après plusieurs optimisations de protocole, notamment le changement de durée et de température de la culture induite (de 1 à 24 heures et 16°C à 37°C), la concentration d'inducteur (0 et 0,1 mM d'IPTG) et la souche d'*E.coli* utilisée (DH5α et BL21(DE3)). Une analyse par spectrométrie de masse a été effectuée après purification et a confirmé toutefois que la protéine récupérée avait la séquence attendue (**Figure 45C et D**).

Un test d'interaction *in vitro* basé sur une approche de type GST pull-down a été mise en place pour tenter de mettre en évidence la capacité de la protéine de fusion GST-EL4 à capturer de la FasC non mannosylée (le substrat PMTub utilisé dans notre test d'activité). Toutefois, aucune interaction entre GST-EL4 et la FasC n'a pu être détectée à l'aide de cette stratégie (données non montrées), renforçant la nécessité d'améliorer la production de la boucle soluble afin d'aller vers des techniques plus sensibles.

Pour obtenir les quantités du peptide EL4 nécessaires à des études plus fines telles que des expériences de titrage par calorimétrie isotherme (ITC) ou de résonnance magnétique nucléaire (RMN), un travail important d'optimisation de la production par *E.coli* aurait été nécessairesans garantie de succès, ce qui nous a donc conduits à nous orienter plutôt vers une production par synthèse chimique du peptide $PMTub_{346\text{-}394}$ de 53 acides aminés, correspondant à la séquence TNSAGNYHPWESKPWTWPMSLRPVLYAIDQQDVAGSGAQSSVKAEMLVGTPAM de EL4.

(2) Étude de la conformation de EL4 en solution.

En première intention, nous avons utilisé la mesure du dichroïsme circulaire (DC) pour obtenir une première estimation de la structure éventuelle adoptée par ce peptide $EL4_{346\text{-}394}$ en solution. Grâce à cette technique, il est possible de distinguer les hélices α, les feuillets β et les peptides non structurés par leurs courbes d'ellipticité, qui présentent des points d'inflexion caractéristiques en fonction des longueurs d'ondes. Ainsi il est attendu pour les

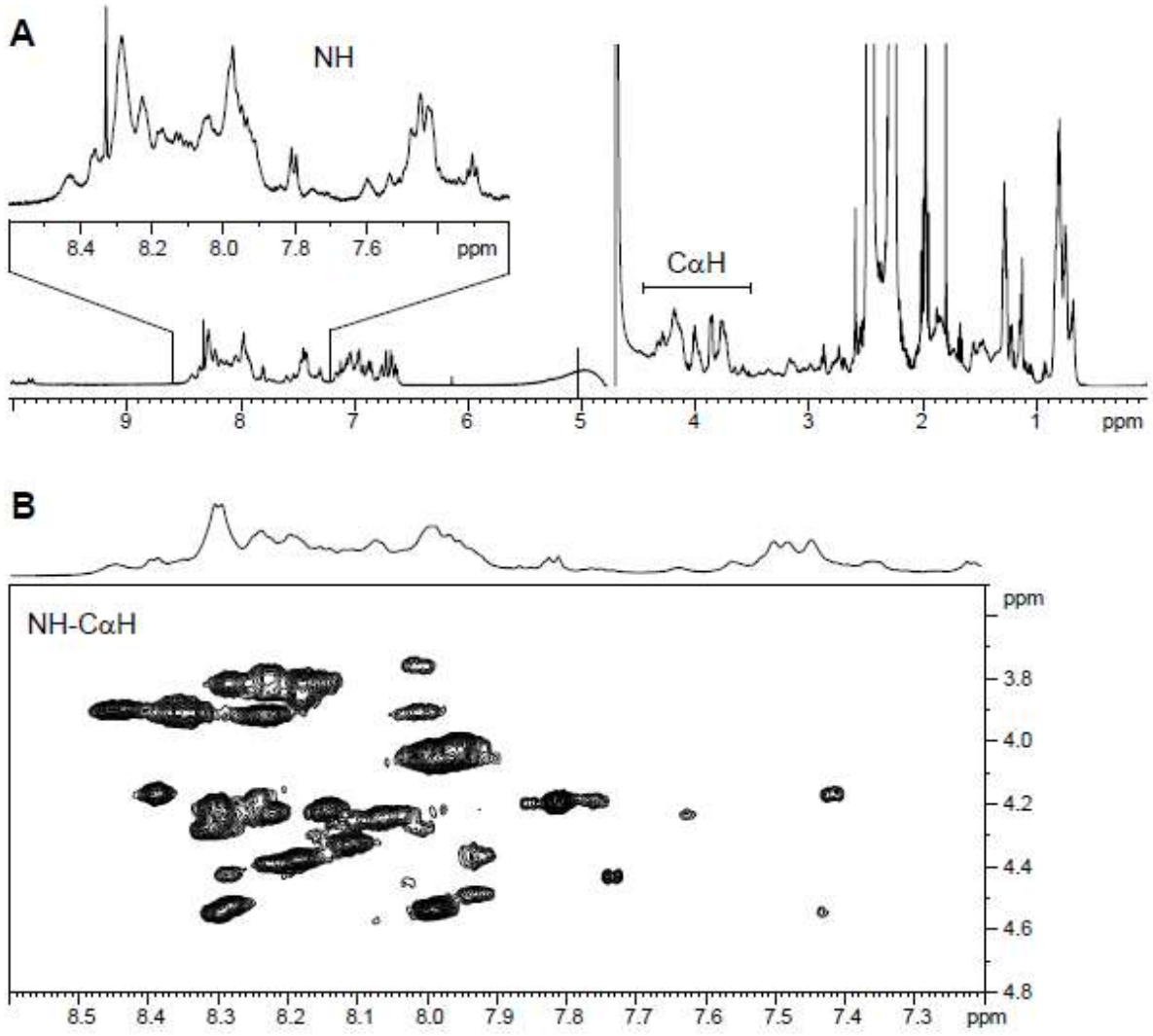


Figure 48. Analyse RMN ^{1}H du peptide EL4$_{346\text{-}394}$. A) Spectre RMN 1D ^{1}H 600MHz (3.2mg/ml, Tampon Phosphate 20mM, pH6.5, 295K) avec agrandissement de la zone des protons NH. **B)** Spectre partiel RMN 2D ^{1}H-^{1}H TOCSY 80ms, (3.2mg/ml, Tampon Phosphate 20mM, pH6.5, 300K) correspondant à la zone des corrélations NH-CαH.

hélices α et les feuillets β une forte ellipticité positive entre 190 et 200 nm avec respectivement 2 minima négatives à 208 et 222 nm pour les hélices α, et un minimum négatif à 218 nm pour les feuillets β. À l'inverse les peptides non structurés se caractérisent généralement par une faible ellipticité négative aux environs de 200 nm. Les résultats obtenus pour le peptide $EL4_{346\text{-}394}$ montrent un signal proche de celui d'une protéine dénaturée avec un minimum négatif à 200 nm, suggérant l'absence de structuration particulière détectable par DC pour ce peptide dans l'eau (**Figure 46**). Toutefois, il a été déjà été observé que la présence de structures secondaires des domaines WW n'est pas toujours détectable par la mesure dichroïsme circulaire DC[182,183]. De plus ce résultat n'est pas complètement surprenant si l'on considère que la conformation adopté par le peptide isolé pourrait être proche de celle prédite par le modèle de PMTub. En effet selon ce modèle, seul un nombre réduit d'acides aminés seraient impliqués dans des structures secondaires (environ une quinzaine sur les 53 que compte le peptide $EL4_{346\text{-}394}$) ; la majorité de la séquence serait désordonnée comme suggéré par le signal observé.

Nous avons tenté d'aller plus loin dans l'étude de la structuration éventuelle de ce peptide en solution en réalisant des analyses par balayage thermique de la fluorescence de tryptophane. Cette technique basée sur le suivi de la fluorescence des tryptophanes en fonction de la température permet de détecter des transitions de phase dans une courbe de dénaturation thermique, témoignant de la présence de conformations préférentielles de la protéine. La présence de 3 tryptophanes dans le peptide EL4 nous a donc incités à appliquer cette technique pour tenter de mettre en évidence une structuration éventuelle attendue pour un domaine WW.

La courbe de dénaturation thermique obtenue entre 35 et 100°C pour le peptide $EL4_{346\text{-}394}$ ne montre aucune transition de phase permettant de mettre en évidence une quelconque conformation adoptée par ce peptide en solution dans l'eau (**Figure 47**). Il est cependant important de noter que les contraintes techniques liées à l'appareillage utilisé pour ces expériences imposent une température initiale du balayage de 35°C : température relativement élevée à laquelle il n'est pas impossible que l'agitation thermique soit déjà suffisante pour avoir entrainé une dénaturation du peptide en solution.

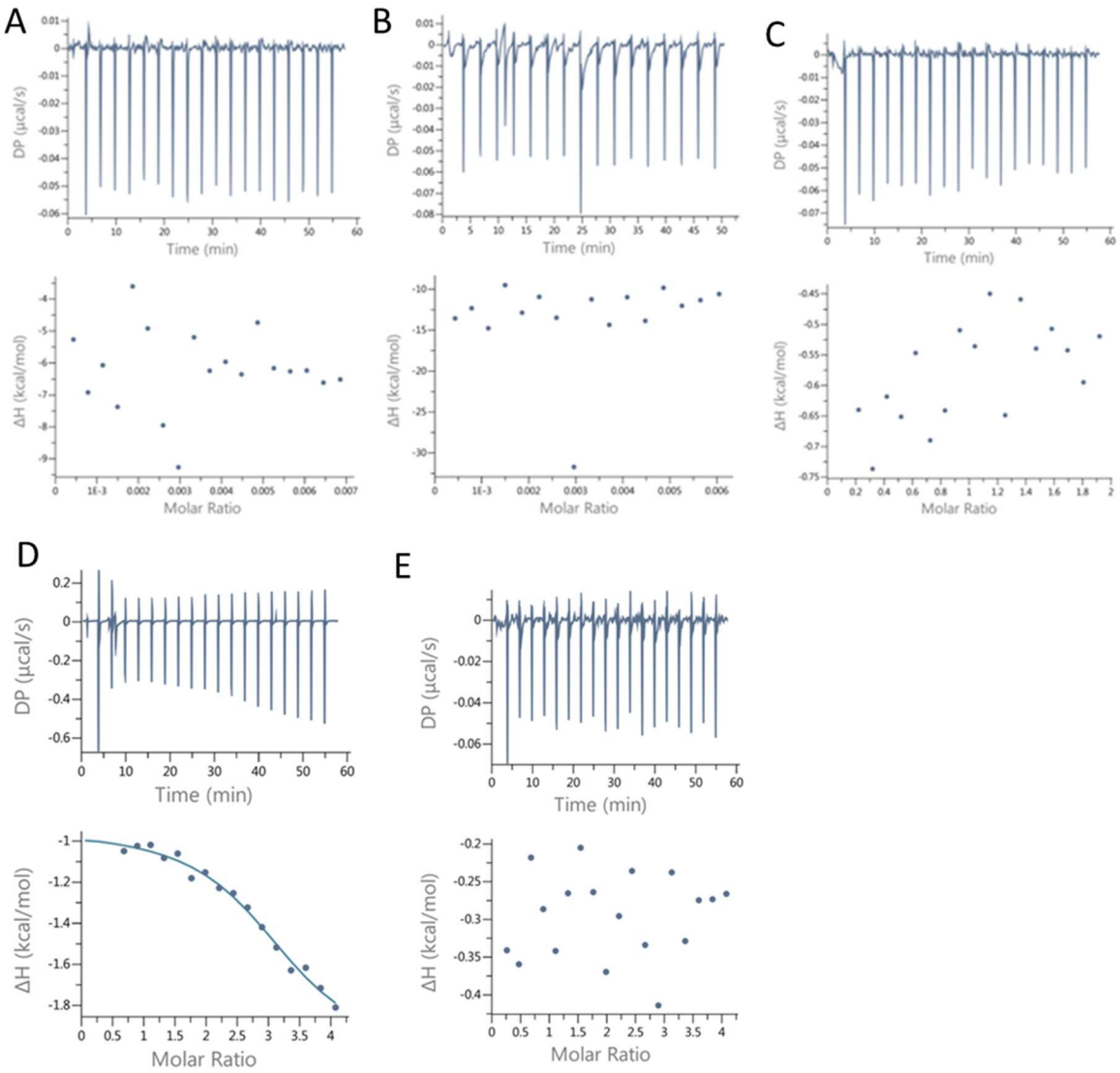


Figure 49. Titration calorimétrique de l'interaction entre le peptide EL4 de PMTub et un site peptide d'Apa de *M. tuberculosis,* porteur d'un site de mannosylation. Chaque pic correspond à une injection du peptide d'Apa dans une solution de peptide EL4 (panels du haut). L'intégration des signaux de titrage, engendrés par l'interaction EL4/Apa, a permis de générer des valeurs de ΔH, représentées en fonction du ratio molaire entre les 2 interactants. Les données de titration résultantes (panels du bas) ont été modélisées par une courbe de tendance selon un modèle d'interaction à un site lorsque cela était possible. **A)** Injection de Apa (250µM) dans une solution de peptide EL4 (25µM), à 25°C **B)** Injection de Apa (700µM) dans une solution de peptide EL4 (25µM), à 25°C **C)** Injection de Apa (700µM) dans une solution de peptide EL4 (69µM), à 25°C **D)** Injection de Apa (700µM) dans une solution de peptide EL4 (33µM), à 37°C **E)** Injection de Apa (700µM) dans une solution contrôle sans peptide EL4, à 37°C.

Enfin, nous avons voulu vérifier s'il était possible de mettre en évidence une quelconque structuration du peptide $EL4_{346-394}$ par des expériences en RMN du proton ^1H, sans avoir à recourir à l'étape laborieuse d'attribution complète des signaux. En effet, il est connu de manière empirique que l'allure générale des spectre 1D ^1H RMN et 2D ^1H-^1H de peptides peut fournir des indices intéressants quant à leur structuration éventuelle en solution. La présence de conformation préférentielle se traduit généralement par une signature particulière se caractérisant entre autres par une dispersion marquée des résonnances des protons amides (NH) et CαH du squelette peptidique[182]. Nous avons alors enregistré les spectres RMN 1D et 2DTOCSY ^1H-^1H du peptide $EL4_{346-394}$ en solution dans du tampon phosphate[184] 20mM D_2O/H_2O. L'analyse du spectre 1D montre une bonne dispersion générale des signaux notamment dans la zone de résonnance des protons amides (NH) qui s'étend sur plus de 1 ppm suggérant un repliement possible du peptide en solution (**Figure 48**). De même, l'examen de la zone du spectre 2D ^1H-^1H TOCSY dans laquelle sont censés se concentrer les signaux de corrélation scalaire NH-CαH montre un grand nombre de taches confirmant également une bonne dispersion des signaux des protons CαH. Près d'une cinquantaine de corrélations distinctes sont discernables qui, compte tenu de la présence de 5 résidus proline, pourraient correspondre à la totalité des corrélations NH-CαH attendues pour le peptide $EL4_{346-}$394. L'ensemble de ces éléments, caractérisant généralement les spectres RMN de protéines structurées, révèlent une probable structuration du peptide $EL4_{346-}$394 en solution que seule une attribution complète des signaux (qui pourra être entreprise hors du cadre de cette thèse) serait à même de confirmer définitivement.

(3) Interaction EL4 – peptide substrat accepteur.

Sur la base de ces observations nous avons voulu vérifier notre hypothèse selon laquelle EL4 pourrait contribuer à la reconnaissance du peptide substrat accepteur. Nous avons donc cherché à démontrer une interaction entre le peptide $EL4_{346-394}$ et un peptide synthétique K-Apa_{1-15}(K P E P A P P V T T A A S P P A synthétisé au laboratoire) dérivé du peptide N-terminal 1-14 de la mannoprotéine Apa de *M. tuberculosis* et comportant un site de mannosylation. Le choix du peptide substrat accepteur était justifié par plusieurs études antérieures dans lesquelles il a servi à prouver l'activité de PMTub ou à caractériser le K_M de la PMT de *M. smegmatis* [75]. Enfin pour détecter une interaction éventuelle entre les peptides $EL4_{346-394}$ et K-Apa_{1-15}, nous avons choisi d'utiliser la méthode de titration calorimétrique isotherme ou ITC (pour Isothermal titration calorimetry). En effet cette méthode a pour

avantage d'être relativement sensible et surtout de ne pas nécessiter de marquage préalable spécifique d'au moins l'un des partenaires comme l'exigent la plupart des autres approches. Dans un premier temps les expériences ont été réalisée à 25C° en utilisant différentes concentrations en peptide $EL4_{346-394}$ (25µM ou 69 µM) et (250µM ou 700µM). Dans ces conditions, aucun signal caractéristique n'a pu être détecté permettant de suspecter une interaction éventuelle. (**Figure 49A, B et C**).

En revanche les mêmes expériences réalisées cette fois ci à 37°c (**Figure 49C et D**) montrent un signal caractéristique d'une réaction endothermique consécutive aux injections (2µl) successives de ligand K-Apa_{1-15} (700µM) dans la solution (200µl) de peptide $EL4_{346-394}$ (33µM). L'absence d'effet dose lors de l'injection du ligand K-Apa_{1-15} dans le tampon seul, confirme la spécificité de la réaction observée en présence du peptide $EL4_{346-394}$. Ces observations démontrent clairement une interaction en solution entre $EL4_{346-394}$ et K-Apa_{1-15} avec une constante de dissociation calculée de l'ordre de la dizaine de µM.

Il semblerait donc que l'ITC permette de suivre une interaction entre le peptide et un peptide substrat interactant. Cette expérience montre aussi la dépendance à la température de l'interaction, celle-ci semble alors sous le contrôle de l'entropie et la réaction devient spontanée à température plus élevée.

Les résultats d'ITC sont très encourageants, mais restent préliminaires. Les mesures d'interaction devront donc être confirmées et optimisées pour conclure définitivement à une interaction entre les deux peptides. Si cette interaction est confirmée, cet outil pourra également être utilisé pour caractériser plus finement les déterminants moléculaires à la base de la reconnaissance. En effet, il sera possible de cartographier précisément l'interaction entre EL4 et le peptide d'Apa en utilisant différents mutants des deux peptides et d'analyser l'effet des mutations sur l'interaction.

Finalement ces résultats préliminaires viennent conforter notre hypothèse concernant le rôle de cette boucle EL4 de PMTub dans la reconnaissance du substrat accepteur de la *O*-mannosylation par analogie les domaines WW d'interaction protéine-protéine. L'utilisation de différents peptides substrats accepteurs pourrait permettre de décortiquer les mécanismes de reconnaissance et de caractériser les propriétés intrinsèques des sites de mannosylation pour mieux comprendre le déterminisme de la sélectivité de PMTub pour ses mannoprotéines cibles.

Partie 3 : Stratégies d'identification d'inhibiteurs de PMTub

Du fait de l'implication de PMTub dans la virulence de *M. tuberculosis*, cette enzyme constitue une cible potentielle pour un futur médicament antivirulence[52]. L'inhibition de PMTub par un composé offrirait ainsi de nouvelles voies thérapeutiques contre la tuberculose. Au-delà de la tuberculose, les PMT sont impliquées dans la virulence de divers pathogènes[84,85,185], l'inhibition spécifique de PMTub fournirait donc une preuve de concept pour le ciblage des PMT dans diverses pathologies.

A. Criblage d'inhibiteurs en employant le test phénotypique d'activité

1. Preuve de faisabilité

Le ciblage des PMT de ces organismes pathogènes constituerait une voie intéressante pour le développement de diverses thérapies antivirulences. À ce jour, un seul inhibiteur de PMT a pu être identifié, OGT2468, un dérivé de rhodanine-3-acide acétique[118]. Cet inhibiteur semble avoir un effet spécifique sur une isoforme chez *Candida albicans*, ce qui apparaît encourageant dans la recherche d'inhibiteurs spécifiques d'une seule PMT. Cependant, l'utilisation en clinique de cette molécule n'a pas été testée, et notablement, son « ciblage spécifique d'une seule isoforme » semble se réduire à *Candida albicans,* puisque cette molécule affecte toutes les PMT de *Saccharomyces cerevisiae*[186].

Le test phénotypique d'activité de PMTub développé au début de ma thèse peut être adapté au crible d'inhibiteurs. En effet, chaque étape a été élaborée dans l'optique d'une potentielle automatisation du crible de composés inhibiteurs. Le test a été conçu sur bactéries vivantes pour permettre de prendre en compte les limitations dues à la faible perméabilité de la paroi mycobactérienne vis-à-vis de potentiels inhibiteurs de PMTub. Les tests ELISA direct et lectinique sont réalisés en plaque 96 puits, permettant une automatisation a moyen débit de la mesure de la mannosylation de la fascicline. Seule l'étape de culture n'est pas adaptée à un débit élevé de criblage. En effet, dans sa forme initiale, le bio-essai que nous avons développé requiert des cultures de volumes trop importants pour permettre d'évaluer en parallèle plusieurs dizaines voire centaines de composés pour le criblage d'inhibiteur.

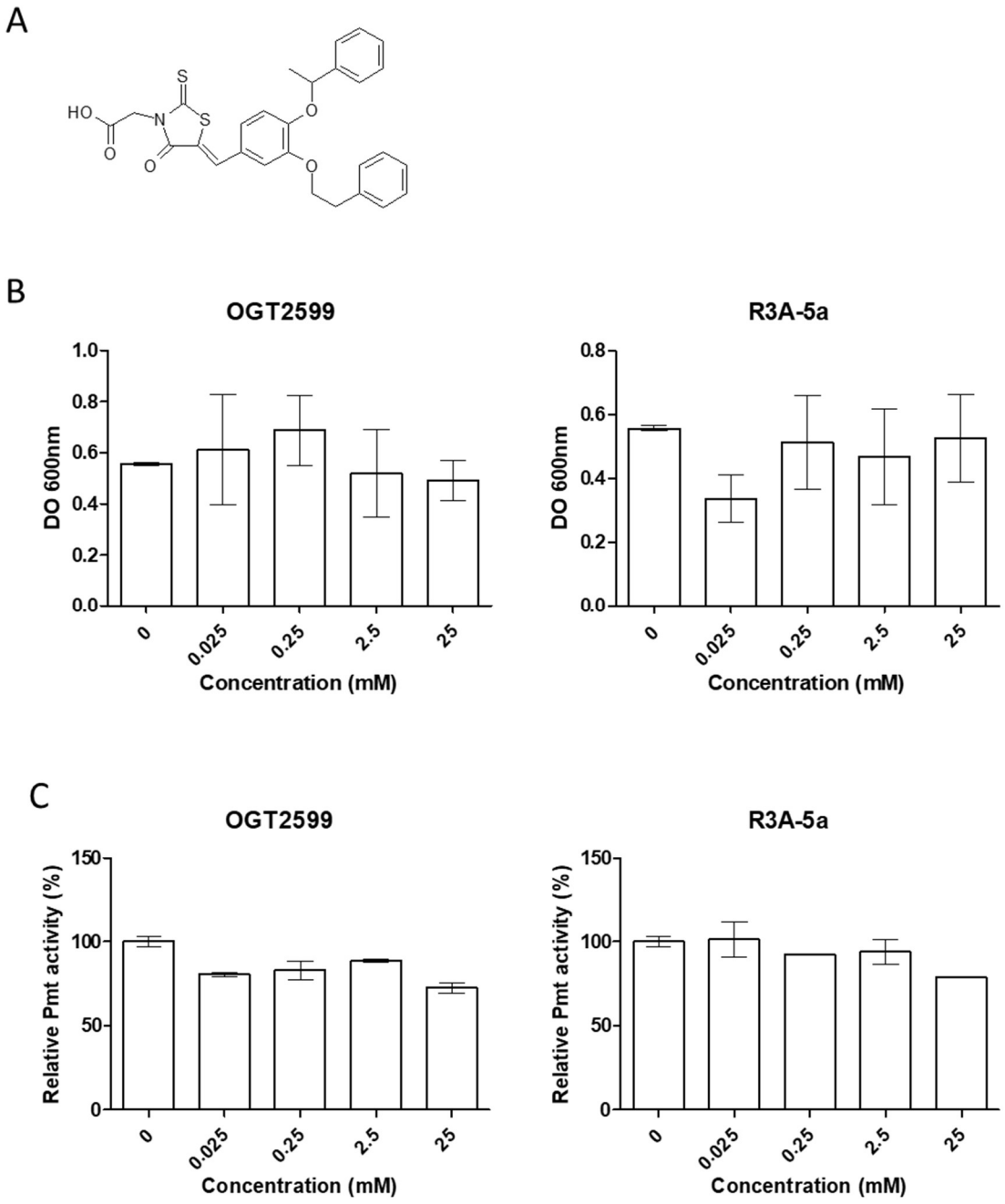


Figure 50. Mesure des effets inhibiteurs de dérivés de rhodanine-3-acide acétique sur PMTub. A) Structure de la molécule testée appelée OGT2599 ou R3A-5 en fonction des sources**. B)** Mesure de la DO_{600nm} de culture de *M. smegmatis* exprimant PMTub native et la $FasC^{6His}$ après 7 jours de culture en présence de différentes concentrations des composés d'intérêt (OGT2599 et R3A-5a). **C)** Mesure du taux de mannosylation de la fascicline produite par *M. smegmatis* exprimant PMTub native et la $FasC^{6His}$ après 7 jours de culture en présence de différentes concentrations des composés d'intérêt. Des triplicats techniques de duplicats biologiques sont présentés.

Pour lever cette limitation et augmenter le débit du test d'activité à un niveau compatible avec l'objectif de criblage d'inhibiteurs potentiels de PMTub, nous avons miniaturisé l'étape initiale de culture de la souche rapportrice sur plaque 96 puits.

Le suivi de la production de fascicline dans le temps par la souche de *M. smegmatis*, modifiée pour exprimer PMTub et $FasC^{6His}$, cultivée en plaque 96 puits, a été réalisé. Cette cinétique a permis de déterminer la durée de culture nécessaire pour obtenir une production de fascicline suffisante pour la détermination du taux de mannosylation de la protéine par ELISA lectinique (**Figure 20A**). Il a pu ainsi être observé que la fascicline est détectable dès le 3e jour de culture pour atteindre son niveau maximum au 7e jour, dernier jour de suivi dans notre cinétique. Afin de limiter autant que possible les variations éventuelles de la vitesse de croissance de la souche rapportrice et de production de FasC d'une culture à l'autre, nous avons choisi de prolonger les temps de culture en plaque 96 puits jusqu'à 7 jours afin d'atteindre la phase de stagnation.

Des essais d'inhibition de PMTub ont donc été réalisés en utilisant le test d'activité phénotypique après optimisation de l'étape de culture. Ces essais d'inhibition ont été réalisés avec un dérivé de rhodanine-3-acide acétique, utilisé précédemment pour l'inhibition de PMT de levure : OGT2599[187] et R3A-5a[118]. Ces deux appellations correspondent à la même molécule provenant de deux sources différentes (**Figure 50A**).

La mannosylation de la fascicline a été suivie sur des cultures de la souche de *M. smegmatis* modifiée pour exprimer PMTub et $FasC^{6His}$ en absence ou en présence de différentes concentrations des composés dérivés de rhodanine-3-acide acétique.

Le suivi de la croissance bactérienne par densité optique à 600 nm (DO_{600}) au septième jour de culture montre que la culture en plaque 96 puits en présence des composés testés à des concentrations de 0 à 25 mM ne permet pas de détecter d'effet dose et suggère que ces molécules n'affecte pas la croissance de la souche mycobactérienne (**Figure 50B**). De plus, la mesure de la concentration en fascicline permet de noter une production suffisante de la protéine dans les surnageants de culture, pour l'évaluation de la mannosylation par ELISA lectinique.

Pour permettre la mesure de l'activité relative de PMTub en présence des dérivés de rhodanine-3-acide acétique par ELISA lectinique, les cultures des souches « No PMT » et

exprimant PMTub native ont aussi été réalisées en absence d'inhibiteur. L'évaluation de la mannosylation de la fascicline produite par ces deux souches permet de définir des activités PMTub relative de 100% et 0%. La mesure du taux de mannosylation de la fascicline après traitement par les composés d'intérêt ne montre pas de différence avec la souche non traitée et démontre que ces molécules n'ont pas d'effet significatif sur le transfert de mannose sur la fascicline (**Figure 50C**).

Les dérivés de rhodanine-3-acide acétique inhibant les PMT de levure ne semblent pas avoir d'effet sur PMTub. Il a précédemment été déterminer que la sensibilité à cet inhibiteur était liée à la présence d'une phénylalanine dans la cinquième boucle externe des PMT[188]. La présence d'un tryptophane à cette position dans PMTub pourrait expliquer l'absence d'activité de ces composés chez la mycobactérie. Cet essai reste néanmoins une preuve de concept pour l'utilisation de ce test d'activité phénotypique pour le criblage de composés inhibiteur de PMTub.

2. Sélection rationnelle d'inhibiteurs putatifs

L'étude des relations structure/fonction décrite dans les parties précédentes de ce manuscrit a permis de mieux comprendre les mécanismes moléculaires enzymatiques de PMTub. Ces connaissances sont importantes pour la sélection rationnelle de molécules avec un pouvoir inhibiteur putatif de PMTub. Nous avons par exemple pu identifier le « pseudo-domaine WW » de la boucle EL4 comme portant des caractéristiques propres à PMTub, et comme impliqué dans l'activité de l'enzyme. Ce domaine de la boucle EL4 semble donc être une cible prometteuse d'un inhibiteur sélectif de PMTub. Nous avons émis l'hypothèse que le « pseudo-domaine WW » participe à l'interaction avec les substrats protéiques riches en proline de PMTub. Pour confirmer ce rôle hypothétique, des approches *in vitro* ont été mises en place qui donnent des résultats préliminaires intéressants (voir p.153). Ceci nous a enjoint à tester la capacité de molécules dérivées de proline, ou portant des noyaux pyrroles, comme inhibiteurs de l'activité de PMTub. En parallèle, des molécules dérivées de rhodanine-3-acide acétique, autres que celle ayant un effet sur les PMT de levure, ont aussi été testées pour leur potentielle capacité à inhiber la réaction catalysée par PMTub.

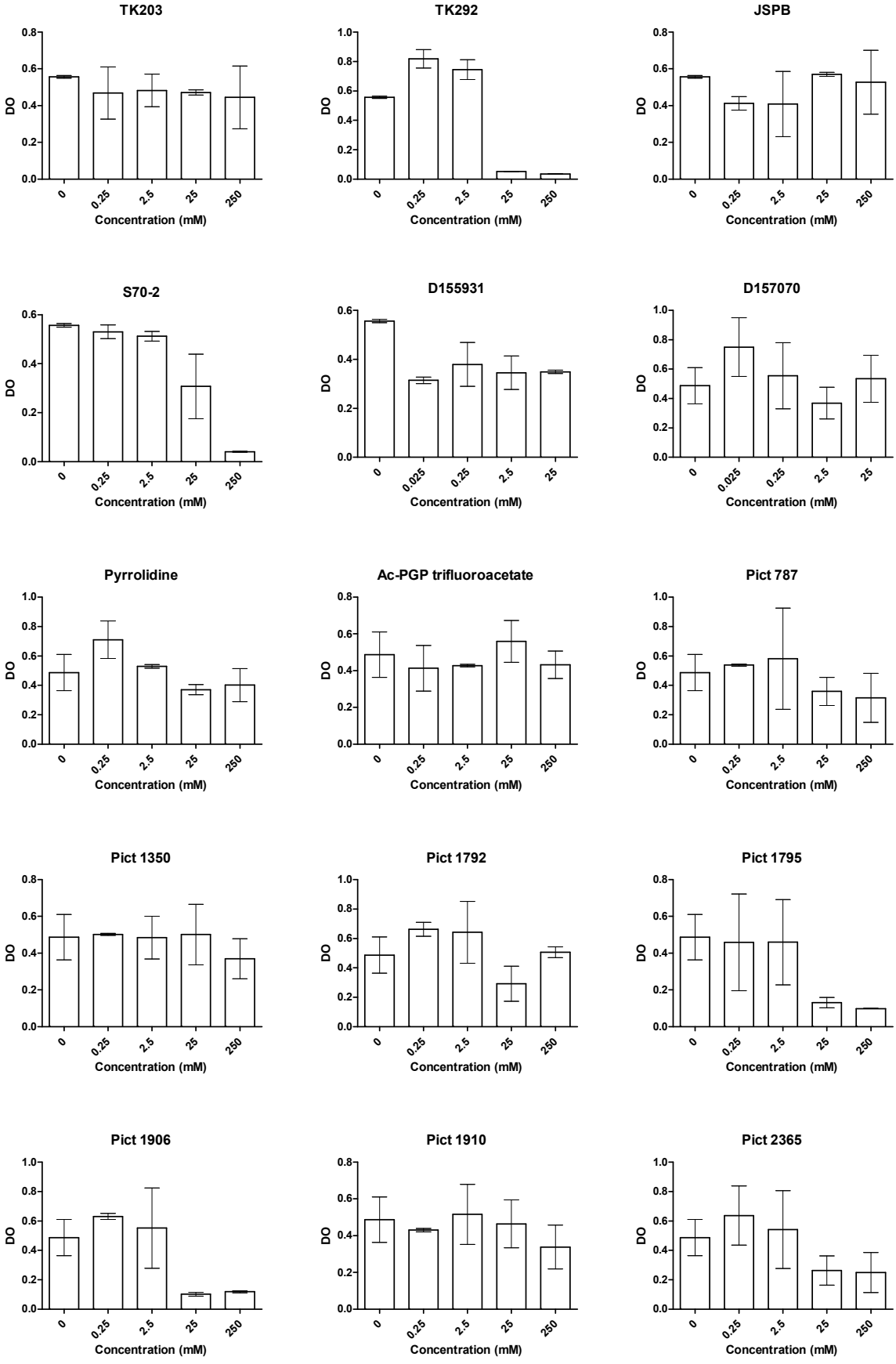
Le suivi de la croissance bactérienne par lecture de la DO_{600nm} a permis de constater un effet dose de certains composés (**Figure 51**). De manière attendue, les composés ayant un effet sur la croissance bactérienne ont aussi un effet sur la production de FasC par les bactéries (**Figure 52**). Néanmoins, la concentration en FasC dans les surnageants est demeurée suffisante pour chaque condition et nous a permis d'appliquer le protocole d'ELISA lectinique pour déterminer son taux de mannosylation. La mesure du taux de mannosylation de la FasC en présence des différents composés testés n'a malheureusement pas permis d'identifier de molécules ayant un effet sur l'activité de PMTub (Figure 53).

3. Perspectives : optimisation pour passage HTS

Plusieurs problèmes techniques ont pu être mis en évidence lors de ces premiers essais de crible d'inhibiteurs à moyen débit, qui devront être résolus pour un passage à plus un haut débit.

Tout d'abord, le système de production de la FasC n'est pas stable au cours du temps. En effet, on observe une perte de la production de la protéine après des cultures trop longues. Cette instabilité de la souche nécessite sa reconstruction systématique pour chaque campagne de crible. Ceci peut avoir pour origine l'utilisation d'un plasmide réplicatif pour la production de la FasC, qui pourrait être perdu ou qui pourrait subir une recombinaison partielle au cours du temps, entrainant finalement la perte du gène de la protéine rapportrice FasC étiquetée His. Une solution possible de ce problème pourrait être d'utiliser un plasmide intégratif pour l'expression de la FasC, qui aurait pour avantage de stabiliser la production de FasC dans le temps. Nous avons donc construit un tel plasmide intégratif pour la production de $FasC^{6His}$ chez *M. smegmatis*. Les premiers essais ont permis de montrer que ce nouveau système permettait en effet la production de $FasC^{6His}$ (données non montrées). La stabilité de cette souche devra être mesurée dans le temps, ainsi que la compatibilité de ce plasmide intégratif avec celui utilisé pour l'expression du gène de PMTub.

Un autre problème identifié, a été « un effet de bord » important du à l'oxygénation non homogène de tous les puits (plus importante sur les bords qu'au centre de la plaque). Notablement, nous avions pu observer précédemment (voir p.79) que l'oxygénation d'une culture pouvait avoir un effet (favorable) sur le taux de mannosylation de la FasC. De plus, lors


TK203
TK292
JSPB
S70-2
D155931
D157070
Pyrrolidine
Ac-PGP trifluoroacetate
Pict 787
Pict 1350
Pict 1792
Pict 1795
Pict 1906
Pict 1910
Pict 2365
DO
Concentration (mM)
0
0.025
0.25
2.5
25
250

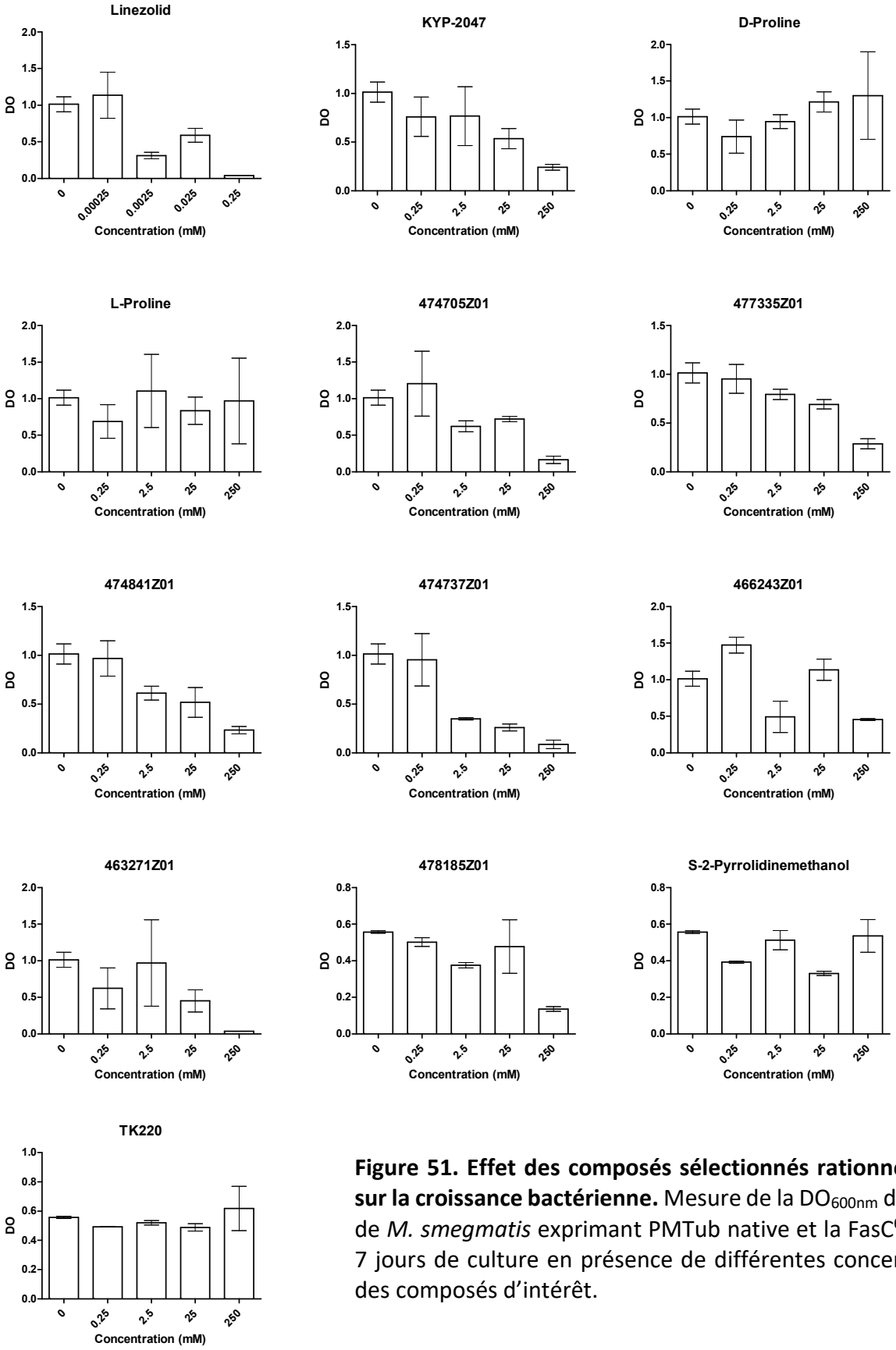
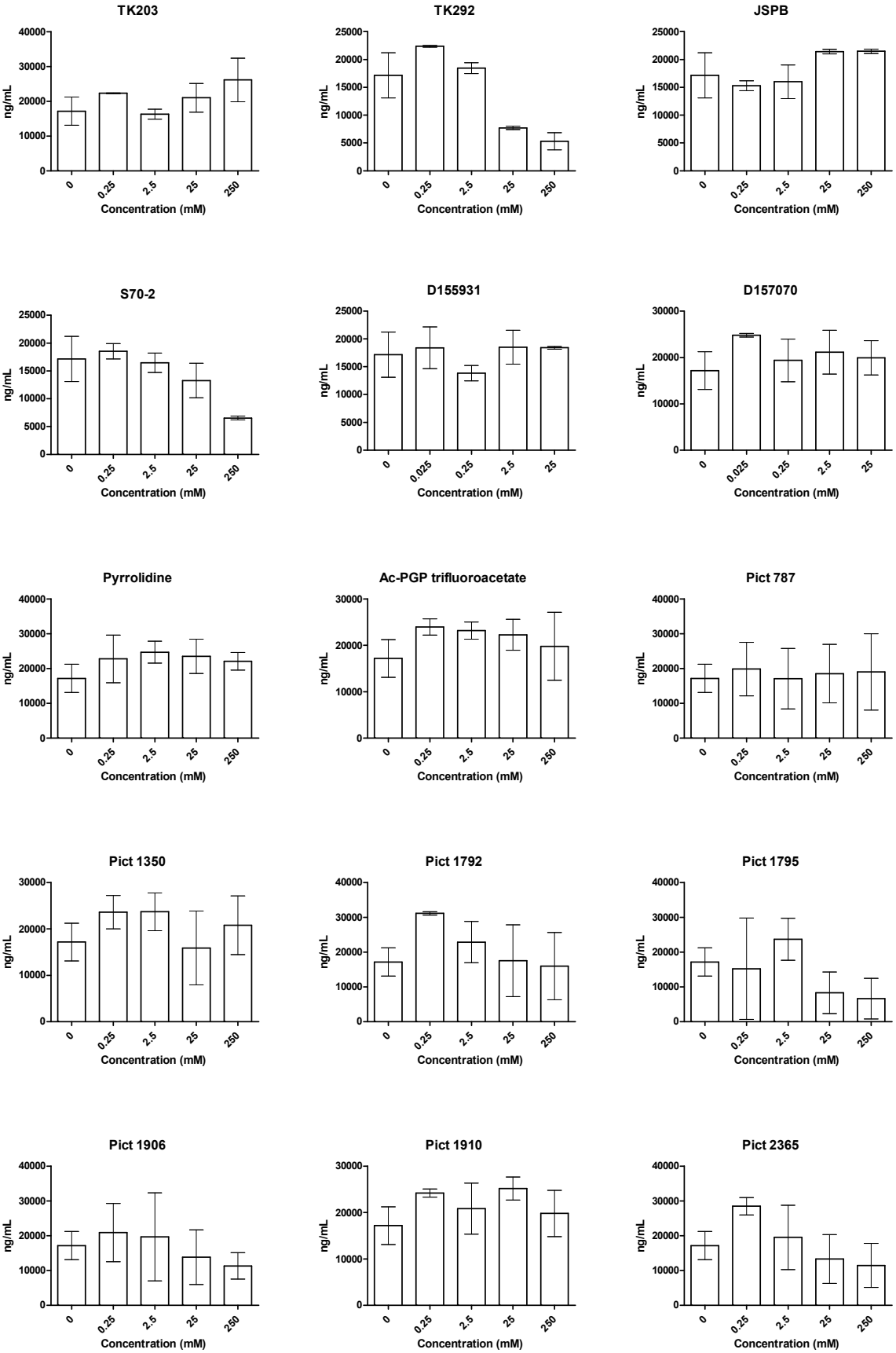


Figure 51. Effet des composés sélectionnés rationnellement sur la croissance bactérienne. Mesure de la DO_{600nm} de culture de *M. smegmatis* exprimant PMTub native et la $FasC^{6His}$ après 7 jours de culture en présence de différentes concentrations des composés d'intérêt.


TK203
TK292
JSPB
S70-2
D155931
D157070
Pyrrolidine
Ac-PGP trifluoroacetate
Pict 787
Pict 1350
Pict 1792
Pict 1795
Pict 1906
Pict 1910
Pict 2365
ng/mL
Concentration (mM)
0
0.025
0.25
2.5
25
250

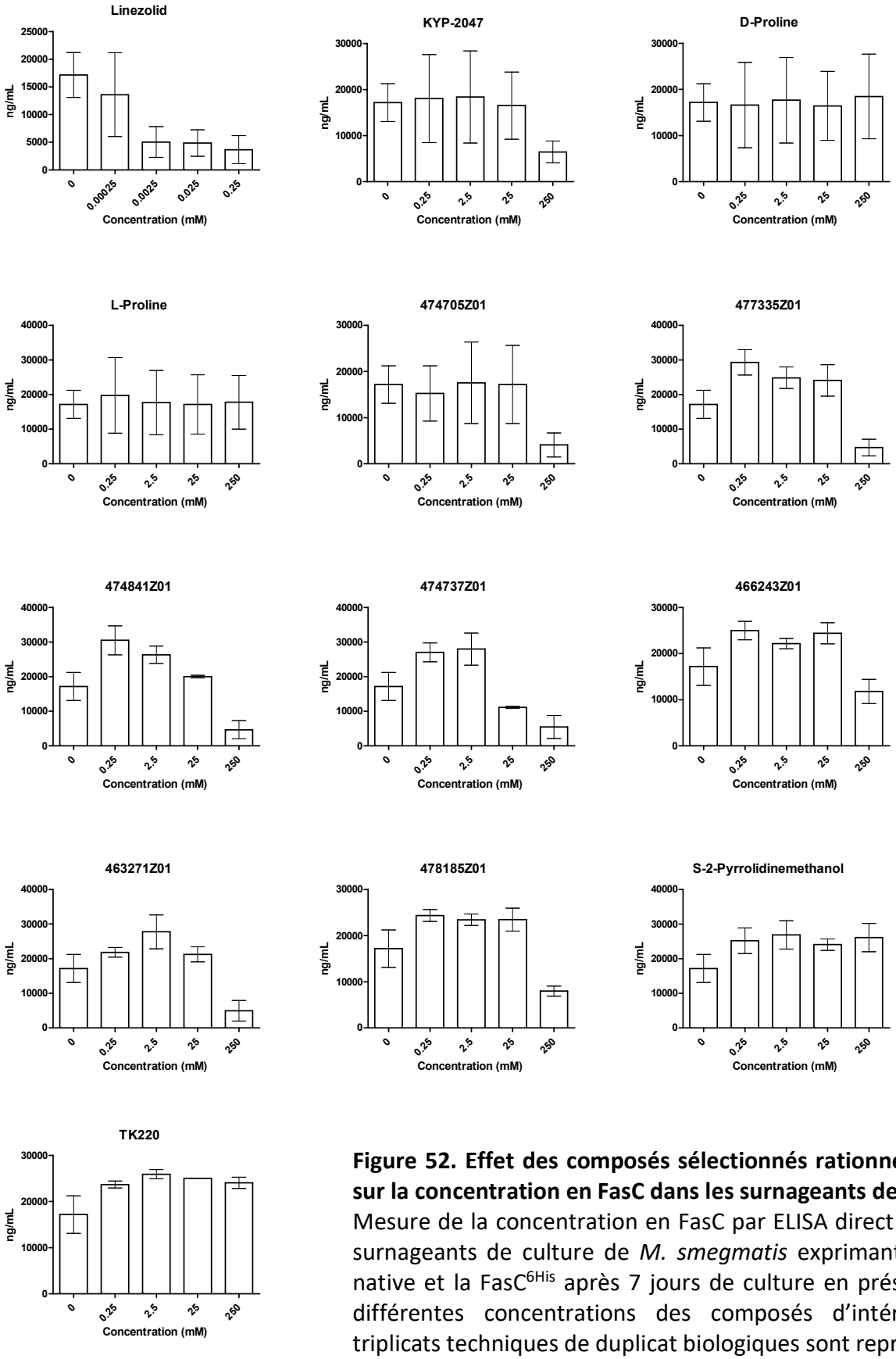
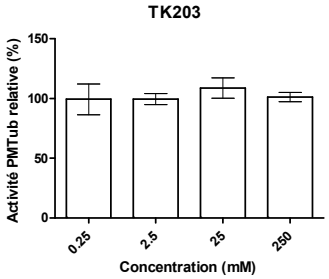
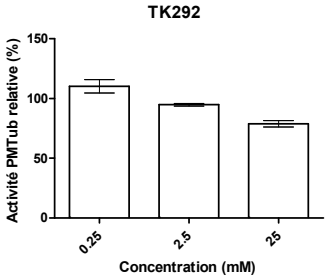
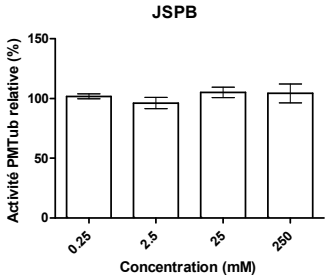
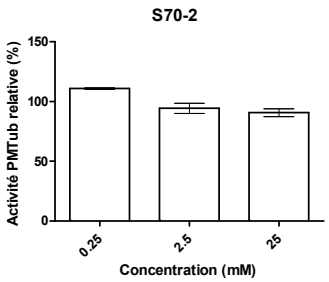


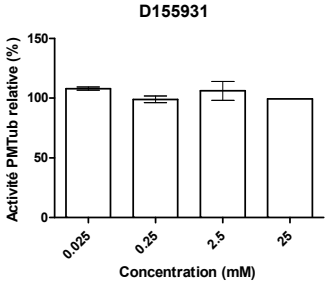
Figure 52. Effet des composés sélectionnés rationnellement sur la concentration en FasC dans les surnageants de culture. Mesure de la concentration en FasC par ELISA direct dans les surnageants de culture de *M. smegmatis* exprimant PMTub native et la FasC^{6His} après 7 jours de culture en présence de différentes concentrations des composés d'intérêt. Des triplicats techniques de duplicat biologiques sont représenté.

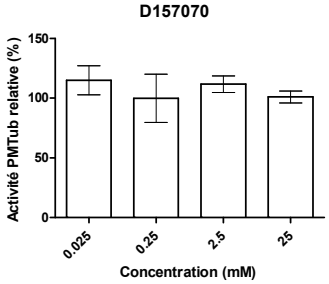

TK203
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

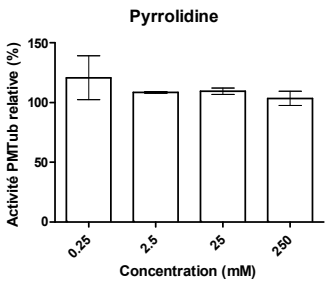

TK292
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
Concentration (mM)

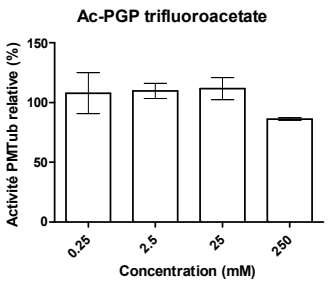

JSPB
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

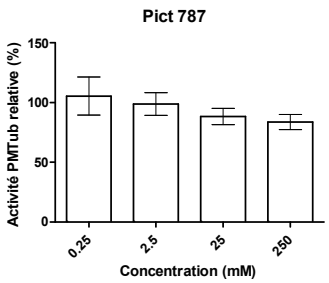

S70-2
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
Concentration (mM)

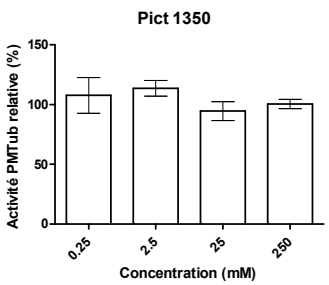

D155931
Activité PMTub relative (%)
150
100
50
0
0.025
0.25
2.5
25
Concentration (mM)

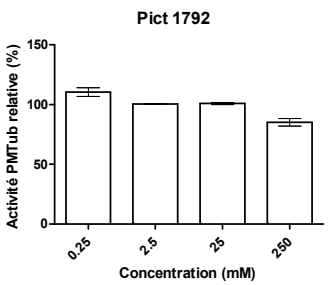

D157070
Activité PMTub relative (%)
150
100
50
0
0.025
0.25
2.5
25
Concentration (mM)

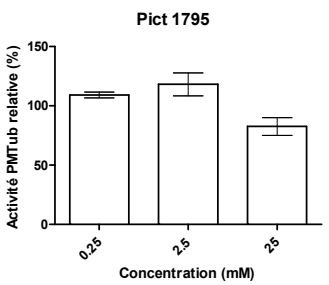

Pyrrolidine
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

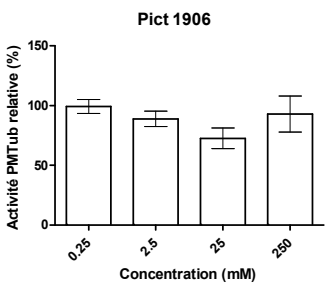

Ac-PGP trifluoroacetate
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

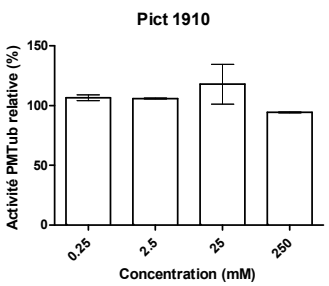

Pict 787
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

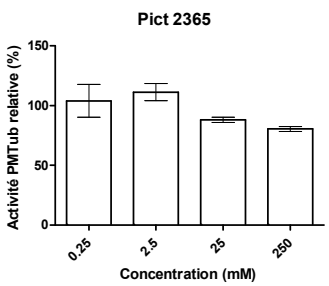

Pict 1350
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)


Pict 1792
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)


Pict 1795
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
Concentration (mM)


Pict 1906
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)


Pict 1910
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)


Pict 2365
Activité PMTub relative (%)
150
100
50
0
0.25
2.5
25
250
Concentration (mM)

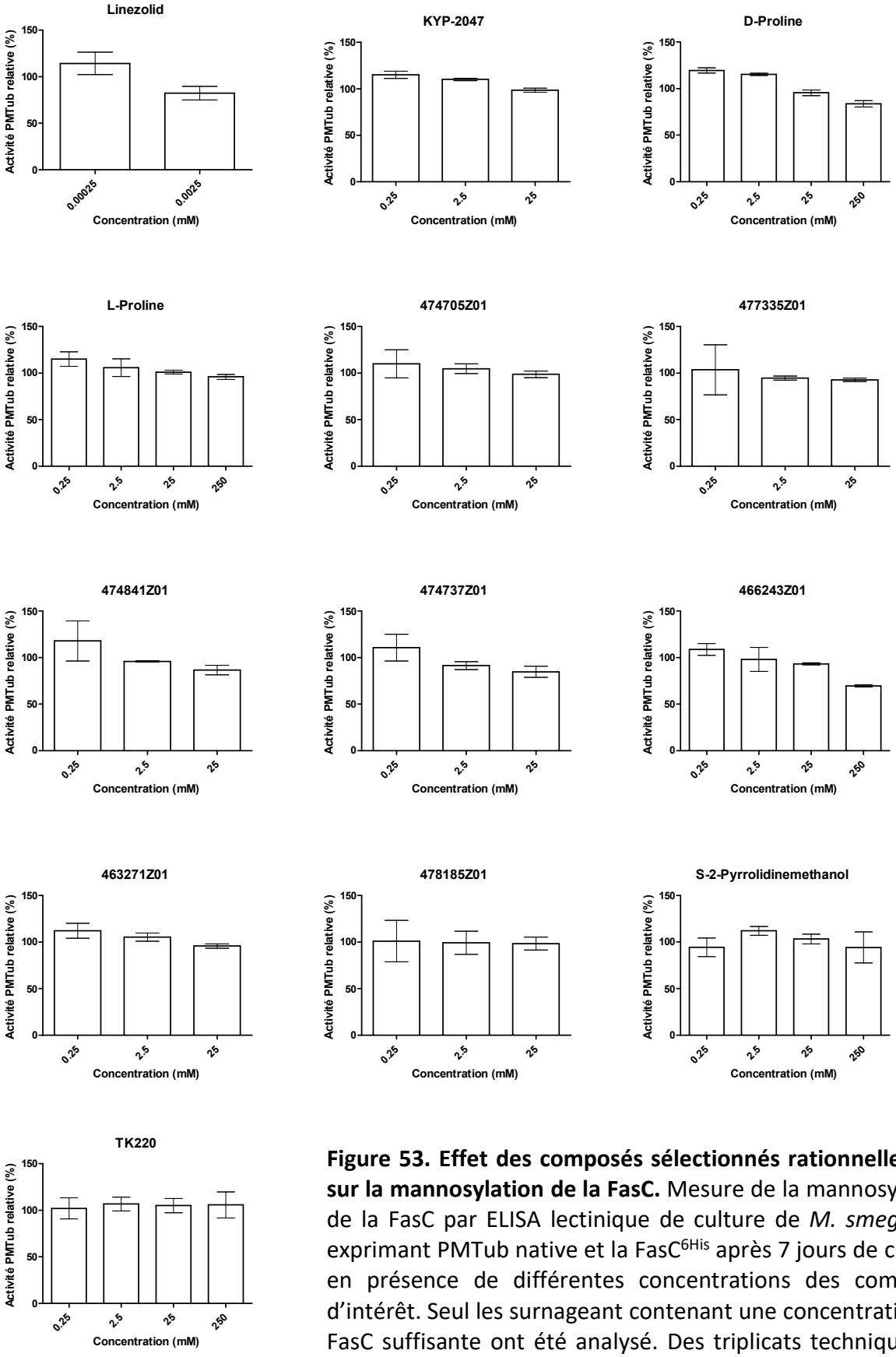


Figure 53. Effet des composés sélectionnés rationnellement sur la mannosylation de la FasC. Mesure de la mannosylation de la FasC par ELISA lectinique de culture de *M. smegmatis* exprimant PMTub native et la FasC^{6His} après 7 jours de culture en présence de différentes concentrations des composés d'intérêt. Seul les surnageant contenant une concentration en FasC suffisante ont été analysé. Des triplicats techniques de duplicat biologiques sont représenté.

de nos expériences de criblage, nous avons pu constater une forte évaporation du milieu de culture dans les puits situés en bords de plaque comparée aux puits en milieu de plaque. Une optimisation de la culture permettant de maintenir une oxygénation homogène de tous les puits de la plaque est donc nécessaire pour obtenir des résultats plus normalisés.

Enfin, l'approche de crible basée sur le test phénotypique mis en place lors de ce travail possède en outre plusieurs limitations majeures à l'identification de molécules ciblant spécifiquement PMTub. En effet, ce test est indirect, de plus il ne permet pas de caractériser des inhibiteurs d'un domaine spécifique de PMTub, ni de confirmer PMTub comme cible si une perte de mannosylation de la FasC est observée. Sa robustesse et son adaptabilité au crible d'inhibiteurs le rendent intéressant, mais cette approche devra être combinée avec plusieurs contrôles permettant de confirmer PMTub comme cible de hits putatifs et de vérifier également que ceux-ci n'affectent pas les PMT humaines.

B. Développement d'une approche de biologie synthétique

Pour contourner les limitations de l'approche précédente, nous avons décidé de mettre en place une approche innovante basée sur l'utilisation d'outil de biologie synthétique. L'objectif de cette approche est de sélectionner sélectivement les seuls composés ayant un effet sur un domaine fonctionnel spécifique de PMTub. De tels composés spécifiques ne devraient donc pas avoir d'effets sur les PMT humaines.

Lors nos travaux décrits plus haut, nous avons pu mettre évidence l'implication fonctionnelle de plusieurs domaines spécifiques de PMTub dans l'activité de l'enzyme. Basés sur le modèle 3D de l'enzyme, des rôles hypothétiques ont pu être assignés à ces domaines notamment au « pseudo-domaine WW » dont nous avons pu montrer qu'il pouvait être impliqué dans l'interaction avec le substrat protéique.

L'originalité de cette nouvelle approche innovante de crible est qu'elle est conçue pour suivre spécifiquement l'interaction entre EL4 et le substrat protéique dans le but de rechercher des composés capables de l'interrompre. Pour cela, nous proposons de « figer », *in vivo,* l'interaction entre la boucle EL4 de PMTub et son substrat, en utilisant un acide aminé non naturel (AAnn) cross linkant et photo-activable. L'insertion stratégique de cet acide aminé dans EL4 doit permettre de capturer les substrats interagissant avec cette boucle spécifique des PMT mycobactériennes. L'utilisation d'un substrat détectable, par fluorescence par exemple, devrait ainsi permettre de suivre cette interaction dans le contexte de l'enzyme entière et de la bactérie vivante. Cette stratégie présente l'avantage majeur de permettre de ne sélectionner uniquement que des composés actifs qui peuvent franchir la paroi mycobactérienne et atteindre PMTub. Il est aussi envisageable d'adapter cette stratégie pour le criblage de composés inhibiteurs de PMTub actifs sur des bactéries intracellulaires internalisées par des macrophages dans un modèle d'infection.

Toutefois le développement de cette approche originale requière la mise au point de différents outils biologiques pour suivre cette interaction. Pour cela, nous proposons donc d'incorporer un AAnn dans la séquence de PMTub par ingénierie génétique en utilisant les outils de la biologie synthétique. L'AAnn choisi est le *p*-Benzoyl-L-phénylalanine (pBpa), une benzophénone dont la photo-activation ($\lambda_{350\text{-}360nm}$) assure un cross-link efficace[189] et qui est indiqué dans l'étude des interactions protéine/protéine[190]. Pour permettre la reconnaissance et l'insertion métabolique d'un AAnn dans un organisme, il est nécessaire de lui fournir la

machinerie biologique adaptée. Cela passe par l'utilisation d'un système orthogonal ARNt/ARNt synthétase spécifique du pBpa[191]. Ce système orthogonal doit complémenter la machinerie traductionnelle endogène de l'organisme pour lui permettant d'insérer l'AAnn dans la séquence d'une protéine en regard d'un codon cible introduit ponctuellement dans le gène de la protéine d'intérêt. Le codon cible utilisé pour définir le site d'insertion de l'AAnn est le codon stop « AMBRE » (TAG), pour lequel ont été développés les outils de biologie synthétique spécifiques actuellement disponibles.

L'utilisation d'AAnn a été originellement développée pour *E. coli*[192]. Le système orthogonal utilisé est alors est le couple MjtRNA/MjTyrRS, dérivé de l'archébactérie *Methanococcus jannaschii*[192]. Ce couple a été choisi en raison de sa capacité à être utilisé efficacement par *E. coli* lors de la traduction protéique, sans interférer avec les tRNA de la bactérie hôte[192]. Ce système orthogonal a été optimisé par une approche de mutagenèse afin d'améliorer son efficacité de reconnaissance du codon ambre, et pour être compatible avec toute une série d'AAnn différents[192–195].

Brièvement, la stratégie expérimentale pour l'insertion d'un AAnn consiste en :

- La génération d'une souche de *E. coli* exprimant le système orthogonal
- La transformation de la souche avec la séquence de la protéine d'intérêt portant un codon ambre à la position souhaité pour l'insertion de l'AAnn
- La culture de la souche dans un milieu supplémentée en AAnn

Depuis, cette approche a été adaptée pour être utilisée dans divers organismes tels que les levures ou les eucaryotes[196–198].

Il est à noter que presque 30% des gènes de *M. tuberculosis* utilisent un codon ambre comme codon-stop[199]. Ce nombre est bien plus faible chez *E. coli*, où seulement 9% des gènes sont terminés par un codon TAG[200]. Chez les mycobactéries, l'utilisation du couple MjtRNA/MjTyrRS reconnaissant les codons ambre peut donc potentiellement avoir un effet délétère plus important sur la biologie de la bactérie. En effet, les codons ambre naturels peuvent être reconnus par la paire orthogonale et conduire à l'insertion de l'AAnn entrainant la production de protéines aberrantes. A l'inverse, le système endogène de traduction peut reconnaitre le codon ambre inséré dans la protéine d'intérêt comme un facteur de

terminaison entrainant la production de protéine d'intérêt tronquée dépourvue d'AAnn. Les deux systèmes sont donc en compétition pour la reconnaissance des codons ambre. Il faut que le couple MjtRNA/MjTyrRS ne soit pas trop prévalent sur le système endogène pour ne pas entrainer d'effets néfastes sur la biologie de la bactérie, mais il faut aussi que ce couple soit exprimé en quantité suffisante pour permettre la production de la protéine d'intérêt portant l'AAnn. Récemment, l'équipe de recherche de Peter Schultz a adapté ce système orthogonal pour l'insertion d'AAnn chez *Mycobacterium smegmatis*[201]. Nous nous sommes donc basés sur ses résultats de biologie synthétique pour concevoir notre stratégie de fonctionnalisation *in vivo* de la boucle EL4 de PMTub par introduction métabolique d'un AAnn photoréactif.

1. Conception des outils génétiques

a) Identification des éléments génétiques d'intérêt

L'objectif de notre stratégie est de piéger un substrat de PMTub sur la boucle EL4 en réalisant un pontage covalent par acide aminé non naturel photoréactif. Pour atteindre cet objectif, nous avons entrepris la construction d'une souche de *M. smegmatis* capable d'exprimer conjointement PMTub, le système orthogonal MjtRNA/MjTyrRS et un substrat protéique de PMTub. De précèdent travaux ont permis de développer les outils nécessaires à l'utilisation du système orthogonale chez les mycobactéries[201] et ces outils ont depuis été utilisés pour étudier l'interactome d'une protéine mycobacterienne[202]. Lors de ces approches, les auteurs ont utilisé deux plasmides distincts permettant chacun d'apporter un gène du système orthogonal MjtRNA/MjTyrRS. Chaque plasmide porte un gène de résistance aux antibiotiques différents pour permettre sa sélection. Afin d'utiliser les outils biologiques existants en l'état, il nous aurait fallu apporter un troisième vecteur codant pour PMTub et le substrat protéique à piéger. Ce vecteur devant avoir une origine de réplication compatible avec les deux autres plasmides déjà employés et apporter un nouveau gène de résistance aux antibiotiques.

Pour éviter la démultiplication des systèmes de sélection pouvant se traduire par une forte instabilité de la souche, nous avons décidé de nous limiter à l'emploi de deux plasmides. La reconstruction des outils de biologie synthétique a alors été nécessaire pour les adapter à nos besoins.

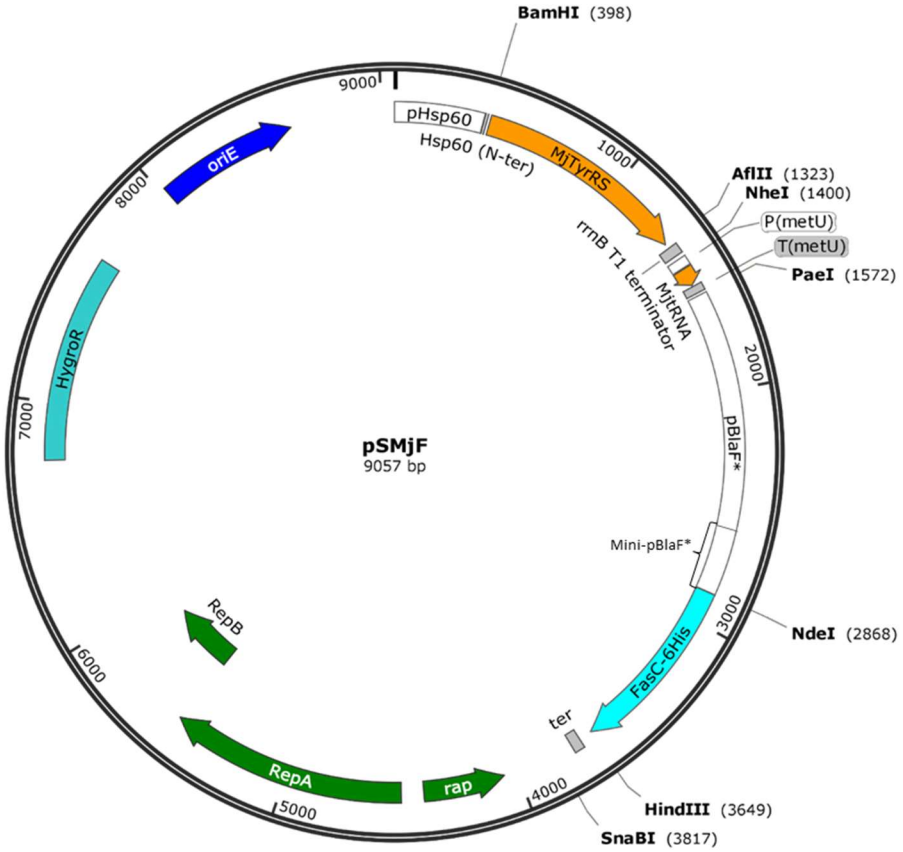


Figure 54. Assemblage *in silico* de pSMjF (« long »). Représentation du plasmide pSMjF pour l'expression du système orthogonale (ARNt synthétase et ARNt, gènes en orange) et du substrat rapporteur (FasC, gène en cyan). Les origines de réplication mycobactérienne, ou de ***E. coli***, et le gène de résistance à l'hygromycine, sont représentés en vert, bleu foncé et bleu clair respectivement. Les promoteurs sont représentés en blanc, et les terminateurs en gris. Notablement, le plasmide pSMjF final ne contient pas la séquence complète du promoteur pBlaF*,mais seulement une portion de celui-ci, appelée « mini-pBlaF* (annotée sur la carte), résultant en un plasmide de 7909pb. Les sites de restriction uniques sont indiqués.

Parmi les éléments génétiques devant être exprimés par notre système, nous avons pu définir deux catégories : 1) les éléments invariants : le système orthogonal et le substrat protéique dont les gènes n'ont pas besoin d'être modifiés au cours des différents développements expérimentaux ; et 2) l'élément variable : le gène de PMTub devant être muté en différentes positions par le codon ambre pour insérer l'AAnn dans des positions choisies de la séquence de pMTub. Compte tenu des contingences spécifiques de notre système : à savoir l'utilisation de 2 plasmides, nous avons décidé de regrouper tous les éléments invariants sur l'un, et de positionner la séquence codant pour PMTub sur le second. Ainsi seul ce plasmide devrait nécessiter d'être modifié pour balayer les positions d'intérêt compatibles pour l'insertion de l'AAnn dans PMTub pouvant être modifiées selon les besoins expérimentaux.

(1) Expression de PMTub :

Pour exprimer PMTub, nous avons choisi d'utiliser le plasmide intégratif construit pour la mise au point notre test d'activité phénotypique. Ce plasmide a par la suite été modifié par mutagenèse dirigée pour permettre l'insertion d'un codon ambre dans la séquence de PMTub à la position souhaitée. Ceci a été réalisé de manière similaire à la méthodologie utilisée pour effectuer des mutations ponctuelles de PMTub lors de l'étude des relations structure/fonction (voir Partie 2.C, p.129).

(2) Expression des éléments invariants :

Les éléments génétiques invariants seront quant à eux portés par un second plasmide dénommé pSMjF. Ce second plasmide code donc pour les gènes du système orthogonal MjtRNA/MjTyrRS et pour le substrat protéique de PMTub. Pour ce dernier, nous avons choisi en première intention, d'utiliser la fascicline (la mannoprotéine rapportrice du test d'activité phénotypique). Dans un second temps (à la suite de ma thèse), le gène de la FasC pourra être remplacé par le substrat détectable choisi. Pour la mise au point du test, la fascicline présentait en effet l'avantage de porter de nombreux sites de mannosylation, maximisant les contacts avec PMTub et donc les chances de capture par l'AAnn. La FasC doit permettre également de vérifier que l'insertion d'un AAnn dans PMTub n'altère pas son activité.

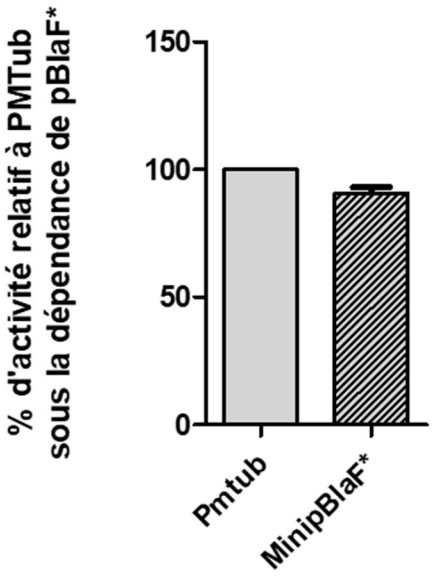


Figure 55. Effet de la réduction de taille du promoteur pBlaF* sur l'activité de PMTub. Résultats du test d'activité phénotypique pour PMTub exprimée sous le contrôle de pBlaF* ou de mini-pBlaF*. 0% et 100% sont définis comme les résultats obtenus pour les souches « No PMT » ou exprimant PMTub sous la dépendance de pBlaF* respectivement. Les résultats présentés sont des triplicats techniques de duplicats biologiques.

Pour obtenir une expression suffisante du substrat protéique et du système orthogonal MjtRNA/MjTyrRS, ce second plasmide, pSMjF sera un plasmide réplicatif à haut nombre de copies ce qui permettra un niveau d'expression élevé des différents éléments génétiques qui le compose.

a) Développement de mini-pBlaF*

La conception *in silico* de ce plasmide conduit à un plasmide pSMjF de 9057 paires de bases (**Figure 54**). Il est communément admis que la stabilité d'un plasmide est inversement corrélé à sa taille[203,204]. Or, lors du développement du test phénotypique (voir p.159) nous avons déjà pu remarquer une certaine instabilité du plasmide réplicatif de presque 9000 paires de bases, utilisé pour l'expression de la fascicline. Nous avons donc décidé de réduire la taille du plasmide pSMjF au maximum pour en diminuer l'instabilité et optimiser son utilisation.

L'élément le plus imposant de ce plasmide est le promoteur pBlaF* qui fait 1.3kb (**Figure 54**). Ce promoteur est l'un des promoteurs constitutifs les plus forts identifiés à ce jour pour l'expression de gènes mycobactériens. Il dérive du promoteur du gène de la β-lactamase de *Mycobacterium fortuitum* (gène blaF). Ce promoteur a été défini comme correspondant aux 1274 nucléotides en amont du codon ATG du gène blaF muté sur deux nucléotides aux positions -9 et -326[205]. La version mutée du promoteur de blaF, nommé pBlaF*, montre une activité 10 fois supérieure au promoteur d'origine[205]. Une étude précédente a permis de mettre en évidence que seule la mutation à la position -9 expliquait ce gain d'efficacité. Du fait de l'absence de rôle de la mutation en -326, nous nous sommes demandés si l'entièreté de la séquence de ce promoteur était nécessaire pour son rôle de promoteur constitutif fort, et si non, si la taille de ce promoteur pouvait être réduite sans affecter son activité.

L'étude de la séquence de pBlaF* montre qu'elle contient plusieurs cadres de lecture. Ces cadres de lecture semblent correspondre à des restes fortuits du premier isolement du promoteur. En effet, le promoteur pBlaF* fut isolé pour la première fois en 1994 en utilisant des enzymes de restrictions[205] et ne semble pas avoir fait l'objet d'une optimisation particulière. Il est de ce fait très probable que la séquence isolée ne corresponde pas à la séquence minimale du promoteur et contienne des gènes en amont.

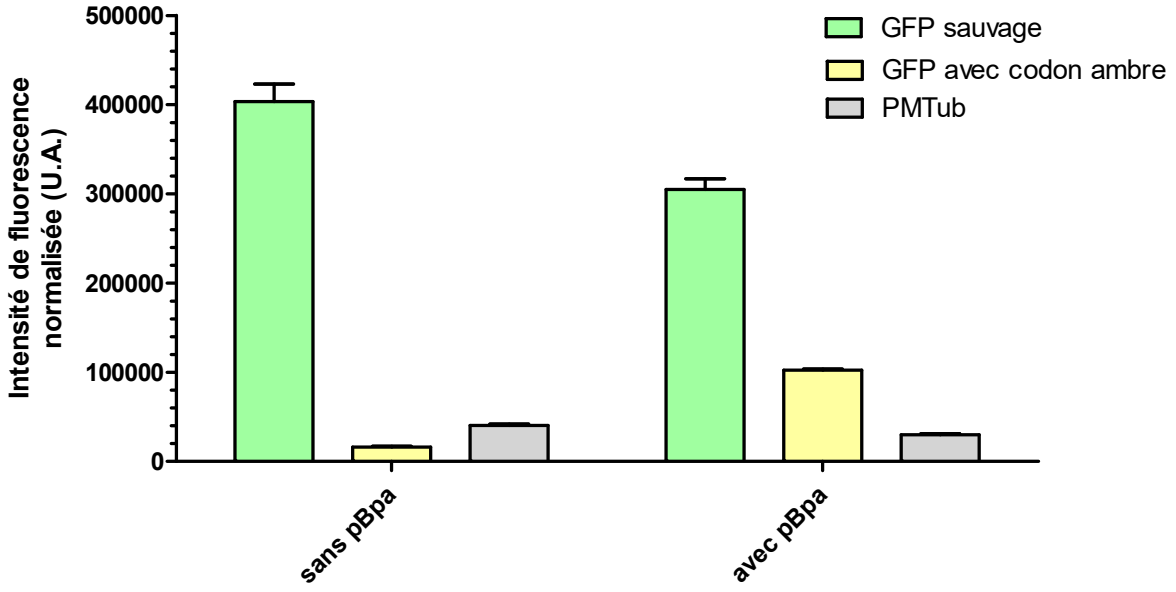


Figure 56. Contrôle de la fonctionnalité du système orthogonal dans la souche de *M. smegmatis* exprimant les éléments génétiques du plasmide pSMJF. Une souche exprimant la GFP sauvage est utilisée comme témoin positif. Une souche distincte, exprimant la GFP mutée par introduction d'un codon ambre en position 47 de la protéine est étudiée. Les deux souches sont également co-transformées avec le plasmide pSMjF. L'intensité de fluorescence (dans le vert) des cellules est mesurée après culture en présence ou en absence de l'acide aminé non naturel pBpa, et est représentée normalisée par la densité optique bactérienne mesurée à DO_{600nm}. Les mesures ont aussi été effectué sur une souche transformée pour exprimer PMTub sauvage et pSMjF comme contrôle.

Nous avons donc décidé de supprimer la séquence correspondant potentiellement à des cadres de lecture, et de vérifier si cela avait un impact sur les propriétés de pBlaF*. Le nouveau promoteur ainsi réduit en taille a été appelé mini-pBlaF*. Nous avons testé l'efficacité de ce promoteur réduit à l'aide de notre test d'activité phénotypique. En effet, le gène de PMTub apporté dans la souche de *M. smegmatis* ΔPMT se situe aussi sous le contrôle du promoteur pBlaF*. Nous avons donc délété de ce promoteur tous les cadres de lecture pour obtenir le promoteur mini- pBlaF* et nous avons évalué l'impact de de cette réduction sur l'expression de PMTub en mesurant l'activité mannosyl-transferase des souches transformées avec cette nouvelle construction. Nous avons ainsi pu observer que la réduction du promoteur pBlaF*n' entrainait pas de perte d'activité mannosyltransférase significative comparée à l'activité mesurée avec le promoteur pBlaF* natif (**Figure 55**).

En utilisant les nouvelles délimitations de ce promoteur réduit, nous avons ainsi pu concevoir *in silico*, un plasmide pSMjF de 7909 paires de bases (**Figure 54**).

2. Construction de pSMjF

a) Assemblage des différents éléments

La conception *in silico* du vecteur a été poursuivie par son assemblage *in vitro*. Pour cela, nous avons utilisé la méthode d'assemblage de Gibson[206]. Cette méthode permet de récupérer différents fragments d'ADN depuis différentes sources afin de les assembler en un produit final. Nous avons récupéré les différents éléments de plasmides utilisés dans le laboratoire ou depuis les plasmides fournis par l'équipe de Peter Schultz[201]. Ce produit a été appelé « pSMjF ». Afin de pouvoir utiliser ce plasmide pSMjF dans des contextes expérimentaux différents, nous avons introduit de multiples sites de restriction uniques à des positions stratégiques sur la séquence (**Figure 54**). Ces éléments rendent ainsi ce plasmide facilement modifiable et cette flexibilité doit permettre par exemple de remplacer, selon les besoins, le gène du substrat protéique, ou bien de le fusionner simplement à un rapporteur fluorescent pour suivre sa localisation dans la cellule.

L'assemblage du vecteur a été validé par son séquençage en intégralité. Il convenait ensuite d'en vérifier la fonctionnalité. Ceci a nécessité le contrôle de l'expression de ses différents éléments génétiques. Il a donc fallu, dans un premier temps, déterminer si pSMjF

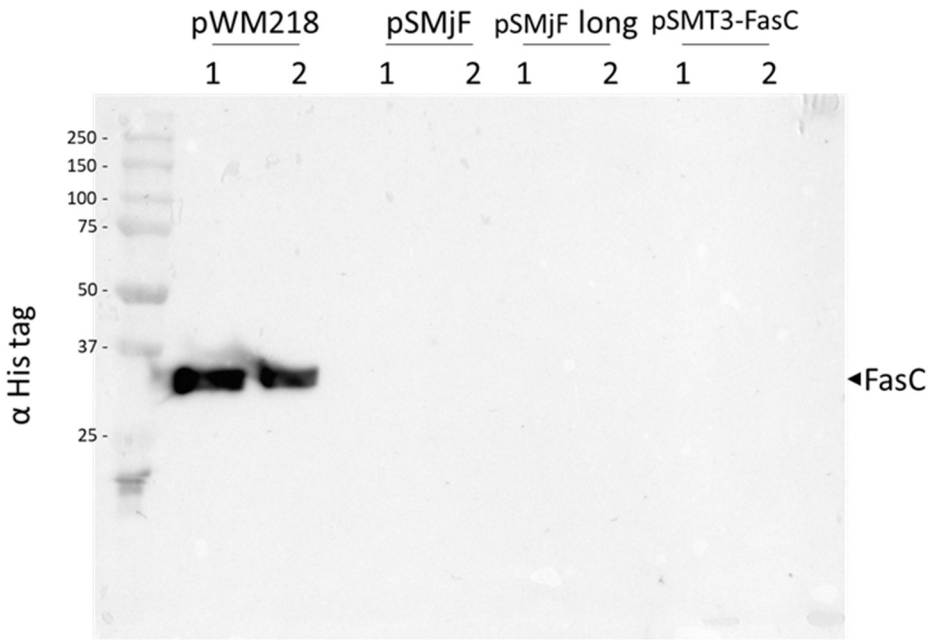


Figure 57. Contrôle de l'expression de la FasC. Western blot anti-His de surnageants de culture de *M. smegmatis* transformée avec les plasmides pWM218 (plasmide contrôle de l'expression de FasC), pSMjF, pSMjF « long » et pSMT3-FasC, en duplicats biologiques.

permettait à une souche de *M. smegmatis* d'exprimer un système orthogonal fonctionnel, et ainsi d'insérer des AAnn dans la séquence d'une protéine d'intérêt. Et dans un second temps, il a fallu vérifier que pSMjF permettait l'expression de la fascicline.

a) Vérification de la fonctionnalité du système orthogonal

Pour vérifier la fonctionnalité du système orthogonal porté par le plasmide pSMjF, nous avons construit un système rapporteur qui utilisait le gène de la GFP dans lequel nous avions introduit un codon ambre à une position stratégique. Nous avons choisi de muter le gène de la GFP en substituant l'isoleucine à la position 47 par un codon ambre, tel que précédemment décrit[202].

Ce contrôle permet de suivre la fonctionnalité du système orthogonal en mesurant la fluorescence des bactéries. En effet, un système orthogonal fonctionnel permet d'insérer un AAnn à la position 47 dans la GFP, entrainant la production d'une GFP entière et fluorescente. S'il est non-fonctionnel, le système orthogonal ne permet pas l'introduction de l'AAnn et le codon ambre en position 47 est alors reconnu comme un codon-stop, entrainant un arrêt de la traduction de la GFP, se traduisant par la production d'une protéine tronquée non fluorescente.

Comme le montre la figure 28 nous avons pu observer que la souche exprimant la GFP mutée n'est fluorescente qu'après addition du plasmide pSMjF et en présence de pBpa (**Figure 56**). Ce résultat a donc confirmé que le plasmide pSMjF permettait l'expression d'un système orthogonal MjtRNA/MjTyrRS fonctionnel et par conséquent l'insertion d'acide aminé non naturel dans une protéine d'intérêt chez *M. smegmatis*.

On peut néanmoins observer une différence d'intensité de fluorescence entre la souche exprimant le contrôle GFP sauvage et la souche exprimant la GFP mutée à la position 47. Cette différence permet d'estimer une efficacité d'insertion de l'AAnn d'environ 30%. Notablement, l'efficacité d'insertion varie en fonction de la protéine cible, mais aussi en fonction du système d'expression du couple MjtRNA/MjTyrRS[201,202]. Le nouvel outil plasmidique que nous avons construit permet donc une efficacité de reconnaissance suffisante pour produire la protéine d'intérêt, sans avoir d'effet sur la croissance bactérienne (donnée non montrée).

a) Vérification de la production de FasC

Le plasmide pSMjF devait, en plus de la paire MjtRNA/MjTyrRS, permettre l'expression de la FasC^{6His}. La FasC est une protéine sécrétée étiquetée His qui peut donc être facilement détectée dans le surnageant de culture. L'analyse par western blot anti-His des surnageants de culture n'a malheureusement pas permis de mettre en évidence la production de ce substrat protéique par pSMjF (**Figure 57**).

Plusieurs hypothèses ont été formulées pour expliquer l'absence de production de la FasC. Dans l'objectif de réduire la taille du plasmide pSMjF nous avons réduit la taille du promoteur pBlaF* utilisé pour exprimer la FasC. Après étude de la littérature et de la séquence de ce promoteur, nous avons pu tronquer sa séquence de 1000pb pour obtenir ce que nous appelons « mini-pBlaF* ». Nous avions testé l'efficacité de ce promoteur réduit à l'aide de notre test d'activité phénotypique (voir p.173 et **Figure 55**). Cependant, l'efficacité de ce promoteur raccourci n'a jamais été testée pour l'expression de la FasC. Nous avons donc émis l'hypothèse que la perte d'expression de la FasC pouvait être due à l'utilisation de ce promoteur « mini-pBlaF* » tronqué. Nous avons donc reconstruit le plasmide pSMjF afin que l'expression de la FasC soit sous le contrôle du promoteur pBlaF* pleine taille. Toutefois ainsi que le montre la Figure 57, ce plasmide pSMjF « long » ne permet pas non plus l'expression de la FasC (**Figure 57**).

Une autre hypothèse a été que le système orthogonal pouvait, pour des raisons indéterminées, réprimer l'expression de la FasC^{6His}. Pour vérifier cette hypothèse, une version de pSMjF a été construite dans laquelle nous avons retiré les gènes pour l'expression du couple MjtRNA/MjTyrRS. Ce nouveau plasmide, nommé pSMT3-FasC, n'a pas non plus permit d'exprimer la FasC^{6His} (**Figure 57**).

Nous avons alors émis l'hypothèse que le plasmide pSMjF pouvait subir une recombinaison dans la bactérie entrainant la perte du gène de la FasC^{6His}. L'étude du plasmide ré-isolé à partir de la souche transformée de *M. smegmatis* a permis de mettre en évidence la perte d'une région d'environs 3000pb correspondant à la séquence codant pour le système orthogonal et la FasC (données non montrées). Ces gènes sont donc excisés de pSMjF chez *M. smegmatis*. Toutefois, il est important de remarquer que cette recombinaison probable semble n'avoir d'impact que sur l'expression de la FasC^{6His} puisque la souche continue

d'exprimer un système orthogonal efficient capable d'introduire l'AANnn dans la séquence de la GFP comme l'ont démontré les résultats du test de fonctionnalité décrit plus haut. Une des explication possible à ce résultat surprenant serait que la surproduction de ces 3 gènes pourrait être défavorable à la croissance bactérienne, entrainant une recombinaison du plasmide. Toutefois rien à ce jour ne permet d'expliquer la perte sélective de la $FasC^{6His}$.

Il apparait donc que malgré la perte des séquences codantes , le plasmide pSMjF retient tout de même la capacité d'induire la production d'un système orthogonal fonctionnel chez *M. smegmatis*, mais pas l'expression du substrat protéique de PMTub

3. Insertion d'AAnn dans PMTub

Un autre défi de cette stratégie est d'insérer l'AAnn pBpa dans la boucle EL4 sans prévenir l'activité de PMTub. Il faut pour cela donc tester plusieurs positions dans la séquence de PMTub. La vérification de l'insertion de l'AAnn peut se faire par la confirmation de la production de PMTub entière. Toutefois, Puisque le plasmide pSMjF ne permet pas l'expression de la FasC, le substrat rapporteur de mannosylation, nous ne pouvons pas, pour le moment, effectuer le contrôle de l'activité pour vérifier que l'insertion de l'AAnn n'interfère pas avec l'activité de PMTub.

Notre objectif étant dans un premier temps de vérifier s'il est possible d'insérer un AAnn dans la séquence de PMTub, nous avons cherché à détecter son extrémité C-terminale. La détection de cette extrémité confirme l'expression de la protéine entière en apportant la preuve qu'elle n'a pas été tronquée au niveau du codon ambre. Pour mettre en évidence l'expression de l'extrémité C-terminale de PMTub, nous nous sommes donc appuyés sur les deux stratégies développées plus haut: l'utilisation de la phosphatase alcaline et l'utilisation de mCherry (voir p.86).

Nous avons donc construit des plasmides exprimant PMTub-PhoA ou PMTub-mCherry portant un codon ambre pour nous permettre de sélectionner le système de contrôle de l'insertion de l'AAnn le plus efficace. La mutagenèse a été effectuée de sorte à introduire le codon ambre dans la séquence en position F117TAG et V378TAG localisées respectivement dans les boucles EL1 et EL4 de PMTub. Deux constructions ont été réalisées pour chaque

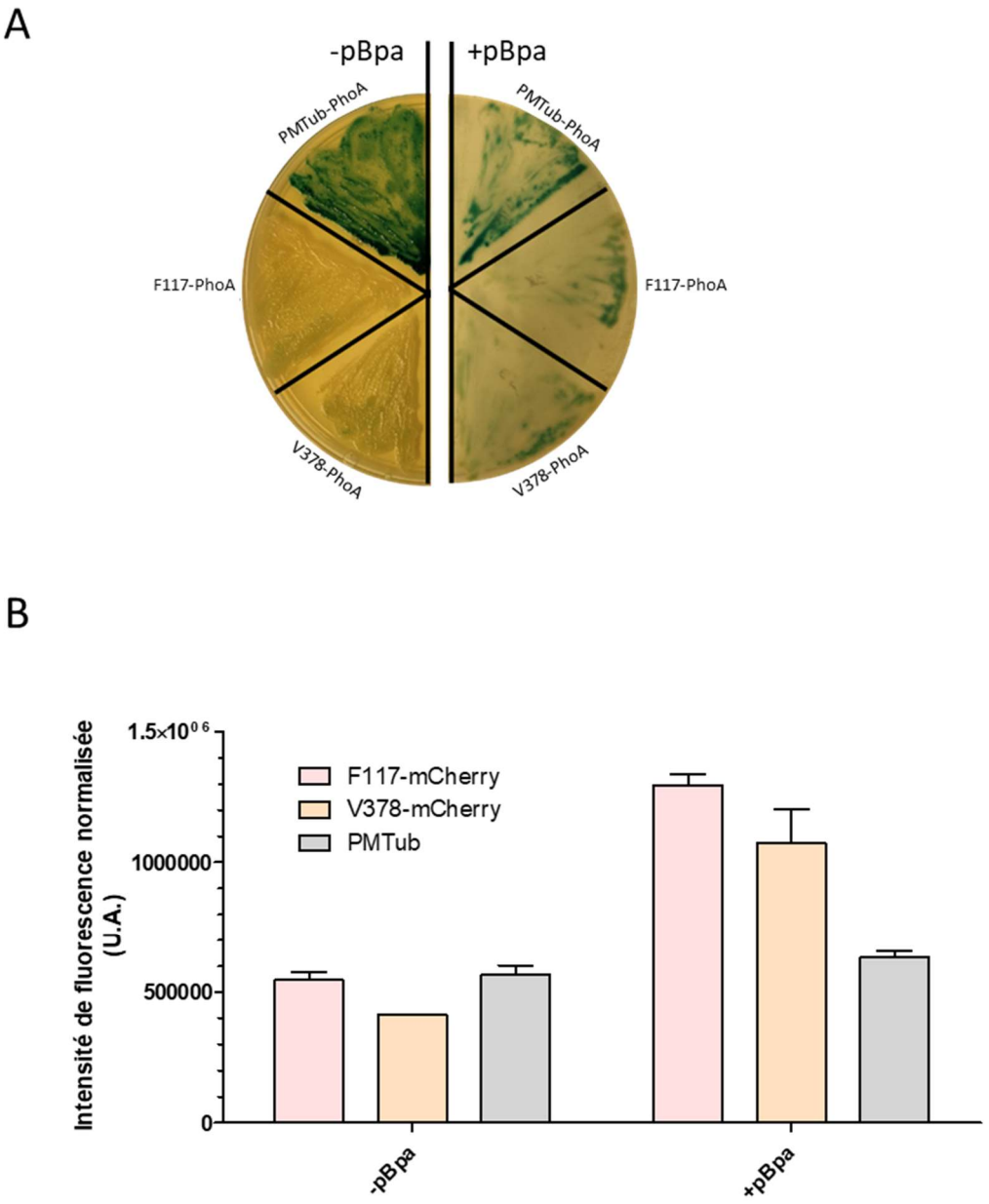


Figure 58. Contrôle de l'insertion d'un acide aminé non naturel dans PMTub. **A)** Etalement des différentes souches sur boite contenant du BCIP permettant de détecter l'activité PhoA par coloration des souches en bleu. La moitié gauche est cultivé en l'absence de pBpa et la partie droite est supplémenté en pBpa. **B)** Mesure d'intensité de fluorescence par spectrofluorimétrie de bactéries exprimant la protéine de fusion PMTub-mCherry en présence ou en absence de pBpa. Normalisé par la quantité de bactéries. Les mesures ont aussi été effectué sur une souche transformée pour exprimer PMTub sauvage et pSMjF comme contrôle.

système de contrôle de l'insertion de l'AAnn : PMTub^{F117TAG}-PhoA, PMTub^{V378TAG}-PhoA, PMTub^{F117TAG}-mCherry et PMTub^{V378TAG}-mCherry.

Pour vérifier l'expression de la phosphatase alcaline, lLa souche *M. smegmatis* ΔPMT a été donc transformé successivement par le plasmide permettant l'expression des protéines PMTub mutées en fusion C terminal avec PhoA, puis par le plasmide pSMjF porteur du système orthogonal permettant l'insertion métabolique de l'AAnn.

L'expression de PhoA a tout d'abord pu être apprécié par le test qualitatif (blanc/bleu) sur milieu de culture supplémenté en BCIP (voir p.96). Les différentes souches ont été cultivé sur boite en l'absence ou en présence de l'AAnn pBpa. Comme il peut être observé, les colonies se sont colorées en bleu pour F117-PhoA et V378-PhoA en présence de pBpa alors qu'elles sont restées blanche en absence de l'AAnn (**Figure 58A**). Ceci témoigne d'une activité PhoA en présence de pBpa et démontre que l'AAnn a été incorporé dans la séquence de PMTub entrainant une expression de l'enzyme à la membrane plasmique avec PhoA correctement localisé dans le periplasme de la bactérie.

Pour quantifier l'expression de PMTub comportant un AAnn, nous avons voulu mettre en place le test quantitatif de détection de l'activité de PhoA en milieu liquide. Ce test indirecte est celui que nous avons utilisé pour démontrer l'expression de mutants inactifs de PMTub. Aucun signal n'ayant pu être mesuré pour PMTub^{F117TAG}-PhoA et PMTub^{V378TAG}-PhoA (données non montrées), ce test ne nous a pas permis de détecter d'activité PhoA et de quantifier l'expression de PMTub après insertion d'un AAnn.

Nous avons donc pu confirmer l'insertion d'un AAnn dans la séquence de PMTub en utilisant la phosphatase alcaline. Cependant, le niveau d'expression de la protéine de fusion PMTub-PhoA est trop faible pour permettre une lecture quantitative de l'expression.

Le second système de vérification de l'insertion d'un AAnn dans PMTub a été l'utilisation d'une protéine fluorescente en fusion C-terminale. Nous avons pu montrer que la fusion de mCherry à l'extrémité C-terminale de PMTub ne permettait pas de confirmer la localisation de PMTub a la membrane plasmique en microscopie (voir p.86 et Figure 22) mais permet en revanche d'obtenir par spectrofluorimétrie une mesure quantitative de l'intensité de fluorescence indicatrice du niveau d'expression de la protéine de fusion.

Nous avons alors comparé l'intensité de fluorescence mesurée pour les souches exprimant les fusions PMTub^{F117TAG}-mCherry et PMTub^{V378TAG}-mCherry après culture en présence ou en absence de pBpa (**Figure 58B**). Pour les deux souches, il a été possible de déterminer une augmentation de l'intensité de fluorescence après culture en présence de pBpa. Ceci témoignait de l'expression de mCherry et donc de l'insertion de l'AAnn dans la séquence de PMTub. L'utilisation de mCherry n'apparait donc pas idéale du fait des difficultés de fluorescence de cette protéine dans le périplasme, mais celle-ci permet tout de même de valider l'insertion de l'AAnn dans la séquence de PMTub.

Les deux systèmes, PhoA et mCherry, semblent donc adaptés à la confirmation de l'insertion d'AAnn dans la séquence de PMTub.

1. Conclusion et perspectives

La stratégie développée dans cette partie avait pour objectif s'initier la mise en place d'un test de criblage innovant d'inhibiteurs de l'interaction entre EL4 et un substrat protéique de PMTub.

Nous avons pu développer un outil plasmidique permettant, théoriquement, de développer ce test. Il était question de piéger un substrat de PMTub sur la boucle EL4 en utilisant un acide aminé non naturel pour pouvoir détecter cette interaction. Le plasmide pSMjF développé a fourni à la bactérie les outils biologiques nécessaires à l'insertion d'un AAnn dans PMTub. Pour vérifier cette insertion, l'utilisation d'une protéine de fusion, PMTub-PhoA ou PMTub-mCherry, s'est montrée effective.

Cependant, il s'est avéré que, bien que capable d'exprimer un système orthogonal fonctionnel, le plasmide pSMjF ne permet pas l'expression du substrat protéique FasC. Or, ce substrat est nécessaire à la détection de l'interaction.

La biologie synthétique semble donc être un outil utilisable pour l'étude de PMTub. Cependant, pour atteindre nos objectifs expérimentaux, plusieurs problèmes limitant son emploi demeurent. Le développement des outils plasmidiques doit être poursuivi pour

permettre l'expression de tous les éléments génétiques nécessaires à notre test. Pour cela, on peut imaginer plusieurs alternatives à la configuration actuelle. Nous avons en effet mis en évidence que la FasC peut être exprimée sur un plasmide intégratif (voir p.159), et son expression pourrait donc être couplée à celle de PMTub sur un plasmide indépendant du plasmide utilisé pour l'expression du système orthogonal. Pour résoudre le problème de perte du gène de la FasC, des systèmes de production alternatifs moins contraignants pour la bactérie pourraient aussi être testés, notamment l'utilisation d'un promoteur inductible ou la production de nos protéines en fusion transcriptionnelle.

Lorsque nous aurons réussi à produire l'ensemble des éléments génétiques nécessaires à la fonctionnalisation ponctuellement localisée de la boucle EL4 de PMTub, il nous faudra alors construire un substrat détectable dans le périplasme mycobactériens. Nous avons pu précédemment observer les difficultés associées à l'expression dans le périplasme de protéines fluorescentes en fusion à PMTub (voir p.86). Ce moyen de détection devra alors être optimisé pour être utilisable dans notre contexte. D'autre système fonctionnant dans le périplasme bactérien pourront aussi être envisagé tel que les luciférases ou les étiquettes « self labelling »[207,208].

La position optimale pour l'insertion de l'AAnn dans EL4 devra aussi être déterminée pour mettre en place le test de crible d'inhibiteur de l'interaction.

Bien que nous n'ayons pu achever le développement de ce test dans le courant de cette thèse, les résultats obtenus par l'approche originale utilisée ici sont prometteurs et constituent les fondements essentiels de la poursuite du projet. En effet, nous apportons ici la preuve de faisabilité de l'insertion métabolique d'acides aminés non naturels pour la fonctionnalisation localement ciblée de PMTub *in-vivo*. Bien que notre objectif premier était de les utiliser pour suivre l'interaction entre EL4 et un substrat protéique détectable, les AAnn sont des outils innovants qui pourront être utilisés pour poursuivre l'étude structurale et fonctionnelle de PMTub. L'utilisation de ces outils permet en effet d'envisager une multitude d'approches expérimentales nouvelles dont, par exemple, une compréhension fine des interactions entre PMTub et ses partenaires putatifs.

Conclusion générale

Le contexte épidémiologique de la tuberculose est alarmant. L'OMS considère que 10 millions de nouveaux cas seront diagnostiqués en 2019 et que la maladie fera plus de 1,5 million de victimes cette même année[1]. La pandémie de Covid-19 ne devrait pas améliorer la situation et pourrait même, selon les estimations de l'OMS, entrainer un recul de la lutte contre la tuberculose.

Depuis 2014, le plan international « End TB » voté par l'OMS vise à contrôler la maladie et mettre fin à l'épidémie mondiale de tuberculose. Ce programme repose sur trois points : une meilleure prévention et prise en charge des patients, un engagement politique dans le soin de la tuberculose, et des efforts dans le développement de nouvelles stratégies thérapeutiques. Le développement de nouvelles molécules actives contre *M. tuberculosis* se concrétise principalement en la recherche de nouveaux antibiotiques. Cependant, l'émergence de souches de *M. tuberculosis* résistantes aux antibiotiques encourage le développement de stratégies alternatives. Parmi ces stratégies, on compte les thérapies anti-virulences. Celles-ci visent des facteurs de virulence non essentiels à la survie de la bactérie permettant de diminuer la pathogénicité de celle-ci sans induire de pression de sélection telle que le ferait un antibiotique.

De nombreux facteurs de virulences ont été identifiés chez *M. tuberculosis*. Parmi ces facteurs de virulence, on trouve les mannoprotéines. Celles-ci sont des protéines *O*-glycosylées, portant des mannoses sur certaines de leurs serines ou thréonines. La preuve de l'implication des mannoprotéines dans la virulence mycobactérienne a été apportée par plusieurs équipes de recherche, dont la nôtre. En effet, la mannosylation de toutes ces protéines est initiée par une enzyme unique : la protéine *O*-mannosyltransférase de *Mycobacterium tuberculosis* (PMTub). Dans ce contexte, l'équipe a démontré que la réaction catalysée par PMTub n'est pas essentielle pour la croissance bactérienne, mais cruciale pour la virulence[52]. PMTub semble donc une cible anti-virulence attractive. Notablement, les PMT humaines sont des enzymes essentielles et elles sont fortement conservées avec PMTub. Par conséquent, pour que PMTub constitue une cible thérapeutique acceptable, il faut donc être capable d'identifier des inhibiteurs spécifiques de l'enzyme mycobactérienne et qui ne ciblent pas les isoenzymes humaines.

Au cours de ma thèse, j'ai donc étudié PMTub dans l'objectif de développer un test de crible d'inhibiteurs spécifiques de cette enzyme. Pour cela, dans un premier temps, nous avons développé les outils nécessaires à l'étude fonctionnelle de PMTub dans la bactérie. Un test d'activité phénotypique fiable et robuste permettant d'évaluer l'activité enzymatique de PMTub *in-situ* sur bactérie vivante a été développé. Ce test a nécessité la mise au point d'un contrôle de l'expression et de la topologie correcte de la protéine PMTub entière et fonctionnelle à la membrane plasmique. Le développement de ce contrôle a apporté de nouvelles connaissances sur PMTub. La localisation membranaire de PMTub, jusqu'alors présumée sur la base d'observations indirectes et d'analogies au système eucaryote[55,75], a pu être confirmée par l'utilisation d'une protéine fluorescente. Aussi, une fusion à l'enzyme PhoA a confirmé la localisation périplasmique de l'extrémité C-terminale de PMTub.

Dans un second temps, ce test phénotypique a été utilisé de façon originale pour caractériser les relations structure-fonction de PMTub. Nous avons élaboré un modèle 3D de PMTub basé sur les structures résolues de PMT de levure et d'autres glycosyltransférases membranaires. Ce modèle nous a permis de définir plusieurs acides aminés comme intéressants à étudier d'un point de vue fonctionnel. Des mutants ponctuels de PMTub correspondants ont donc été construits et l'effet de ces mutations sur l'activité de l'enzyme a été mesurées grâce au test d'activité développé. En particulier, nous avons identifié quatre acides aminés comme essentiels pour la catalyse de la réaction de transfert de mannose : D74, D176, R441 et Y444. La littérature et la mise en place d'une stratégie de « chemical rescue » a guidé l'attribution de rôles pour ces acides aminés dans la réaction catalysée par PMTub et a permis de proposer, pour la première fois, un mécanisme catalytique pour les PMT.

En utilisant la même méthodologie, nous avons aussi pu identifier la boucle EL4 comme un domaine fonctionnel de PMTub. Cette boucle présente un intérêt tout particulier puisqu'elle est divergente entre les PMT bactériennes et eucaryotes et du fait de son implication dans l'activité, elle pourrait être la cible d'inhibiteurs spécifiques des PMT bactériennes. Grace à l'étude par mutagenèse de EL4, nous avons pu formuler une hypothèse sur le mode d'action de cette boucle. Elle pourrait, en effet, participer à l'interaction avec les substrats protéiques de PMTub par un processus similaire à celui décrit pour les domaines WW. Ces domaines sont des domaines d'interaction protéine-protéine particuliers récurrents

dans le monde du vivant, mais jamais décrits dans les bactéries, et EL4 pourrait donc agir comme un « pseudo domaine WW ». Des études préliminaires d'interaction *in vitro* semblent confirmer cette hypothèse.

Les résultats obtenus par l'étude des relations structure-fonction de PMTub grâce au test phénotypique d'activité développé sont en accord avec les données de la littérature sur les PMT de levure et les autres glycosyltransférases membranaires. Cette concordance permet donc de valider le test phénotypique comme outil d'étude quantitatif de l'activité de PMTub qui peut alors être adapté au crible d'inhibiteurs.

Les connaissances acquises lors de l'étude des relations structure-fonction nous ont tout d'abord permis d'effectuer une sélection rationnelle d'inhibiteurs putatif de EL4. Pour mesurer l'effet de ces molécules sur l'activité de PMTub, nous avons adapté le test d'activité phénotypique à un test de criblage a moyen débit. Bien que les molécules sélectionnées n'aient pas eu l'effet escompté sur l'activité de PMTub, nous avons apporté la preuve de concept de l'utilisation de notre test d'activité pour le crible d'inhibiteurs putatifs de PMTub sur bactéries vivantes.

De plus, nous avions pour objectif d'élaborer une stratégie innovante pour le crible d'inhibiteurs spécifiques de l'interaction entre EL4 et la protéine substrat. Cette stratégie a nécessité la mise au point d'outils moléculaires nouveaux incluant notamment l'utilisation d'acides aminés non naturels chez *M. smegmatis*. Lors de cette thèse, j'ai pu mettre au point un outil plasmidique permettant l'expression de la machinerie biologique nécessaire à l'incorporation d'acides aminés non naturels dans PMTub. Le développement de ce test n'a pas pu être finalisée pendant la durée de cette thèse et est toujours en cours dans l'équipe. Cependant, j'ai pu mettre en avant la faisabilité de l'utilisation d'acides aminés non naturels pour l'étude de PMTub. L'utilisation de ces outils innovants permet d'envisager une grande diversité d'approches expérimentales qui pourraient permettre une compréhension fine des mécanismes réactionnels de PMTub.

Matériels et méthodes

Tableau 5. Amorces sens utilisées pour la mutagénèse dirigée du vecteur codant pour PMTub. Les séquences soulignées correspondent aux mutations et celles surlignées en gris mettent en évidence les sites de restriction ajoutés ou supprimés par mutation silencieuse pour le contrôle de l'insertion de la mutation par profil de restriction.

Mutation	Sequence	Site de restriction
Site actif		
D74N	5' 5294 CC GAT GCC GGG ACC CCC ATC TTC AAC GAG AAG CAT TAC G 5332 3'	Retire un site NaeI
D74E	5' 5294 CC GAT GCC GGG ACC CCC ATC TTC GAG GAG AAG CAT TAC 5331 3'	Retire un site NaeI
D74A	5' 5294CC GAT GCC GGG ACC CCC ATC TTC GCT GAG AAG CAT TAC5331 3'	Retire un site NaeI
DE74-75ED	5' 5294CC GAT GCC GGG ACC CCC ATC TTC GAG GAC AAG CAT TAC5331 3'	Retire un site NaeI
D176N	5' 5610 G ACC GCG CTG TTA AAC GGC TTC CTG 5634 3'	Retire un site MseI
D176E	5' 5612CC GCG CTG CTC GAG GGC TTC CTG 5634 3'	Ajoute un site XhoI
D176A	5' 5612CC GCG CTG CTG GCC GGC TTC CTG5634 3'	Ajoute un site NaeI
R441L	5' 6410 GAC ATC GAC CTG CAG ATG TAC 6429 3'	Ajoute un site PstI
R441K	5' 6408CC GAC ATC GAT GCG CAG ATG TAC6429 3'	Ajoute un site ClaI
R441A	5' 6408CC GAC ATC GAT AAG CAG ATG TAC6429 3'	Ajoute un site ClaI
Y444A	5' 6405 GCC GAC ATC GAT CGG CAG ATG GCC TTC TTC TAC6439	Ajoute un site PvuI
Y444F	5' 6405 GCC GAC ATC GAT CGG CAG ATG TTC TTC TTC TAC6439	Ajoute un site PvuI
EL4		
W355A	5' **6145** GGC AAC TAC CAC CCT GCA GAA TCC AAA CCG TGG **6177** 3'	Ajoute un site PstI
W355F	5' **6145** GGC AAC TAC CAT CCT TTC GAA TCC AAA CCG **6174** 3'	Ajoute un site FokI
P359A	5' **6166** TCC AAA GCG TGG ACC TGG CCC ATG TCG TTG CGC CCG GTG **6204** 3'	Ajoute unsite Eco52I
W362A	5' 6172 CCG TGG ACC GCG CCC ATG TCG TTG CGG CCG GTG CTC TAC GCC ATC 6216 3'	Ajoute un site SduI
W362F	5' 6172 CCG TGG ACC TTT CCC ATG TCG TTG CGG CCG GTG CTC TAC GCC ATC 6216 3'	Ajoute un site SduI
M364A	5' 6179 CC TGG CCC GCG TCG TTG CGG CCG GTG CTC TAC GCC ATC G 6217 3'	Ajoute un site SduI
M364V	5' 6179 CC TGG CCC GTG TCG TTG CGG CCG GTG CTC TAC GCC ATC G 6217 3'	Ajoute un site SduI
S365A	5' 6179 CC TGG CCC ATG GCG TTG CGG CCG GTG CTC TAC GCC ATC G 6217 3'	Ajoute un site SduI
S365T	5' 6179 CC TGG CCC ATG ACG TTG CGG CCG GTG CTC TAC GCC ATC G 6217 3'	Ajoute un site SduI
L366A	5' 6182 GG CCC ATG TCG GCG CGG CCG GTG CTC TAC GCC ATC G6217 3'	Ajoute un site SduI
Y371A	5' 6192 G TTG CGG CCG GTG CTC GCC GCC ATC GAC CAG 6222 3'	Ajoute un site SduI
Y371F	5' 6192 G TTG CGG CCG GTG CTC TTC GCC ATC GAC CAG C 6223 3'	Ajoute un site SduI
I373Q	5' 6192 G CCG GTG CTC TAC GCC CAG GAC CAG CAA G 6227 3'	Ajoute un site SduI
I373L	5' 6192 G CCG GTG CTC TAC GCC CTC GAC CAG CAA G 6227 3'	Ajoute un site SduI
C381A	5' 6224 GTT GCA GGC GCC GGT GCA CAG TCG TGC 6252 3'	Ajoute un site ApaLI
C381S	5' 6224 GTT GCA GGC AGC GGT GCA CAG TCG TGC 6252 3'	Ajoute un site ApaLI
C386A	5' 6230GC TGC GGT GCA CAG TCG GCC GTC AAG G6259	Ajoute un site ApaLI
C386S	5' 6230GC TGC GGT GCA CAG TCG TCC GTC AAG G6259	Ajoute un site ApaLI
L392A	5' **6259** AAG GCC GAG ATG GCG GTG GGC ACG CCC GCG ATG **6291** 3'	Ajoute un site SduI
L392W	5' **6259** AAG GCC GAG ATG TGG GTG GGC ACG CCC GCG ATG **6291** 3'	Ajoute un site SduI
P396A	5' **6265** GAG ATG CTG GTG GGC ACG GCC GCG ATG TGG **6297** 3'	Ajoute un site BseSI
Extrémité C-terminale		
W517A	5' 6634 G AAC CTG GAG ATC GCG CTG CCC AGC 6661 3'	Retire un site BglII
W517F	5' 6634 G AAC CTG GAG ATA TTC CTG CCC AGC 6661 3'	Retire un site BglII
P519A	5' 6637 G AAC CTG GAG ATA TGG CTG GCC AGC TGG C 6662 3'	Retire un site BglII
P519D	5' 6637 G AAC CTG GAG ATA TGG CTG GAC AGC TGG CG 6665 3'	Retire un site BglII

A. Cultures bactériennes et génération de souches mutées

Mutagenèse dirigée. Les mutations de la séquence du gène *pmtub* ont été obtenues par mutagenèse dirigée, selon la méthode de « quickchange » de Stratagen. Des amorces complémentaires sens et antisens ont été conçues afin d'introduire la mutation souhaitée, à côté d'une mutation silencieuse qui ajoute ou supprime un site de restriction (voir la séquence des amorces, **Tableau 5**). Des mutations ont été introduites dans le gène *pmtub* codé par le plasmide pMW158 (voir caractéristiques des principaux plasmides, **Tableau 10**), par PCR sur plasmide entier en utilisant la polymérase Phusion (ThermoFisher) avec le tampon commercial high-GC et en présence de 3% de DMSO. Les cycles PCR étaient les suivants : une première étape de 3 minutes à 98°C, suivie de 30 cycles composés d'une étape de dénaturation de 20 secondes à 98°C, d'une étape d'hybridation de 30 secondes à 65°C et d'une étape d'allongement de 5 minutes à 72°C, et une étape finale d'allongement de 10 minutes à 72°C. Des modifications ont été apportées à ce protocole afin de mieux s'adapter aux caractéristiques de chaque amorce (longueur et température d'hybridation) : la température d'hybridation a varié entre 55°C et 72°C et la concentration en DMSO a été testée de 0% à 10%. Les produits de PCR ont été incubés avec l'enzyme DpnI (suivant les recommandations du fabricant, ThermoFisher) afin de dégrader le plasmide parental puis utilisés pour transformer la souche *E. coli* électro-compétente TOP10 (suivant les recommandations du fabricant, **Tableau 6**). Les plasmides pWM158 et pWM158 mutés ont été amplifiés et purifiés à partir de cultures d'*E.coli* pendant la nuit (sur milieu LB, en présence d'un antibiotique approprié pour permettre la sélection, **Tableau 7**) en utilisant le kit Qiagen QIaprep Spin Miniprep. Les concentrations et la pureté du plasmide ont été contrôlées par nanodrop. Leurs profils de restriction ont été établis afin de confirmer leur identité et de vérifier éventuellement l'acquisition des mutations souhaitées. Cette vérification qualitative a été complétée, pour trois clones de chaque construction, par le séquençage global de l'ensemble de la séquence du gène *pmtub* (par Eurofins).

Transformation de *E. coli* et production d'ADN. Pour obtenir un taux de transformation de *M. smegmatis* optimal, les plasmides à transformer ont été produit pour obtenir une concentration élevée. Les plasmides sont transformés dans *E. coli* Dh5α (**Tableau 6**). Les bactéries *E. coli* sont préalablement rendues compétentes selon le protocole suivant : culture dans milieu LB jusqu'à atteindre une DO_{600nm} de 0,5, centrifugation (1000 xg pendant

Tableau 6. Souches bactériennes utilisées dans cette étude et leurs principales caractéristiques.

Nom	Génotype	Source
PMM186 (appelée ΔPMT)	Souche mc²155 de *M.smegmatis* avec son gène de PMT (*M.SMEG*_5447) interrompu	Liu et al, 2013[52]
Dh5α	**Souche de clonage** *Φ80 Δ(lacZ)M15 fhuA2 Δ(argF-lacZ)U169 phoA glnV44 gyrA96 recA1 relA1 endA1 thi-1hsdR17*	Taylor *et al*, 1993 [209]
Stellar	**Souche de clonage chimiocompétente** *F-, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ(lacZYA-argF) U169, Δ(mrr-hsdRMS-mcrBC), ΔmcrA, λ-*	Takara bio
One Shot Top10	**Souche de clonage electrocompétente** *F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu) 7697 galU galK rpsL (StrR) endA1 nupG*	Invitrogen
BL21-DE3	**Souche de clonage electrocompétente** *F– ompT HSDS$_B$ (r_B–, m_B–) gal dcm (DE3)*	SigmaAldrich

Tableau 7. Antibiotiques utilisés pour les cultures bactériennes.

Nom	Dilution	Source
Kanamycine	*E. coli* : 50µg/mL *M. smegmatis* : 25µg/mL	Fisher Scientific
Hygromycine	*E. coli* : 100µg/mL *M. smegmatis* : 50µg/mL	Invivogen
Ampicilline	*E. coli* : 100µg/mL	Fisher Scientific
Streptomycine	*E. coli* : 100µg/mL *M. smegmatis* : 50µg/mL	Fisher Scientific

10 minutes) et récupération du culot, resuspendre dans 1/10 du volume dans du tampon de stockage (10%PEG (MW 3350), 5% DMSO, 10mM $MgCl_2$, 10mM $MgSO_4$ dans du LB), stockage à -80°C des aliquots de bactéries. 100µl de bactéries compétentes sont mis en présence de 1µl de plasmide à produire, 79µl d'eau stérile et de 20µl de solution « KCM » (0,5M KCl, 0,15M $CaCl_2$ et 0,25M $MgCl_2$) et un choc thermique est réalisé (20 minutes sur glace suivit de 30secondes à 37°C). Une expression phénotypique est réalisée à 37°C pendant 1heure et les bactéries sont étalées sur milieu LA avec les antibiotiques appropriés (**Tableau 7**). Après une incubation d'une nuit à 37°C, des colonies isolées sont repiquées dans du milieu LB avec les antibiotiques appropriés (**Tableau 7**) et incubées à 37°C sur la nuit. Le lendemain, les plasmides sont récupérés en utilisant le kit Qiagen QIaprep Spin Miniprep. Les concentrations et la pureté du plasmide ont été contrôlées par nanodrop.

Transformation et culture de *Mycobacterium smegmatis*. Les plasmides et souches utilisés dans cette étude sont présentés dans les **Tableau 8**, **Tableau 10** et **Tableau 6**. Des souches de *M. smegmatis* compétentes ont été préparées en suivant les étapes suivantes : les bactéries à transformer ont été cultivées pendant une semaine dans 10 ml de milieu LB avec du Tween 80 à 0,05 % (appelé milieu LBT) jusqu'à une DO_{600nm} de 1. 500 µL de la préculture ont été repris pour ensemencer 50 ml de LBT jusqu'à une DO_{600nm} de 0,2-0,8 puis centrifugées (8 minutes, 4°C, 500 xg). Les culots cellulaires ont été lavés 3 fois avec des volumes décroissants de tampon de lavage froid (Tween 80 0,05 %, glycérol 10 %) : 50 ml, 15 ml et 1 ml avec centrifugation (8 minutes, 4°C, 500 xg) entre chaque étape de lavage. Les cellules compétentes lavées ont été divisées en aliquots de 200 µl. 1 µL de plasmide (à 200 ng/µL) a été ajouté dans chaque aliquot de cellules compétentes de *M. smegmatis* et l'électroporation a été réalisée à 2500 V, 25 µf, 200 Ω. Immédiatement après la transformation, 800 µl de milieu LBT ont été ajoutés aux bactéries et celles-ci ont été incubées pendant 3 heures à 37°C sous agitation (180 rpm). Les bactéries transformées ont ensuite été étalées sur du milieu LA contenant les antibiotiques appropriés (**Tableau 7**) et incubées à 37°C pendant 5 jours. Les colonies isolées de la boite de LA ont été cultivées dans 10 mL de milieu LBT contenant les antibiotiques appropriés (**Tableau 7**) à 37°C pendant 9 jours (jusqu'à DO600nm 1) pour être soit transformées à nouveau (afin d'introduire le plasmide codant FasC ou pSMjF) soit analysées en termes de phénotype. Pour ce dernier cas, les cellules ont été centrifugées (10 minutes, 4°C, 500 xg) et les surnageants et culots stockés séparément à -20°C.

Tableau 8. Plasmides utilisés dans cette étude et leurs principales caractéristiques.

Nom	Caractéristiques principales	Résistance	Source
pWM158	Plasmide intégratif portant le gène rv1002c sous contrôle du promoteur pBlaF*	Kanamycine	Liu et al, 2013[52]
pMV361	Plasmide intégratif vide utilisé comme contrôle	Kanamycine	Liu et al, 2013[52]
pWM218	Plasmide réplicatif portant le gène MSMEG_5196 (FasC) sous contrôle du promoteur pBlaF*	Hygromycine	Liu et al, 2013[52]
pMIPS4PhoA	Plasmide réplicatif portant le gène de mmpSa en fusion avec PhoA	Kanamycine	Dr Chalut, IPBS
pMIPKatGPhoA	Plasmide réplicatif portant le gène de KatG en fusion avec PhoA	Kanamycine	Dr Chalut, IPBS
pGMCS-GFP-PMbcA-RFP	Plasmide intégratif portant les gènes de mCherry et de GFP	Streptomycine	Dr Gutierrez, IPBS

Tableau 9. Anticorps utilisés dans cette étude.

Nom	Description	Dilution	Source
Anti-Strep	Mouse	1/500	Sigma, SAB2702216
Anti-HA	Mouse	1/3000	Invivogen, ab-hatag
Anti-Flag	Mouse, HRP	1/500	Sigma, A8592
Anti-Mouse	Goat, HRP	1/2000	Sigma, A4416
Anti-His	Mouse, HRP	1/5000	Invitrogen, 3D5
Anti-GST	Rabbit	1/1000	Invitrogen, CAB4169
Anti-rabbit	Goat, HRP	1/5000	Biorad, 172-1019

« Chemical rescue » *in vivo* pour la restauration de l'activité de PMTub mutée sur la triade catalytique. 9mL de cultures de 1 semaine des souches de *M. smegmatis* recombinantes produisant les enzymes portants les mutations D74A, D176A et R441A, ainsi que les deux souches contrôles PMTub native et No PMT (Table 2), ont été centrifugées 10 minutes à 4000 rpm et 4°C. Les culots ont été repris à l'aide de billes de verre stériles (de 4 mm) dans 20mL de LBT, puis centrifugés à nouveau pendant 10 minutes à 800 rpm pour retirer les agrégats cellulaires et ne garder que des surnageants de cellules supposées dissociées, dont les A_{600nm} ont été mesurées. Ces surnageants de cellules dissociées ont alors été employés pour démarrer des micro-cultures en plaques 96 puits. Pour cela, les puits contenant du milieu LBTKH (LB, Tween 80 0.05%, Kanamycine 25µg/mL, Hygromycine 50µg/mL) ont été inoculés à une A_{600nm} de 0,01. Les plaques contenant les micro-cultures ont été incubées à 37°C avec re-suspension des bactéries par agitation douce quotidienne. Les A_{600nm} des micro-cultures ont été mesurées au bout de 48h, et les composés chimiques ont été ajoutés dans le milieu de culture selon le plan suivant : les souches exprimant PMTub-R441A, PMTub native et No PMT ont été traitées avec 2,5 mM ou 5mM d'Imidazole ou de Guanidine Hydrochloride ; les enzymes PMTub-D74A, D176A, native et No PMT ont été traitées avec 2,5mM ou 5mM d'Acétate de Sodium ou d'Azide de Sodium. Chaque culture a également été cultivée sans traitement, en contrôle.

B. Test d'activité phénotypique de PMTub

1. Étude de la mannosylation de la FasC

Étude de la mannosylation de la FasC par ELISA direct et lectinique. Les surnageants ont d'abord été analysés par ELISA direct afin de doser la concentration en $FasC^{6His}$. 100 µl de surnageant de culture dilués deux fois dans du tampon (0,53 % Na2C03, 0,42 % NaHC03, pH 9,6) ont été incubés en triplicat, pendant une nuit à température ambiante, dans des plaques 96 puits. Après 3 lavages avec 200 µL de PBS, la plaque a été incubée 1 heure à température ambiante avec 200 µL de PBS contenant 0,05 % de Tween20 et 5 % de lait. Après 3 lavages avec 200 µL de PBS contenant 0,05 % de Tween20, la plaque a été incubée 2 heures à température ambiante avec 50 µL d'anticorps anti-His (Anti-His (C-term)-HRP anticorps, Invitrogen) dilué 1000x dans PBS contenant 0,05 % de Tween20 et 0,5 % de lait. Après quatre lavages avec 200 µl de PBS, la révélation a été réalisée avec 50 µl de solution de TMB

Tableau 10. Plasmides construits dans cette étude et leurs principales caractéristiques.

Nom	Caractéristiques principales	Résistance
pGex-EL4	Plasmide commercial pGex-6P dans lequel la séquence codant le peptide EL4 a été incluse	Ampicilline
pWM158-minipBlaF*	Plasmide intégratif portant le gène rv1002c sous contrôle du promoteur mini-pBlaF*	Kanamycine
Mutants de PMTub		
pWM158-D74N	Plasmide pWM158 portant la mutation D74N sur le gène de pmtub	Kanamycine
pWM158-D74E	Plasmide pWM158 portant la mutation D74E sur le gène de pmtub	Kanamycine
pWM158-D74A	Plasmide pWM158 portant la mutation D74A sur le gène de pmtub	Kanamycine
pWM158-DE74-75ED	Plasmide pWM158 portant la mutation DE74-75ED sur le gène de pmtub	Kanamycine
pWM158-D176N	Plasmide pWM158 portant la mutation D176N sur le gène de pmtub	Kanamycine
pWM158-D176E	Plasmide pWM158 portant la mutation D176E sur le gène de pmtub	Kanamycine
pWM158-D176A	Plasmide pWM158 portant la mutation D176A sur le gène de pmtub	Kanamycine
pWM158-R441L	Plasmide pWM158 portant la mutation R441L sur le gène de pmtub	Kanamycine
pWM158-R441K	Plasmide pWM158 portant la mutation R441K sur le gène de pmtub	Kanamycine
pWM158-R441A	Plasmide pWM158 portant la mutation R441A sur le gène de pmtub	Kanamycine
pWM158-Y444A	Plasmide pWM158 portant la mutation Y444A sur le gène de pmtub	Kanamycine
pWM158-Y444F	Plasmide pWM158 portant la mutation Y444F sur le gène de pmtub	Kanamycine
pWM158-W355A	Plasmide pWM158 portant la mutation W355A sur le gène de pmtub	Kanamycine
pWM158-W355F	Plasmide pWM158 portant la mutation W355F sur le gène de pmtub	Kanamycine
pWM158-P359A	Plasmide pWM158 portant la mutation P359A sur le gène de pmtub	Kanamycine
pWM158-W362A	Plasmide pWM158 portant la mutation W362A sur le gène de pmtub	Kanamycine
pWM158-W362F	Plasmide pWM158 portant la mutation W362F sur le gène de pmtub	Kanamycine
pWM158-M364A	Plasmide pWM158 portant la mutation M364A sur le gène de pmtub	Kanamycine
pWM158-M364V	Plasmide pWM158 portant la mutation M364V sur le gène de pmtub	Kanamycine
pWM158-S365A	Plasmide pWM158 portant la mutation S365A sur le gène de pmtub	Kanamycine
pWM158-S365T	Plasmide pWM158 portant la mutation S365T sur le gène de pmtub	Kanamycine
pWM158-L366A	Plasmide pWM158 portant la mutation L366A sur le gène de pmtub	Kanamycine
pWM158-Y371A	Plasmide pWM158 portant la mutation Y371A sur le gène de pmtub	Kanamycine
pWM158-Y371F	Plasmide pWM158 portant la mutation Y371F sur le gène de pmtub	Kanamycine
pWM158-I373Q	Plasmide pWM158 portant la mutation I373Q sur le gène de pmtub	Kanamycine
pWM158-I373L	Plasmide pWM158 portant la mutation I373L sur le gène de pmtub	Kanamycine
pWM158-C381A	Plasmide pWM158 portant la mutation C381A sur le gène de pmtub	Kanamycine
pWM158-C381S	Plasmide pWM158 portant la mutation C381S sur le gène de pmtub	Kanamycine
pWM158-C386A	Plasmide pWM158 portant la mutation C386A sur le gène de pmtub	Kanamycine
pWM158-C386S	Plasmide pWM158 portant la mutation C386S sur le gène de pmtub	Kanamycine
pWM158-L392A	Plasmide pWM158 portant la mutation L392A sur le gène de pmtub	Kanamycine
pWM158-L392W	Plasmide pWM158 portant la mutation L392W sur le gène de pmtub	Kanamycine
pWM158-P396A	Plasmide pWM158 portant la mutation P396A sur le gène de pmtub	Kanamycine
pWM158-W517A	Plasmide pWM158 portant la mutation W517A sur le gène de pmtub	Kanamycine
pWM158-W517F	Plasmide pWM158 portant la mutation W517F sur le gène de pmtub	Kanamycine
pWM158-P519A	Plasmide pWM158 portant la mutation P519A sur le gène de pmtub	Kanamycine
pWM158-P519D	Plasmide pWM158 portant la mutation P519D sur le gène de pmtub	Kanamycine
Insertion du codon ambre dans PMTub		
pWM158-F117-mCherry	Plasmide pWM158-C-mCherry portant un codon ambre à la position 117 sur le gène de pmtub	Kanamycine
pWM158-F117-PhoA	Plasmide pWM158-PhoA portant un codon ambre à la position 117 sur le gène de pmtub	Kanamycine
pWM158-V378-mCherry	Plasmide pWM158-C-mCherry portant un codon ambre à la position 378 sur le gène de pmtub	Kanamycine
pWM158-V378-PhoA	Plasmide pWM158-PhoA portant un codon ambre à la position 378 sur le gène de pmtub	Kanamycine

Nom	Caractéristiques principales	Résistance
PMTub étiquetée		
pWM158-Strep	Plasmide pWM158 avec le gène de pmtub portant une étiquette strep en C-terminal	Kanamycine
pWM158-TwinStrep	Plasmide pWM158 avec le gène de pmtub portant une étiquette twin strep en C-terminal	Kanamycine
pWM158-Flag	Plasmide pWM158 avec le gène de pmtub portant une étiquette flag en C-terminal	Kanamycine
pWM158-HA	Plasmide pWM158 avec le gène de pmtub portant une étiquette HA en C-terminal	Kanamycine
pWM158-3xHA	Plasmide pWM158 avec le gène de pmtub portant une étiquette triple HA en C-terminal	Kanamycine
pWM158-His	Plasmide pWM158 avec le gène de pmtub portant une étiquette His en N-terminal	Kanamycine
PMTub en fusion		
pWM158-N-mCherry	Plasmide pWM158 avec le gène de pmtub en fusion avec mCherry à son extrémité N-terminale	Kanamycine
pWM158-C-mCherry	Plasmide pWM158 avec le gène de pmtub en fusion avec mCherry à son extrémité C-terminale	Kanamycine
pWM158-sfTq2ox	Plasmide pWM158 avec le gène de pmtub en fusion avec sfTq2ox à son extrémité C-terminale	Kanamycine
pWM158-mNG	Plasmide pWM158 avec le gène de pmtub en fusion avec mNeonGreen à son extrémité C-terminale	Kanamycine
pWM158-PhoA	Plasmide pWM158 avec le gène de pmtub en fusion avec PhoA à son extrémité C-terminale	Kanamycine
pWM158-D74N-PhoA	Plasmide pWM158-PhoA portant la mutation D74N sur le gène de pmtub	Kanamycine
pWM158-D176N-PhoA	Plasmide pWM158-PhoA portant la mutation D176N sur le gène de pmtub	Kanamycine
pWM158-R441L-PhoA	Plasmide pWM158-PhoA portant la mutation R441L sur le gène de pmtub	Kanamycine
pWM158-Y444A-PhoA	Plasmide pWM158-PhoA portant la mutation Y444A sur le gène de pmtub	Kanamycine
pWM158-P359A-PhoA	Plasmide pWM158-PhoA portant la mutation P359A sur le gène de pmtub	Kanamycine
pWM158-W355A-PhoA	Plasmide pWM158-PhoA portant la mutation W355A sur le gène de pmtub	Kanamycine
pWM158-W355F-PhoA	Plasmide pWM158-PhoA portant la mutation W355F sur le gène de pmtub	Kanamycine
pWM158-P359A-PhoA	Plasmide pWM158-PhoA portant la mutation P359A sur le gène de pmtub	Kanamycine
pWM158-W362F-PhoA	Plasmide pWM158-PhoA portant la mutation W362F sur le gène de pmtub	Kanamycine
pWM158-Y371A-PhoA	Plasmide pWM158-PhoA portant la mutation Y371A sur le gène de pmtub	Kanamycine
pWM158-Y371F-PhoA	Plasmide pWM158-PhoA portant la mutation Y371F sur le gène de pmtub	Kanamycine
pWM158-L392A-PhoA	Plasmide pWM158-PhoA portant la mutation L392A sur le gène de pmtub	Kanamycine
pWM158-P519A-PhoA	Plasmide pWM158-PhoA portant la mutation P519A sur le gène de pmtub	Kanamycine
Contrôles de l'expression des protéines de fusion		
pmmpS4-PhoA	Plasmide pWM158 dans lequel le gène de PMTub a été remplacé par celui de mmpS4 en fusion avec PhoA a son extrémité C-terminale	Kanamycine
pKatG-PhoA	Plasmide pWM158 dans lequel le gène de PMTub a été remplacé par celui de KatG en fusion avec PhoA a son extrémité C-terminale	Kanamycine
pmmpS4-mNG	Plasmide pWM158 dans lequel le gène de PMTub a été remplacé par celui de mmpS4 en fusion avec mNeongreen a son extrémité C-terminale	Kanamycine
pSMjF		
pSMjF	Plasmide réplicatif portant les gène du système orthogonal et de la FasC sous contrôle du promoteur mini-pBlaF*	Hygromycine
pSMjF long	Plasmide réplicatif portant les gène du système orthogonal et de la FasC sous contrôle du promoteur pBlaF*	Hygromycine
pSMT3-FasC	Plasmide pSMjF duquel les gènes du système orthogonal ont été retiré	Hygromycine

(eBioscience), puis 25 µL de solution d'arrêt (1 mM d'acide sulfurique) ont été ajoutés avant d'atteindre la saturation (10 à 20 minutes). La gamme standard de FasC purifiée a été réalisée pour chaque ELISA direct avec les concentrations suivantes (en ng/ml) : 50, 100, 200, 300, 500, 700, 900 et 2000. Dans un deuxième temps, des surnageants avec des concentrations normalisées de $FasC^{6His}$ (déduites de l'ELISA direct) ont été analysés pour la mannosylation $FasC^{6His}$ par ELISA lectinique. Pour ce deuxième test, des plaques 96 puits avec des puits revêtus d'anticorps anti-His (Genscript) ont été utilisés. 100 µl de surnageants de culture (dilués deux fois dans du PBS) ont été ajoutés par puits en triple. La plaque a été incubée durant 3 heures à température ambiante puis lavée 4 fois avec 260 pl de PBS contenant 2 % de Tween20. 100 µl de solution de Concanavaline A-HRP (0,5 µg/mL de ConA-HRP, 1 mM de $CaCl_2$, 1 mM de $MgCl_2$, 1 mM de $MnCl_2$, 0,05 % de Tween20, dans du PBS) ont été ajoutés dans chaque puits et la plaque a été incubée à température ambiante pendant 2 heures. Après quatre lavages avec 260 µl de PBS, la révélation a été réalisée avec 100 µl de solution de TMB (eBioscience), après quoi 50µL de solution d'arrêt (acide sulfurique 1 mM) ont été ajoutés avant d'atteindre la saturation (10 à 20 minutes). Les absorbances des deux ELISA ont été mesurées à 450nm en utilisant le lecteur de microplaques CLARIOstar BMG LABTECH.

Purification de la $FasC^{6His}$. 3mL de surnageant de culture de 9 jours de *M. smegmatis*, dilués trois fois dans du tampon de liaison (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, pH 8) ont été chargés sur une colonne de centrifugation Ni-NTA (Qiagen). Après 5 minutes de centrifugation à 270 xg, le flowthrough a été jeté. Ensuite, la colonne a été lavée deux fois par centrifugation (2 minutes, 890 xg) avec 600 µl de tampon de lavage (50 mM NaH2P04, 300 mM NaCl, 20 mM imidazole, pH 8). Les protéines ont été éluées deux fois avec 100 µl de tampon d'élution (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, pH 8) par centrifugation (2 minutes, 890 xg). Pour chaque échantillon, des aliquots de chaque étape de purification ont été évaluées par SDS-PAGE. Le premier éluat (contenant la plupart du FasC purifié) a ensuite été évalué en utilisant le test de protéine BCA (Thermo Fisher, suivant les instructions du fabricant) pour mesurer la concentration de FasC.

Étude de la mannosylation de $FasC^{6His}$ par spectrométrie de masse (réalisé par la plateforme protéomique de l'IPBS). Des échantillons de FasC purifiés ont été dilués dans du tampon de charge (2% ACN, 0,1% TFA) pour atteindre une concentration finale de 0,4 M. Les analyses NanoLC-MS ont été réalisées sur un système nanoRS UHPLC (Dionex) couplé à un

spectromètre de masse LTQ-Orbitrap Velos (Thermo Fisher Scientific). 5 µL d'échantillon ont été chargés sur une pré-colonne C4 en phase inverse (300 m i.d. x 5 mm ; Thermo Fisher Scientific) à 20 µL/min dans 2 % d'acétonitrile et 0,05 % de TFA. Après 5 minutes de dessalage, la précolonne a été couplée avec une nanocolonne analytique C4 (75 µm i.d. x 15 cm; Reprosil-C4) équilibrée dans 95% de solvant A (0,2% d'acide formique) et 5% de solvant B (0,2 % d'acide formique dans l'acétonitrile). Les protéines ont été éluées en utilisant un gradient linéaire de 5 % de B à 100 % en 38 minutes à un débit de 300 nL/min. Les scans MS ont été acquis dans la plage 800-2000 m/z avec une résolution fixée à une valeur de 60 000, et les spectres ont été déconvolués à l'aide du logiciel Xtract (Thermo Fisher Scientific).

Etude de la FasC^{6His} par profil électrophorétique. Les échantillons sont déposés sur gel SDS-PAGE. La séparation est réalisée par un gel de polyacrylamide 13% précédé d'un gel de concentration de polyacrylamide 5%. La migration est effectuée dans un tampon de migration (Tris Glycine et SDS). La coloration des gels est réalisée par une solution de bleu colloïdale (Instant Blue) puis rincée à l'eau. Pour les analyses par western-blot, les protéines sont transférées sur une membrane de nitrocellulose pour analyse par western blot. Les membranes sont incubées en présence de l'anticorps anti-His couplé HRP pendant 2h dans une solution de PBS avec 1% de BSA et 0,05% de Tween20. La membrane est ensuite lavée 3 fois avec du PBS puis la révélation est réalisée à l'aide du réactif Amersham ECL Prime en suivant les instructions du constructeur.

2. Contrôle de l'expression de PMTub à la membrane

a) Approche par western blot

Construction des souches exprimant PMTub étiquetée. Les séquences des étiquettes ajoutées aux extrémités N- et C-terminales de PMTub ont été apportées à l'aide d'amorces à queue flottante (**Tableau 11**). Les PCR sont réalisées avec la Platinum SuperFi II DNA Polymerase (invitrogen). Les amorces sont phosphorylées à leur extrémité 5' permettant de circulariser le produit d'amplification par la T4 DNA ligase (Fisher scientific). Le produit de ligation est transformé chez *E. coli* pour production comme décrit plus haut (voir p.193 et est vérifié par profil de restrictions puis par séquençage.

Tableau 11. Amorces utilisées pour la construction PMTub étiquetée.

Nom	Séquence	Site de restriction
Rev_TwinStrep	5'- [PHO] GCTGCCACCGCCGCTGCCACCGCCTTTTTCGAACTGCGGGTG -3'	FokI
For_TwinStrep	5'- [PHO]GGCGGTAGCTCGGCATGGAGCCATCCGCAGTTCGAGAAGTGAAGCTTATCGATGTCG -3'	FokI
Rev_HA	5'- [PHO] GCCGCCGAACCCCAGCTGGGCAGCCAGATC -3'	NdeI
For_HA	5'- [PHO]CTACCCATATGATGTGCCGGACTACGCGTGAAGCTTATCGATGTCGAC -3'	NdeI
Rev_3HA	5'- [PHO]TGCGTAATCCGGGACGTCGTACGGGTACGAGCCCGCGTAGTCCGGCACATAGTATGGGT-3'	NdeI
For_3HA	5'- [PHO]GGGTCCTATCCGTACGACGTACCTGATTATGCATGAAGCTTATCGATGTCGAC-3'	NdeI
Rev_Flag	5'- [PHO] CGAACCGCCCCAGCTGGGCAGCCAGATC -3'	FokI
For_Flag	5'- [PHO] GACTACAAGGATGACGACGACAAGTGAAGCTTATCGATGTCGAC -3'	FokI
Rev_His	5'- [PHO] GTGATGCATATGTACTTCTCCTTCTAATCCATTGGACCC -3'	KpnI
For_His	5'- [PHO] CATCACCATCACCATCACCATCACGTACCCGTCGTCAGCCCC GGC -3'	KpnI

Préparation d'extraits enrichis de la paroi cellulaire, de la membrane et du compartiment cytosolique de *M. smegmatis*. À cet effet, des extraits ont été préparés en modifiant le protocole décrit précédemment[75]. 200 mL de cultures de 2 jours en milieu LB Tween80 0,05 % (contenant 25 µg/mL de Kanamycine), ont été centrifugés, à 1500 xg pendant 10 minutes. Les cellules ont été remises en suspension dans 800 µL de tampon A (contenant 50 mM de MOPS, pH 7,9 ajusté par KOH, 5 mM de β-mercaptoéthanol, 10 mM de MgCl2, un cocktail d'inhibiteur de protéase sans EDTA (Sigma aldrich) et 1 µL de benzonase (5 KU de Sigma Aldrich) pour 1 g de culot cellulaire. Ensuite, une lyse cellulaire par batteur à billes a été réalisée, avec 3 cycles d'une durée de 60 s chacun et suivie de 120 s sur glace entre les agitations, à 35 % de l'amplitude d'agitation maximale. Les lysats cellulaires ont ensuite été centrifugés à 3000 xg pendant 10 minutes à 4°C pour éliminer les cellules non lysées. Les lysats ont été incubés avec 200 µL de solution de 2,2 mM de Triton-X100 pendant 2 h sous agitation douce, puis centrifugés à 23 000 xg pendant 20 minutes à 4°C afin d'obtenir la fraction « paroi cellulaire » dans les culots. Ceux-ci ont été remis en suspension dans 250 µL de tampon A additionné de de 2,2 mM de Triton-X100. Pour préparer la fraction membranaire et le cytosol, le surnageant de l'étape de centrifugation précédente a été soumis à une ultracentrifugation à 100000 xg pendant 1h à 4°C. Le culot gélatineux jaune-brun résultant a été remis en suspension dans 83,4 µl de tampon A complété avec 2,2 mM de Triton-X100. Ce culot repris est appelé fraction membranaire et le surnageant clair cytosol.

Western Blot sur lysats cellulaires. Les bactéries ont été cultivées et lysées pour obtenir les fractions « paroi mycobactérienne », « membrane plasmique » et « cytoplasme ». Les différentes fractions ont été déposées sur SDS-PAGE et, après migration, transférées sur membrane de nitrocellulose. Les membranes ont été incubées en présence de l'anticorps primaire adapté dilué dans du PBS avec 1% de BSA et 0,05% de Tween20 pendant 3 heures. Le blocage a été ensuite effectué pendant 3 heures avec du PBS avec 5% de BSA et 0,1% de Tween20. Une dernière étape d'incubation a été réalisé avec un anticorps secondaire adapté dilué dans du PBS avec 1% de BSA et 0,1% de Tween20 pendant 1h. La membrane a été révélée à l'aide du réactif Amersham ECL Prime en suivant les instructions du constructeur.

Détection de PMTub marquée par spectrométrie de masse (réalisée en collaboration avec la plateforme protéomique de l'IPBS). Après migration SDS-PAGE, le gel est coloré par bleu de coomassie. Des bandes de gel correspondant aux bandes d'intérêts sont découpées et déshydratées dans de l'acétonitrile (ACN) 100%. Elles sont ensuite réhydratées dans du

Tableau 12. Amorces utilisées pour la construction de protéines de fusion. Pour les protéines fluorescentes mNG et SfTq2ox, les amorces utilisées pour la fusion avec mCherry sont aussi utilisées.

Protéine de fusion	Nom	Séquence
PMTub-Cter-mCherry	Rev_mC_C_plasmide	CCT TGC TCA CGG ATC CGC TGC TGC CCC AGC TGG GCA GCC AGA TC
	For_mC_C_plasmide	TGA AGC TTA TCG ATG TCG ACG TAG
	Rev_mC_C_insert	GTC GAC ATC GAT AAG CTT CAC TTG TAC AGC TCG TCC ATG
	For_mC_C_insert	GCC CAG CTG GGG CAG CAG CGG ATC CGT GAG CAA GGG CGA GGA G
mCherry-Nter-PMTub	Rev_mC_N_plasmide	CAT ATG TAC TTC TCC TTC TAA TCC ATT GG
	For_mC_N_plasmide	GCT GTA CAA GGG CAG CAG CGG ATC CGT ACC CGT CGT CAG CCC C
	Rev_mC_N_insert	CGA CGG GTA CGG ATC CGC TGC TGC CCT TGT ACA GCT CGT CCA TG
	For_mC_N_insert	TAG AAG GAG AAG TAC ATA TGG TGA GCA AGG GCG AGG AG
PMTub-PhoA	Rev_PhoA_plasmide	CCA GCT GGG CAG CCA GAT C
	For_PhoA_plasmide	AAG CCG CTC TGG GGC TGA AAT GAA GCT TAT CGA TGT CGA CGT AG
	Rev_PhoA_insert	TTT CAG CCC CAG AGC GGC
	For_PhoA_insert	AGA TCT GGC TGC CCA GCT GGG ACT CTT ATA CAC AAG TAG CGT CC

bicarbonate d'ammonium 100mM et incubées 3h à 37°C. Les bandes sont ensuite séchées au speed-vac. Les protéines contenues dans les bandes de gel sont réduites par ajout d'une solution de DTT 10mM dans du bicarbonate d'ammonium 100 mM en incubant 30minutes à 56°C. Après incubation, les bandes de gel sont déshydratées dans de l'acétonitrile 100% puis réhydratées dans du bicarbonate d'ammonium 100 mM. L'alkylation des protéines est réalisée par ajout d'iodoacétamide 55mM dans du bicarbonate d'ammonium 100mM. Après 30 minutes d'incubation à température ambiante, les bandes de gel sont déshydratées dans de l'acétonitrile 100% puis réhydratées dans du bicarbonate d'ammonium 100 mM et séchées au speed-vac. La digestion trypsique est réalisée sur la nuit à 37°C par la tryspine de Promega selon les recommandations commerciales. Les peptides trypsiques sont extraits par incubations et sonications successives : d'abord dans du bicarbonate d'ammonium 50 mM et ensuite dans un mélange acide formique/acétonitrile. Les extraits obtenus sont évaporés au speed-vac. Les échantillons sont repris dans du tampon de charge (5% ACN, 0,1% TFA) pour atteindre une concentration finale de 0,4 M. Les analyses NanoLC-MS ont été réalisées sur un système nanoRS UHPLC (Dionex) couplé à un spectromètre de masse LTQ-Orbitrap Velos (Thermo Fisher Scientific) tel que décrit précédemment (voir p.200).

b) Approches phénotypiques

Construction des souches pour l'expression des protéines de fusion PMTub. Les protéines de fusion ont été construites à l'aide de la technique du « Gibson assembly »[206]. Les différents fragments d'ADN linéaire ont été amplifiés par PCR en utilisant la Platinum SuperFi II DNA Polymerase (invitrogen). Les amorces utilisées sont recensées **Tableau 12**. Les gène s codant pour les protéines fluorescentes et la phosphatase alcaline ont été récupérées de plasmides décrits **Tableau 8**. Les gènes de PMTub et des protéines contrôles mmpS4 et KatG ont été récupérés des plasmides décrits **Tableau 8**. Les différents amplicons linéaires ont été assemblés à l'aide du Kit « NEBuilder HiFi DNA Assembly Cloning » (New England Biolabs) selon le protocole commercial et la solution de réaction a été directement transformé dans *E. coli* stellar (**Tableau 6**). Les plasmides ainsi construits ont été récupérés et produits tels que décrits précédemment (voir p.193). Les constructions ont été vérifiées dans un premier temps par profil de restriction et ensuite par séquençage du gène de la protéine de fusion.

Mesure de l'activité de la phosphatase alcaline en milieu liquide. Les activités PhoA ont été mesurées sur des souches de *M. smegmatis* cultivées en milieu liquide (LB Tween80 0,05 % avec antibiotique Kanamycine 25 µg/mL) à 37°C, 180 rpm, pendant 5 jours. De l'iodoacétamide a été ajouté à une concentration finale de 10 mM, 20 minutes avant la récolte et maintenu tout au long de l'expérience pour empêcher le repliement spontané et l'activation de PhoA cytoplasmique. Les cellules ont été récoltées par centrifugation (5000 tr/min, 10 minutes, à 4°C) et lavées deux fois dans du tampon de lavage (1M Tris pH8, Tween 80 0,05 %, iodoacétamide 10 mM, 10 mM MgCl2). Les concentrations cellulaires dans le tampon de lavage ont été normalisées à 0,4 g/mL de culots humides. 200 µl de suspensions cellulaires normalisées ont été mélangés avec 1 ml de tampon de réaction (10 mM de phosphate para-nitrophénylphosphate dans du tampon de lavage) et incubés à 37°C dans l'obscurité, sous agitation douce. Un suivi cinétique de deux heures a ensuite été réalisé en ajoutant 50 µl de solution d'arrêt (1M K_2HPO_4) aux échantillons de suspension cellulaire. Les cellules ont été éliminées par centrifugation (13 000 tr/min, 10 minutes, à 4°C) et la DO_{420nm} des surnageants restants a été mesurée pour surveiller l'hydrolyse du phosphate para-nitrophénylphosphate par PhoA.

Détection de l'activité de la phosphatase alcaline sur milieu de culture solide. La présence d'une activité phosphatase alcaline peut être détectée par une lecture phénotypique sur colonie. Les souches à tester ont été étalées sur boite de LA supplémentée par 60 µg/ml de BCIP (Fisher Scientific) et en antibiotiques adaptés (**Tableau 7**). Après 7 jours de cultures, les boites de Petri ont été analysées pour la détection d'une coloration bleue des colonies. Dans le cas du test d'insertion d'un acide aminé non naturel, les bactéries ont été étalées sur deux boites de Petri avec une des deux boites supplémentées avec pBpa 2mM (Fisher Scientific).

Coloration au PKH67 des bactéries. Les membranes mycobactériennes ont été marquées avec un colorant fluorescent vert (PKH67, Sigma-Aldrich). La culture de mycobactéries a été culottée et lavée une fois avec du PBS avant de remettre le culot en suspension dans 100 µl de solution de PKH (1 µl de fluorophore et 99 µl de diluant C). La coloration a été effectuée à l'obscurité à température ambiante pendant 15 minutes avec des mélanges fréquents. 200 µl de BSA ont été ajoutés pour arrêter la réaction. Les bactéries ont ensuite été sédimentées et remises en suspension dans du milieu minimal M9 (Sigma).

Préparation d'échantillon pour la microscopie de super-résolution. Une culture bactérienne de DO_{600nm} d'environs 0,05 a été culottée. Pour obtenir des bactéries fixées, 200µl de tampon PFA 2% dans du PBS ont été utilisé pour reprendre le culot, et le mélange a été incubé 30 minutes à température ambiante sur table agitante. Les bactéries ont été ensuite culottées et lavées dans 20µL de PBS. Cette opération a été réalisée trois fois puis les bactéries ont été reprises dans 200µl de PBS final et déposé sur lame pour observation en microscopie. Pour visualiser des bactéries non fixées, le culot a été remis en suspension dans du milieu minimal M9 (Sigma). La suspension cellulaire a été déposée sur un pad d'agarose composé de 1 % (p/vol) d'agarose à basse température de fusion (Fisher Scientific) dans un milieu minimal M9 (Sigma).

Microscopie de super-résolution (réalisée avec la plateforme TRI-GENOTOUL). Les images de mycobactéries ont été abtenues à l'aide du microscope Zeiss ELYRA 7 Sim Lattice avec un objectif Zeiss 63× NA = 1.4. L'analyse des images a été réalisée avec le logiciel ZEN de Zeiss.

Mesure de la fluorescence des bactéries par spectrofluorimétrie. La DO600nm de cultures bactériennes a été mesurée, puis ces cultures ont été culottées. Les culots ont été lavés dans du PBS 3 fois puis repris dans du milieu minimal M9 (Sigma) de façon à normaliser les DO60nm des différentes cultures à comparer. La fluorescence a alors été mesurée selon les recommandations du fabricant pour chaque protéine fluorescente à l'aide du lecteur de microplaques CLARIOstar BMG LABTECH.

Tableau 13. Listes des composés inhibiteurs putatifs sélectionnés

Nom	Source	Masse molaire (g/mol)	Caractéristique
OGT2599	Szafranski-Schneider, *et al., 2012*[187]		Inhibiteur de PMT de levure
R3A-5a	Orchard, *et al., 2004*[118]		Inhibiteur de PMT de levure
TK203	CT	557	Inhibiteur de glycosyltransférase
TK220	CT	603	Inhibiteur de glycosyltransférase
TK292	CT	617	Inhibiteur de glycosyltransférase
JSPB	CT	299	Inhibiteur de glycosyltransférase
S70-2	CT	299	Inhibiteur de glycosyltransférase
D155931	KBH	499,99	Dérivé de rhodanine
D157070	KBH	499,99	Dérivé de rhodanine
Pyrrolidine	Sigma	71,12	Dérivé de Proline
Ac-PGP trifluoroacetate salt hydrate	Sigma	311,33	Dérivé de Proline
Pict787	Chimiothèque de la plateforme PICT	381,49	Dérivé de Proline
Pict1350	Chimiothèque de la plateforme PICT	397,47	Dérivé de Proline
Pict1792	Chimiothèque de la plateforme PICT	231,25	Dérivé de Proline
Pict1795	Chimiothèque de la plateforme PICT	261,34	Dérivé de Proline
Pict1906	Chimiothèque de la plateforme PICT	347,46	Dérivé de Proline
Pict1910	Chimiothèque de la plateforme PICT	251,31	Dérivé de Proline
Pict2365	Chimiothèque de la plateforme PICT	250,34	Dérivé de Proline
Linezolid	Sigma	337,35	Dérivé de Proline
KYP-2047	Sigma	339,43	Dérivé de Proline
D-Proline	Sigma	115,13	Dérivé de Proline
L-Proline	Sigma	115,13	Dérivé de Proline
474705Z01	Pierre Fabre	405,47	Dérivé de rhodanine
477335Z01	Pierre Fabre	325,41	Dérivé de rhodanine
474841Z01	Pierre Fabre	446,6	Dérivé de rhodanine
474737Z01	Pierre Fabre	405,47	Dérivé de rhodanine
466243Z01	Pierre Fabre	379,84	Dérivé de rhodanine
463271Z01	Pierre Fabre	356,43	Dérivé de rhodanine
478185Z01	Pierre Fabre	337,42	Dérivé de rhodanine
S-2-pyrrolidinemethanol	Sigma	101,15	Dérivé de Proline

3. Test d'inhibition de PMTub

Choix rationnel d'inhibiteurs putatifs. Des molécules ont été sélectionnées rationnellement pour le potentiel effet sur l'activité de PMTub basé sur les connaissances acquises sur cette enzyme lors de cette étude. Les molécules choisies et leurs caractéristiques sont décrites **Tableau 13**.

Crible d'inhibiteurs avec le test phénotypique d'activité. Des plaques 96 puits ont été ensemencés à partir d'une pré culture d'une semaine environs (DO_{600nm} au moins de 1.5). Les puits ont été préalablement remplis de milieu LBT supplémenté en antibiotiques (**Tableau 7**) et en composé inhibiteur putatif à la concentration souhaitée. Afin d'ensemencer tous les puits de la même façon, un protocole de resuspension des bactéries est appliqué afin d'obtenir une solution homogène sans agrégats de bactéries. La pré culture de 25mL est centrifugé 10 minutes à 4°C 500xg. Le surnageant est jeté et 4-5 billes stériles (4mm) sont ajouté dans le tube contenant le culot. Le tube est agité doucement afin « d'étaler » le culot sur les parois du tube avec les billes. 20mL de milieu LBT froid est ajouté au tube et le liquide est transféré dans un tube neuf pour éliminer les billes et restes de culot. Le nouveau tube est centrifugé 10 minutes à 800rpm pour retirer les agrégats. Le surnageant est récupéré est sa DO_{600nm} est mesurée. Les différents puits sont alors ensemencés par un volume de surnageant nécessaire pour atteindre une DO_{600nm} finale de 0,1 dans les puits. Après 7 jours de culture à 37°C, la DO_{600nm} de chaque puit est mesurée et les surnageant des cultures sont récupérées et analysées comme décrit précédemment (voir p.197).

Construction d'un plasmide intégratif pour l'expression de la FasC. Les plasmides pWM158 et pWM218 ont été digérés par les enzymes de restriction XbaI et SfiI (Fisher Scientific). Les deux produits de digestion d'intérêt ont été purifiés sur gel par le kit Monarch DNA Gel Extraction (New England Biolabs). Les fragments purifiés ont été ligués par la T4 DNA ligase (Fisher Scientific) selon le protocole commercial. Les produits de ligation a été amplifié comme décrit plus haut (voir p.193) et le gène de la FasC a été vérifié par séquençage.

MVPVVSPGPLVPVADFGPLDRLRGWIVTGLITLLATVTRFLNLGSLTDAGTPIFDEKHYAPQAWQVL
NNHGVEDNPGYGLVVHPPVGKQLIAIGEAIFGYNGFGWRFTGALLGVVLVALVVRIVRRISRSTLVGA
IAGVLLICDGVSFVTARTALLDGFLTFFVVAAFGALIVDRDQVRERMHIALLAGRSAATVWGPRVGVR
WWRFGAGVLLGLACATKWSGVYFVLFFGAMALAFDVAARRQYQVQRPWLGTVRRDVLPSGYALG
LIPFAVYLATYAPWFASETAIDRHAVGQAVGRNSVVPLPDAVRSLWHYTAKAFHFHAGLTNSAGNY
HPWESKPWTWPMSLRPVLYAIDQQDVAGCGAQSCVKAEMLVGTPAMWWLAVPVLAYAGWRM
FVRRDWRYAVVLVGYCAGWLPWFADIDRQMYFFYAATMAPFLVMGISLVLGDILYHPGQGSERRT
LGLIVVCCYVALVVTNFAWLYPVLTGLPISQQTWNLEIWLPSWR

Figure 59. Séquence en acide aminé de PMTub présent sur le plasmide pWM158.

5’-ATGGCAGGTCTGACCAATAGCGCAGGTAATTATCATCCGTGGGAAAGCAAACCGTGGACCT
GGCCGATGAGCCTGCGTCCGGTTCTGTATGCAATTGATCAGCAGGATGTTGCAGGTTGTGGTG
CACAGAGCTGTGTTAAAGCAGAAATGCTGGTTGGTACCCCGGCAATG-3’

Figure 60. Séquence ADN insérée dans pGex-6P, correspondant au peptide EL4.

C. Modélisation

Construction des différents modèles de PMTub. La modélisation par homologie de PMTub a été réalisée à l'aide des serveurs web : Phyre2[154], I-tasser [155–157] et Robetta [158,159]. Le modèle avec le meilleur score de fiabilité a été sélectionné pour chaque serveur. La séquence de PMTub utilisée correspond à la séquence présente sur le plasmide pWM158 (**Figure 59**). Pour chaque modèle, une représentation de Ramachandran a été établie grâce au serveur PDBsum[210].

D. Production et étude *in vitro* du peptide EL4

Construction et production de pGex-EL4. Le plasmide pGex-AL4 a été construit par digestion du plasmide commercial pGex-6P et d'une séquence synthétique (**Figure 60**) par les enzymes de restriction XhoI et NdeI (Fisher Scientific). Les deux produits de digestion d'intérêt ont été purifiés sur gel par le kit Monarch DNA Gel Extraction (New England Biolabs). Les fragments purifiés ont été ligués par la T4 DNA ligase (Fisher Scientific) selon le protocole commercial. Les produits de ligation ont été amplifié comme décrit plus haut (voir p.193) et l'insertion de la séqueunce de EL4 a été vérifié par séquençage. Le plasmide pGex-EL4 a alors servi à transformer différentes différentes souches bactériennes : DH5α et BL21-DE3 (**Tableau 6**). Les bactéries transformées ont été cultivées en milieu LB en présence d'ampicilline (**Tableau 7**) pour la production de la protéine GST-EL4. Plusieurs protocoles de production ont été testés : culture à 37°C, 30°C ou 16°C, avec ou sans IPTG à différentes concentrations (0,05 à 0,5 mM) et culture longue ou courte (de 1h à sur la nuit). Les bactéries ont été récoltées par centrifugation de 5 minutes à 5000 xg. Le culot bactérien a alors été conservé à -20°C pour usage ultérieur ou directement repris dans du tampon d'extraction (PBS, 5% glycerol, 2mM dithiothreitol, 1% Triton X-100, inhibiteur de protéases, 0, 25mg/mL de lysozyme et benzonase). Les bactéries ont été lysées par sonication brève (4cycles de 15s de sonication entrecoupés de 30secondes sur glace). Les fractions soluble et insoluble ont ensuite été séparées par une centrifugation à 27000 xg pendant 20 minutes. Le culot a été repris dans du tampon d'extraction.

Purification et analyse du peptide EL4. La purification de GST-EL4 a été réalisée sur les fractions solubles sur milieu glutathion sépharose (Cytiva) selon le protocole commercial. Le peptide EL4 a été récupéré en utilisant la protéase PreScission (Cytivia). 100 µl de milieu

sépharose ont été incubés avec 8 µl de PreScission et 92 µl de tampon de clivage pendant 4 heures à 4°C avec agitation douce. Le milieu glutathion sépharose a été sédimenté par centrifugation à 500 xg et le surnageant a été récolté. Le surnageant est dans un premier temps analysé par coloration au nitrate d'argent. Après migration du surnageant sur SDS-PAGE (gel précoulé 16,5%, biorad), le gel a été incubé dans le tampon de fixation (40% éthanol et 10% d'acide acétique) pendant 30 minutes. Le gel a ensuite été incubé dans du tampon de sensibilisation (30% éthanol, 6M thiosulfate de sodium, pH 6) pendant 30 minutes. Après 3 lavages dans de l'eau, le gel a été incubé dans une solution de nitrate d'argent (11 M de nitrate d'argent, 0,01% de formaldéhyde). Le gel a ensuite été lavé à deux reprises dans de l'eau et incubé dans un tampon de révélation (1,2% de carbonate de sodium, 0,01% de formaldéhyde) jusqu'à apparition d'un signal. La révélation est arrêtée par incubation dans un tampon stop (50 mM EDTA). Dans un deuxième temps, une analyse par spectrométrie de masse est réalisée (par la plateforme de l'IPBS) dans des conditions similaires à celles décrites pour l'étude de PMTub-mCherry par spectrométrie de masse (voir p. 203).

GST Pull down. Les fractions solubles contenant GST-EL4, obtenues comme décrit précédemment, sont utilisées ici (voir p.211). La protéine recombinante GST-EL4 a été isolée en utilisant un milieu glutathion sépharose (Cytiva) tel que décrit dans le protocole commercial en utilisant le tampon de lyse (PBS, 5% glycerol, 2mM dithiothreitol, 1% Triton X-100, inhibiteur de protéases, 0, 25 mg/mL de lysozyme et benzonase). Le milieu portant la GST-EL4 a alors été utilisé pour le GST pull down. Pour cela, celui-ci a été incubé sur la nuit avec agitation douce en présence de surnageant de culture de PMM186:pMV361:pWM218 (souche No PMT exprimant de la FasC non glycosylée) préalablement concentré 50 fois à l'aide d'un centricon 10 kD (Millipore) puis dilué 3 fois dans du tampon de lyse. Le milieu a ensuite été sédimenté par centrifugation à 500xg pendant 5 minutes et le culot a été lavé dans du tampon de lyse. Le lavage a été réalisé trois fois. Après le troisième lavage, le milieu a été récupéré et directement analysé par western blot anti-His et anti-GST pour détecter la présence de FasC et de GST-EL4.

Isothermal Titration Calorimetry (ITC) (en collaboration avec le Dr Virginie Gervais de la plateforme PICT). Les mesures d'ITC ont été réalisées à l'aide du système VP-ITC de MicroCal Inc. L'appareil utilisé possède une cellule de mesure de 1,4-2mL. Les peptides EL4 et d'Apa ont été solubilisés dans le tampon de titration (phosphate de potassium 5mM pH6,5)

aux concentrations indiquées dans la Figure 49. La solution de peptide EL4 dans la cellule de mesure a été directement titrée avec le peptide Apa pour définir la courbe de titration. L'intégration des signaux de titrage, engendrés par l'interaction EL4/Apa, a permis de générer des valeurs de ΔH, représentées en fonction du ratio molaire entre les 2 interactants. Les données de titration résultantes ont été modélisées par une courbe de tendance selon un modèle d'interaction à un site lorsque cela était possible.

RMN (en collaboration avec le Dr Martine Gilleron, collaboratrice de l'équipe). Les échantillons de peptide EL4 ont été solubilisés à3,2 mg/mL dans du tampon phosphate 20mM, pH 6,5. Les expériences de RMN ont été effectuées sur une spectromètre Brucker 600 à 295 ou 300K. Des spectres RMN 1D ^{1}H 600MHz et RMN 2D ^{1}H-^{1}H TOCSY 80 ms (partiels, correspondant à la zone des corrélations NH-CαH) ont été acquis.

Dichroïsme circulaire (réalisé avec la plateforme PICT). La structure secondaire du peptide EL4 a été étudiée par dichroïsme circulaire. Le peptide a été repris dans de l'eau et les acquisitions ont été réalisées à 25°C. Les résultats sont exprimés par l'ellipticité moyenne des résidus [Θ] (en deg.cm^2/dmol) en fonction de la longueur d'onde.

Expérience de suivi de dénaturation thermique (réalisé avec la plateforme PICT). La dénaturation du peptide EL4 par la température a été suivie par le ratio 350nm/330nm sur un spectrofluorimétrie Eclipse (Varian Cary). Le peptide EL4 a été repris dans un tampon phosphate 20mM, pH 7.0, 25 °C, à une concentration protéique de 20 µM.

Tableau 14. Amorces utilisées pour la construction du plasmide pSMjF et la mutagenèse du gène *gfp*. Les amorces permettant de générer les Fragments n°1-4 ont été conçues par le logiciel NEBuilder®. Leur queue flottante (lettres minuscules) sert à la construction du pSMjF selon une méthode de clonage par assemblage. Les amorces complémentaires de la mutagenèse dirigée introduisent le codon ambre (souligné) et le site de restriction de Alw44I (surligné en gris) dans le gène cible gfp.

Nom	Sequence	Site de restriction
pSMjF		
1_for	ccgcggtgatTGGACAAGCCAGTTGGATTC	/
1_rev	attcgtccatctgcagGTGATTCCTCCGGATCGG	Ajoute un site PstI
2_for	caatggccaagacaattgcggatccGGACGAATTTGAAATGATAAAGAG	Ajoute un site BamHI
2_rev	cagtcttaagTTATAATCTCTTTCTAATTGGCTCTAAAATC	Ajoute un site AflII
3_for	gagattataacttaagACTGCCAGGCATCAAATAAAAC	Ajoute un site AflII
3_rev	tagagcatgcAAAATGACCCCAGAGGTCG	/
4_for	gggtcattttgcatgcTCTAGAAGAGGTCAACGCC	Ajoute un site PaeI
4_rer	gctcgatgatccccgggtacgtaATCACCGCGGCCATGATG	/
GFP		
Froward	CTGACCCTGAAGTTCTAGTGCACCACCG	Ajoute un site Alw44I
Reverse	CGGTGGTGCACTAGAACTTCAGGGTCAG	Ajoute un site Alw44I

Tableau 15. Plasmides utilisés pour la construction du plasmide pSMjF

Nom	Caractéristiques principales	Source
pWM218	pBlaF*-FasC-His^6-T_{rrnBt1}, OriE, OriM, Hyg^R	Liu *et al*, 2013[52]
pSMT3-M	pHsp60, OriE, OriM, Hyg^R	Carroll et al, 2010[211] Addgene plasmid # 26589
pEVOL-pBpa	pBAD-MjTyrRS (2 copies), p15Aori, Cm^R	Chin *et al*, 2002[212] Addgene plasmid # 31190

5’-ACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCG
CTAGCGATTTGGCGAGCTTCGTGCGTGTTCGGTAGCCTGGCATTTACCGACCGGCGGTAGTTCAGCAGGGCAG
AACGGCGGACTCTAAATCCGCATGGCGCTGGTTCAAATCCGGCCCGCCGGACCAGCGGAAATGACCCCAGAGG
TGCGACCTCTGGGGTCATTTT-3’

Figure 61. Séquence synthétique commandée pour la construction de pSMjF

E. Approche de biologie synthétique

Construction du plasmide pSMjF et dérivés. Le plasmide pSMjF a été construit à l'aide de la technique du « Gibson assembly »[206]. Les différents fragments d'ADN ont été amplifiés par PCR en utilisant la Platinum SuperFi II DNA Polymerase (invitrogen) à partir de 4 sources plasmidiques (**Tableau 15** et **Figure 61**) à l'aide d'amorces adaptées (**Tableau 14**). Le Kit « NEBuilder HiFi DNA Assembly Cloning » (New England Biolabs) a été utilisé selon les recommandations du fabricant et la solution de réaction a été directement transformée dans *E. coli* stellar (**Tableau 6**). Les plasmides ainsi construits ont été récupérés et produits tels que décrits plus haut (voir p.193). Les constructions ont été vérifiées dans un premier temps par profil de restriction et ensuite par séquençage du plasmide en intégralité. Le plasmide pSMT3-FasC a été construit par digestion du plasmide pSMjF par XbaI (Fisher Scientific). Cette digestion a résulté en deux fragments, dont un portant les gènes du système orthogonal. Le second fragment a été purifié sur gel en utilisant le kit Monarch DNA Gel Extraction (New England Biolabs) et recircularisé en utilisant la T4 DNA ligase (Fisher Scientific). Le plasmide ainsi obtenu a été produit comme décrit précédemment (voir p.193) et vérifié par séquençage intégral. Le plasmide pSMjF-long a été construit par digestion de pWM158 et pSMjF par XbaI. Les fragments d'intérêt (plasmide linéarisé pour pSMjF et pBlaF* pour pWM158) ont été purifiés sur gel en utilisant le kit Monarch DNA Gel Extraction (New England Biolabs) et ligués en utilisant la T4 DNA ligase (Fisher Scientific). Le produit de ligation a été transformé chez *E. coli* pour production comme décrit plus haut (voir p.193) et vérifié par séquençage.

Construction et test du promoteur mini-pBlaF*. Le plasmide pWM158-mini-pBlaF* codant pour PMTub sous la dépendance du promoteur réduit « mini-plaF* », a été construit par digestion du plasmide pWM158. La digestion a été réalisée par XbaI (Fisher Scientific) selon le protocole commercial. Les différents fragments de digestion ont été séparés sur gel d'agarose, et le fragment de taille la plus importante a été purifié en utilisant le kit Monarch DNA Gel Extraction (New England Biolabs). Ce fragment a été recircularisé par la T4 DNA ligase (Fisher Scientific). Le plasmide a ensuite été produit comme décrit plus haut (voir p.193) et vérifié par séquençage. L'activité de PMTub sous la dépendance du promoteur mini-pBlaF* a été mesurée en utilisant le test d'activité phénotypique, par 2 ELISA successifs, tel que décrit précédemment (voir p.197).

Tableau 16. Amorces utilisées pour l'introduction d'un codon ambre dans PMTub. Les séquences soulignées correspondent aux mutations, celles surlignées en gris mettent en évidence les sites de restriction ajoutés ou supprimés par mutation silencieuse pour le contrôle de l'insertion de la mutation par profil de restriction et en gras est indiqué le codon ambre inseré.

Nom		Sequence	Site de restriction
F117TAG	for	GAGGCCATC**TAG**GGTTATAACGGATTCGGC	Ajoute un site PsiI
	rev	GCCGAATCCGTTATAACC**CTA**GATGGCCTC	
V378TAG	rev	CAACGACATGGGCCAGGTCCACGGTTTG	Retire un site Eco52I
	for	TGGCCG**TAG**CTGTACGCCATCGACCAGCAAG	

Construction du rapporteur GFP de fonctionnalité du système orthogonal. Le plasmide muté en position 47 de la GFP est dérivé du plasmide pGMCS-GFP-PMbcA-RFP qui nous a été généreusement fourni par le Dr Claude Gutierrez de l'IPBS. Il a été muté par quickchange en utilisant les amorces décrites **Tableau 14** et utilisant la Platinum SuperFi II DNA Polymerase (invitrogen). L'introduction de la mutation a été vérifiée par profil de restriction puis séquençage.

Insertion d'un codon ambre dans la séquence de PMTub. La mutagenèse de PMTub pour l'insertion d'un codon ambre a été réalisée par PCR. Le plasmide pWM158 a été amplifié par un couple d'amorces construit de sorte à insérer la mutation souhaitée et une mutation silencieuse ajoutant un site de restriction (Tableau 16). La PCR a permis d'obtenir un produit linéaire dont les extrémités sont phosphorylées par la T4 polynucléotide kinase (Fisher Scientific) puis circularisé par la T4 DNA ligase (Fisher Scientific). Le plasmide ainsi obtenu est produit comme décrit précédemment (voir p.193) et la séquence du gène de PMTub est vérifié par séquençage.

Expression de protéines contenant un acide aminé non naturel. Les bactéries exprimant un gène portant un codon ambre ont été cultivée à 37°C avec agitation dans du milieu LBT (LB et Tween80 80 à 0,05 %) supplémenté en antibiotiques (**Tableau 7**) jusqu'à atteindre un DO_{600nm} de 1. À partir de ces précultures, une culture de 20 mL de 7H9 a été ensemencée à une DO_{600nm} de 0,05. Lorsque la DO_{600nm} a atteint 0,2, du pBpa a été ajouté au milieu pour obtenir une concentration finale de 2 mM. Les cultures ainsi supplémentées ont été incubées à 37°C dans le noir avec agitation sur la nuit. Le lendemain, les cellules ont été récoltées et, selon les besoins expérimentaux, analysées par spectrofluorimétrie ou utilisées pour un test d'activité PhoA tel que décrit précédemment (voir p.207 et 206).

Annexe – Article en préparation

Insights into the M tuberculosis protein o mannosyl transferase structure function relationships: a step forward rational search for specific inhibitors of high therapeutic potential.

Nicolas Géraud, Camille Falcou, …..Jérôme Nigou, Emeline Fabre✉, Michel Rivière✉

Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France.

✉To whom correspondence may be addressed: Institut de Pharmacologie et de Biologie Structurale, CNRS – Université de Toulouse, 205 route de Narbonne, BP 64182, F-31077 Toulouse Cedex 4, France.

Emeline Fabre✉ E-mail: emeline.fabre@ipbs.

Michel Rivière, Tel.: 33-561-175-558; Fax: 33-561-175-994; fr

Introduction

With 28,000 new patients falling ill and nearly 4,000 deaths every day, tuberculosis (TB) remains one of the world's deadliest infections with devastating health, social and economic consequences(WHO, 2021). This dramatic situation is worsened by the threatening emergence of antibiotic resistant M. tuberculosis (Mtb) strains challenging our therapeutic arsenal currently based on poly-antibiotherapy regimens. The increasing incidence of antibio-resistant TB then highlights the urgent need for innovative therapeutic approaches to achieve the End-TB program target of reducing the number of new cases by 80% by 2030(WHO, 2015).

Currently, two complementary routes to conventional treatment regimens raise great hope and are actively explored: i) host-directed therapies able to potentialize the natural defenses of the infected host, and ii) anti-virulence factor approaches capable of disarming the infectious agent by targeting non-essential factors that are unlikely to induce mutational escapes leading to the selection of resistant bacteria. In a previous study, we showed that mycobacterial mannoproteins meet this criterion. Indeed, we found that the invalidation of the protein-O-mannosylation post-translational pathway does not affect the growth in vitro of Mtb but leads to a drastic reduction in its virulence in mice. This original observation makes the enzyme responsible for this process, namely MtPMT, the protein mannosyl transferase (PMT) of Mtb, a potential target of therapeutic value for the search for alternative strategy to fight against Mtb. This assertion is underpinned by the fact that MtPMT belongs to the C-type glycosyl transferase (GT-C) superfamily just like the arabinosyl-transferases EmbA, EmbB and EmbC involved in Mtb cell wall biosynthesis and identified as the primary targets of the first-line anti-tuberculosis drug ethambutol. However, to date, MtPMT remains an orphan target against which no specific inhibitor is available to disrupt pharmacologically the protein-O-mannosylation pathway to assess its relevance as potential therapeutic target. Moreover, the scarce knowledge existing about the structure-function relationships of the PMTs in general and in particular of the bacterial PMT of 522 amino acids coded by the rv1002c gene of Mtb, strongly hinders the rational search for potential specific inhibitors.

PMTs have been experimentally assessed in animals (Human, mouse and rat, zebra fish, drosophila) and fungi (Saccharomyces, Aspergillus, Candida, Trichoderma, Cryptococcus) and PMT genes are found in the vast majority of the opisthokonta. In contrast, up to now, PMT mediated protein-*O*-mannosylation (EC2.4.1.109) process in bacteria have been experimentally detected into the actinobacteria *phylum* only, including C. glutamicum, *S. coelicolor* and *M. tuberculosis*. ((Lara, 2004, Mahne et al., 2006, Wehmeier et al., 2009). From the functional point of view, PMTs are “non-Leloir” glycosyltransferases of the Cazy GT-39 family that shares with all members of the GT-C superfamily, the peculiarity of using an activated phospho-isoprenoid sugar donor substrate to transfer a mannosyl residue to the acceptor hydroxyl group of a threonine or a serine of a nascent protein substrate (Fig.1A). This GTC’s characteristic feature is likely supported by a preserved fold characterized by an integral polytopic membrane topology resulting from the presence of 11 to 15 hydrophobic transmembrane helices (TMH) delimiting a variable number of soluble loops facing the cytosol or protruding in the endoplasmic reticulum (ER) lumina of eukaryotes or in the extracellular

periplasmic space of bacteria. Due to this strong embedment in the membrane that greatly limits their isolation and their subsequent structural analyzes, the GT-C are much less well characterized compared to the two other GT families, namely the GT-A and the GT-B. Indeed, to date, less than fifteen GT-C structures are deposited in protein structure data banks, among which two PMT only. These correspond to the eukaryotic *S. cerevisiae* ScPMT1 and ScPMT2 recently resolved by CryoEM in a dimeric complex encompassing a lipid donor analogue and peptide acceptor substrate (a dolichyl-phosphate and the tetrapeptide PYTV associated with the ScPMT1 and ScPMT2 subunits respectively). Compared to the two presumed topologies previously proposed for these PMTs, the resolved complex definitively confirms that both subunits adopt very similar 11 TMH based folds. In addition, these structures reveal that the PMTs highly conserved catalytic amino acid doublet DE (D77E78 in ScPMT1 and D92E93 in ScPMT2) located in the external loop EL1, are likely to be assisted mechanistically by the conserved spatially closely located amino acids doublets D189-R649 and D203-R671 (in ScPMT1 and ScPMT2 respectively). These two amino acids (aligned with D16 and R441 of MtPMT) located distantly in the external loops EL2 and EL5 respectively have been proposed to complete the catalytic residues as also reported for several oligosaccharyltransferase enzymes of the GT-C superfamily (Albuquerque 21, 24, 29) (**Fig.1B, upper panel**). Besides, scrutiny of the spatial (3D) surrounding of the ScPMT1/2 active sites also allowed to propose several amino acids including in particular the aromatic phenylalanine/tyrosine F652/Y674 from ScPMT1/2 (conserved among PMTs), to contribute to the binding site of the acceptor threonine (Bai 2019) (**Fig.1B, upper panel**). However, it is worth mentioning that, despite their presumed determinant contributions, remarkably, the functional implications of these residues has never been yet experimentally demonstrated for the PMT family members. As a last noteworthy point, the resolution of these structures describe the large soluble luminal fraction of ScPMTs corresponding to the EL 4 outer loops, carrying the 200 amino acid long MIR domain absent from bacterial PMT sequences.

In the light of these results and the subsidiary hypotheses they arise, we therefore undertook a mutational analysis of MtPMT with the aim of a better understanding of its structure-function relationships required for a rational search for inhibitory molecules in the future. For this, we conceived an original and robust bioassay that allows an indirect quantitative measuring of the *in-situ* activity of MtPMT expressed in a heterologous mycobacterial reporter strain. Owing to this bioassay described herein and to the drawing up of a predictive 3D model of MtPMT, we have been able to evaluate the impact of key amino acid mutations on the enzyme activity. Our results support the existence of an extended catalytic pocket that can be targeted by exogenous molecules capable of crossing the mycobacterial cell wall. Moreover, we afford new clues on the functional role of the bacterial MtPMT external loop EL4 devoid of the MIR triplicated domains of the eucaryote PMTs, opening the way for a rational search for specific inhibitors of the bacterial MtPMT.

Materials & Methods

Strains and culture: DNA constructs and site-directed mutagenesis. Mutations of the pmtub gene sequence were achieved by site-directed mutagenesis, following the quickchange method from Stratagen. Complementary Forward and Reverse primers were designed in order to introduce the desired mutation, alongside a silent mutation that adds or removes a restriction site (see primers sequence, Table S1). Mutations were introduced into the pmtub gene encoded by the pMW158 plasmid (see main plasmids characteristics, Table S2), by whole-plasmid PCR using the Phusion polymerase (ThermoFisher) with the commercial high-GC buffer and in the presence of 3% DMSO. PCR cycle were as follow: a first 3 minutes step at 98°C, followed by 30 cycles composed a 20 seconds denaturation step at 98°C, a 30 seconds annealing step at 65°C and a 5 minutes elongation step at 72°C, and a final 10 minutes elongation step at 72°C. Changes have been made to this protocol in order to better suit each primers characteristics (length and melting temperature): annealing temperature was varied between 55°C and 72°C and DMSO concentration was varied from 0% to 10%. PCR products were incubated with the DpnI enzyme (following manufacturer recommendations, ThermoFisher) in order to degrade the parental plasmid and then used to transform the E. coli electro-competent TOP10 strain (following manufacturer recommendations, LifeTechnologies). Plasmids pWM218, pWM158 and mutated-pWM158 were amplified and purified from E.coli overnight cultures (onto LB medium, in the presence of appropriate antibiotic to allow selection; Kanamycin 50 µg/mL or Hygromycine 200µg/ml) using the Qiagen QIaprep Spin Miniprep kit. Plasmid concentrations and purity were controlled by nanodrop. Their restriction profiles were established in order to confirm their identity and to verify eventually the acquisition of the desired mutations. This qualitative verification was further completed, for three clones of each construct, with the overall sequencing of the entire pmtub gene sequence (by Eurofins).

Mycobacterium Smegmatis transformation and culture. Plasmids and strains used in this study are presented in Table S2 and S3. Competent M. smegmatis strains were prepared ensuing the following steps: bacteria to transform were grown during one week in 10 mL of LB medium with Tween 80 0.05% (called LBT medium) until an OD600nm of 1. 500 µL of the week-old pre-culture were taken up to inoculate 50 mL of LBT until an OD600nm of 0.2-0.8 and then centrifuged (8min., 4°C, 500 xg). Cell pellets were washed 3 times with decreasing volumes of cold Washing Buffer (Tween 80 0.05%, Glycerol 10%) : 50mL, 15mL and 1mL with centrifugation (8min., 4°C, 500 xg) between each washing step. Washed competent cells were divided in 200 µl aliquots. 1 µL of plasmid (at 200 ng/µl) was added into each aliquot of competent M. smegmatis cells and electroporation was performed at 2500V, 25µf, 200Ω. Immediately after transformation, 800 µl of LBT medium were added to bacteria and those were incubated during 3 hours at 37°C while shaking (180 rpm). Transformed bacteria were then plated onto LB-agar containing the appropriate antibiotics (25 µg/mL of Kanamycin and/or 50 µg/mL of Hygromycin) and incubated at 37°C for 5 days. Isolated colonies from the LB-agar plate were grown in 10 mL of LBT medium containing the appropriate antibiotics at 37°C during 9 days (up to OD600nm 1) to be either transformed again (in order to introduce the FasC encoding plasmid) or further analyzed in terms of phenotype. For the latter case, cells were centrifuged (10 min., 4°C, 500 xg) and the supernatants and pellets stored separately at -20°C.

In vivo « Chemical rescue » “Chemical rescue” for the restoration of the activity of PMTub mutated on the catalytic triad. 9 mL of 1 week old cultures of the recombinant M. smeg strains containing the enzymes carrying the D74A, D176A and R441A mutations, as well as the two control strains, native PMTub and No PMT (Table 2), were centrifuged for 10 minutes at 4000 rpm and 4 ° C. The pellets were taken up using sterile glass beads (4 mm) in 20 mL of LBT, then centrifuged again for 10 minutes at 800 rpm to remove cell aggregates and keep only supernatants of cells supposed to be dissociated. , whose A600nm were measured. These dissociated cell supernatants were then used to start micro-cultures in 96-well plates. For this, the wells containing LBTKH medium (LB, Tween 80 0.05%, Kanamycin 25 µg / mL, Hygromycin 50 µg / mL) were inoculated at an A600nm of 0.01. The plates containing the micro-cultures were incubated at 37 ° C. with the bacteria being resuspended by gentle daily shaking. The A600nm of the micro-cultures were measured after 48h, and the chemical compounds were added to the culture medium according to the following plan: the strains expressing PMTub-R441A, native PMTub and No PMT were modified with 2.5 mM or 5mM imidazole or guanidine hydrochloride; PMTub-D74A, D176A, native and No PMT enzymes were transformed with 2.5mM or 5mM Sodium Acetate or Sodium Azide. Each culture was also grown without treatment, in control.

Bioessay: Study of FasC6His mannosylation by Direct- and Lectin-ELISA. Supernatants were first analyzed by Direct-ELISA in order to dose FasC6His concentration. 100µl of culture supernatant diluted twice in coating buffer (0.53% Na2CO3, 0.42% NaHCO3, pH 9.6) were incubated in triplicate, overnight at room temperature, in 96-well plates. After washing 3 times with 200 µL of PBS, the plate was incubated 1 hour at room temperature with 200 µL of PBS containing 0.05% of Tween20 and 5% of milk. After washing 3 times with 200 µL of PBS containing 0.05% of Tween20, the plate was incubated 2 hours at room temperature with 50 µL of anti-His antibody (Anti-His (C-term)-HRP antibody, Invitrogen) diluted 1000x in PBS containing 0.05% of Tween20 and 0.5% of milk. After four washing with 200 µl of PBS, revelation was performed with 50 µl of TMB solution (eBioscience), after which 25 µL of stop solution (1mM sulphuric acid) was added before reaching saturation (10 to 20 minutes). Standard range of purified FasC was performed for each Direct-ELISA with the following concentrations (in ng/ml): 50, 100, 200, 300, 500, 700, 900 and 2000. Secondly, supernatants with normalized concentrations of FasC6His (deduced from the previous Direct-ELISA) were analyzed for FasC6His mannosylation by Lectin-ELISA. For this second test, 96-well plates with anti-His antibody-coated wells (Genscript) were used. 100 µl of culture supernatants (diluted twice in PBS) were added per well in triplicate. The plate was incubated 3 hours at room temperature and then washed 4 times with 260 µl of PBS containing 2% Tween20. 100µl of Concanavalin A-HRP solution (0,5 µg/mL ConA-HRP, 1mM CaCl2, 1mM MgCl2, 1mM MnCl2, 0.05% Tween20, in PBS) were added in each well and the plate was incubated at room temperature during 2 hours. After four washing with 260 µl of PBS, revelation was performed with 100 µl of TMB solution (eBioscience), after which 50 µL of stop solution (1mM sulphuric acid) was added before reaching saturation (10 to 20 minutes). Absorbances from the two ELISAs were measured at 450 nm using CLARIOstar BMG LABTECH Microplate reader.

FasC6His purification. 3 mL of 9-days-old culture supernatant of M. smegmatis, diluted three times in binding buffer (50 mM NaH2PO4, 300 mM NaCl, 10 mM imidazole, pH 8) were loaded onto Ni-NTA spin column (Qiagen). After 5 minutes of centrifugation at 270 xg, the flowthrough was discarded. Then, the column was washed twice by centrifugation (2 min., 890 xg) with 600 µL of washing buffer (50 mM NaH2PO4, 300mM NaCl, 20mM imidazole, pH 8). Proteins were eluted twice with 100 µL of elution buffer (50 mM NaH2PO4, 300 mM NaCl, 500 mM imidazole, pH 8) by centrifugation (2 min., 890 xg). For each sample, aliquots of each purification step were evaluated by SDS-PAGE. The first eluate (containing most of the purified FasC) was then assessed using the BCA protein assay (Thermo Fisher, following manufacturer instructions) to measure FasC concentration.

Mass spectrometry Study of FasC6His mannosylation by Mass spectrometry (work supervised by Dr Julien Parra). Purified FasC samples were diluted in loading buffer (2% ACN, 0.1% TFA) to reach a final concentration of 0.4 µM. NanoLC-MS analyses were performed on a nanoRS UHPLC system (Dionex) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific). 5 µL of sample was loaded on a reverse-phase C4-precolumn (300 µm i.d. x 5mm; Thermo Fisher Scientific) at 20 µL/min in 2% acetonitrile and 0.05% TFA. After 5min of desalting, the precolumn was switched online with an analytical C4 nanocolumn (75 µm i.d. x 15 cm; in-house packed with C4 Reprosil) equilibrated in 95% solvent A (0.2% formic acid) and 5% solvent B (0.2% formic acid in acetonitrile). Proteins were eluted using a linear gradient from 5% B to 100% in 38 minutes at a flow rate of 300 nL/min. MS scans were acquired in the 800-2000 m/z range with the resolution set to a value of 60,000, and spectra were deconvoluted using Xtract software (Thermo Fisher Scientific).

Contrôle of MtPMT expression by WB : Anti-strep Western Blot on cell lysate. 9-days-old 100mL cell culture were lysed and cell fractions were prepared as described in Cooper et al, 200252. Briefly, cell pellet is sonicated after centrifugation and cell lysates are separated in sub fractions (cell wall, membrane and cytoplasm) by a succession of ultra-centrifugations. Cell fractions, separated by SDS-PAGE and transferred onto nitrocellulose membranes, were probed during 3 hours with anti-strep tag antibodies produced in mouse (1:5000 in PBS with 1%BSA and 0.05% Tween 20, Sigma) after blocking 3 hours in blocking solution (PBS with 5% BSA and 0.1% Tween 20). Proteins of interest were detected by 1 hour incubation with HRP-conjugated goat anti-mouse IgG antibody (1:5000 in PBS with 1%BSA and 0.05% Tween20, Sigma) and visualized with the Amersham ECL Prime Western blotting detection reagent (GE Healthcare), according to the provided protocol.

To be continued

Results

Model and Transmembrane localization and topology of MtPMT

For higher confidence, we used two conceptually different protein 3D structure prediction web-softwares, *i.e.* Phyre2 (Soding 2005) and I-TASSER (Yang et al, 2015) to generate three-dimensional models of the bacterial PMT. In both cases the maximum alignment coverage was obtained with the recently resolved cryo-EM structures of the Saccharomyces cerevisiae ScPMT1 and ScPMT2 (Ref PDB) (Bai, 2019) which share respectively 26% and 23% sequence identity with MtPMT.

Consistently, with these strong homologies found notably in the pairwise sequence alignments of the hydrophobic domains of MtPMT with those of yeast's homologues, the generated models predict the presence of 11 most probable transmembrane helices (TMHs) as previously predicted by VanderVen, 2005, (see alignment with ScPMT1 in supplementary data). These hydrophobic segments delineate a fairly constrained fold characteristic of the minimal core shared among the GT-C proteins. The predicted non-even number of TMHs constrains to a trans-topology presenting the protein extremities emerging on opposite side of the membrane. Therefore, by analogy with the demonstrated topology of eukaryotic PMTs embedded in the ER membrane, the bacterial MtPMT is predicted to adopt a plasma membrane-spanning configuration exposing its N-terminal extremity in the cytosolic compartment and its catalytic domain and C-terminal end into the periplasmic space.

To verify these predictions, we constructed a panel of recombinant mycobacterial strains expressing several MtPMT fusion proteins designed to report localization, level of expression or activity of MtPMT (Fig. 2A). All of these strains were constructed from a model organism derived from *M. smegmatis,* engineered to have its own PMT gene disrupted (*M. smegmatis* Δ(MsPMT)). For the control of the localization at the plasma membrane, we constructed a mCherry-N-terminal PMT fusion and then implemented microscopy experiments to detect red fluorescence at the membrane (Fig. 2B). Results are in perfect agreement with membrane insertion of MtPMT, as predicted by the 3D model. Notably, in parallel, we also constructed a MtPMT C-terminal mCherry fusion, but this construct lacked drastically of fluorescence intensity (not shown). Thus, to confirm and quantify periplasmic expression, we further constructed an Alkaline Phosphatase (PhoA) C-terminal fusion, and measured the corresponding enzymatic activity from mycobacterial periplasmic fractions (Fig. 2C). Results were compared to activities obtained for Mmps4 and KatG respectively, two mycobacterial control proteins with well-defined bacterial localization, Mmps4 being reported as membrane-inserted through a unique N-terminal TMH, with a mainly periplasmic orientation, and KatG as a cytosolic soluble protein.

To verify whether this transmembrane topology corresponds to the functional one we developed an original phenotypic bioassay that allows monitoring the in situ MtPMT activity in a recombinant *M. smegmatis* reporter strain.

Development of a novel quantitative PMT phenotypic assay

To monitor the transfer of mannose residue to protein by MtPMT, we co-expressed MtPMT with the *M. smegmatis* Fasciclin (FasC) a highly mannosylated reporter acceptor substrate (with 7 mannosylation sites precisely identified by LC-MS, see supplementary data). We previously demonstrated that this natural substrate of the endogenous MsPMT (Liu, 2013) (Fig.3D) is also mannosylated similarly by MtPMT expressed in the *M. smegmatis* MsPMT deficient strain (supplementary data). To allow the isolation of FasC from the culture medium, *M. smegmatis* Δ(MsPMT) was transformed with C-terminal 6xHis-tagged FasC. After normalization of the FasC-6xHis concentration (measured by anti-6xHis direct ELISA), its mannosylation was monitored by a lectin-ELISA using peroxidase coupled Concanavalin A (ConA), a lectin that binds to terminal mannose residues of oligosaccharides (Fig. 2D). Since MtPMT is the only enzyme that initiates protein-mannosylation, and considering that protein-mannosylation is linear in mycobacteria (ref), the measurement of mannose recognition by ConA allows the evaluation of the PMT activity. The sensitivity and reliability of this *in bacteria* protein mannosylation assay were first verified by measuring its response to different doses of FasC-6xHis (ranging fron 1 to 500 ng/well) produced by control strains, including *M. smegmatis* Δ(MsPMT) expressing WT-MtPMT, Y444F-mutated-MtPMT (showing a decreased PMT activity compared to WT-MtPMT), or No PMT (transformed with the vector control) (Fig. 2D) (see also results for S365A-MtPMT, L366A-MtPMT and Y371A-MtPMT in supplementary data). Results showed a dose-response effect of this assay, and that for each concentration of FasC-6xHis, the proportion of ConA binding detected for the 3 control strains is similar. Indeed, Y444F always represents around 30% of the response measured for the WT-MtPMT strain, for any FasC-6xHis concentration, from 50 to 500 ng per well. To validate the quantitative aspect of this assay, measures were also achieved by Liquid Chromatography (LC) coupled to entire-protein mass spectrometry (MS) after FasC-6xHis purification onto Ni-NTA columns. Results obtained for the WT-MtPMT strain confirmed that FasC-6xHis produced by this bacteria is highly glycosylated, with up to 17 linked glycosyl units detected (with mass differences of about 162 between each, ie the mass of a hexose residue) (Fig.2D). The observed mass of the most intense peak (23843.67 Da) was in agreement with the monoisotopic mass calculated for the non-glycosylated 6xHis-FasC (23844 Da). As expected, the spectrum obtained for the purified FasC-6xHis produced by the strain expressing no PMT showed only one peak, corresponding to the not-glycosylated protein, in agreement with the complete abrogation of protein glycosylation in the absence of PMT enzyme (Liu 2013). The absence of ConA binding for the No PMT strain underline the absence of background noise of this method. Results obtained for these two control strains are reliable in terms of presence or absence of FasC-6His glycosylation and thus can be used as a positive and a negative control for protein mannosylation. Besides, it clearly appears that the mutation Y444F affects the glycosylation of the FasC-6xHis, as fewer different glycoforms were detected (with up to 10 linked glycosyl units), and the sum of their respective abundances is inferior, compared to the WT-MtPMT strain. . In the same way, the mutations S365A, L366A and Y371A caused more or less attenuated PMT activities, following the same pattern of attenuation as detected both by ConA lectin assay and native MS based quantitative methods, (S365A>L366A>Y371A, see supplementary data)

Essentiality of conserved amino acids and targeting of MtPMT in living bacteria

Inspection of the conformational structural model of MtPMT, reveals in particular thatthe three conserved amino acids D176, R441 and Y444 distant in the sequence are spatialy close to the catalytic doublet DE74-75, and form a core-pocket at the periplasmic face of the MtPMT (Fig.1B) consistently with what observed in the ScPMT1 and ScPMT2 resolved structures.

The PMT phenotypic assay was used to assess the importance of the above-mentioned conserved residues for MtPMT activity. The amino acids residues D74, D176, R441 and Y444 were individually mutated into alanine, in order to disrupt the chemical functions of their respective lateral chains. Alternatively, aspartic acids were mutated into asparagine, and arginine into leucine to restore steric hindrance. Results showed that all of these mutations inactivated totally MtPMT (Fig. 3A and 3B). To substantiate these results, the mannosylation of purified FasC-6xHis secreted by the strain expressing Y444A-MtPMT was further analysed by LC-MS. As expected, this strain showed almost unsignificant protein glycosylation properties, with only two very weakly abundant glycoforms detected (corresponding to a maximum of two glycosyl units). Notably, when the DE74-75 was reversed into ED74-75, no PMT activity was detected neither, suggesting that the respective functions of these two residues might not be redundant nor exchangeable. The results indicate thus that these residues, conserved among GT-Cs (D74, D176, R441) or either PMTs only (Y444), are essential for *O*-mannosylation and most likely have similar biochemical functions within the different enzymes.

In order to establish this hypothesis, by excluding the influence of mutations on enzyme expression or localization, we quantified the periplasmic expression of all of the corresponding mutated enzymes, using PhoA C-terminal fusions (Fig. 3C). The results showed that the catalysis-disruptive mutations D74, D176, R441 and Y444 do not markedly affect the PhoA activity, emphasizing thus the importance of these residues for PMT catalysis but not expression. Furthermore, additional substitutions by equivalent biochemical functions, in mutants D74E, D176E, R441K and Y444F, rescued partially, but significantly, PMT activity (up to 6±1%, 4±1%, 83±1% and 30±0% respectively, Fig.3D), highlighting the essentiality of acid, basic or hydrophobic functions at these positions respectively.

To confirm the importance of the chemical function at these positions, we tentatively performed chemical exchange experiments by providing the corresponding alanine-mutant strains with various chemicals miming the lacking functions. Indeed, we expected that small chemical compounds could accommodate into the cavity formed at the extremity of the alanine lateral chain, and thus could compensate the mutationally disrupted chemical function, to restore PMT activity. In this aim, we used sodium acetate and azide (Camille 20, 21, 22), or guanidinium and imidazole (Camille, 23-25), to replace acid or basic groups respectively as reported in litterature. From all the combinations tested, only the activity of the mutant strain expressing R441A-MtPMT was rescued, by imidazole and guanidinium (up to 29±5%)(Fig. 3E or Fig. 3'A), evidencing that these compound not only cross the mycobacterial cell wall but are

able to reach membrane embedded MtPMT to modulate its activity. This remarkable result constitutes an indisputable proof of concept of the possibility of targeting MtPMT by small molecules *in-vivo.*

As these results confirmed the feasibility of targetting MtPMT in a biological context, we used the MtPMT phenotypic assay to screen for 30 putative inhibitors. We choose the candidate compounds based on two main criteria, a low molecular mass (below 500 KDa, in order to allow them to cross the mycobacterial membrane), and structure similarities with either rhodanine, or proline. Indeed, rhodanine derivatives were reported previously as inhibitors of yeast PMTs (refs), while Proline analogs were choosen to mimick the residues that surround preferentially the mannosylation sites (S/T) in MtPMT peptide substrates (sup data or fig. 5B). Among this set of compounds, the molecule Pict 787 standed out, providing up to 25±3% of inhibition, with logarithmic dose-effect (Fig. 3'B).

Specificities and essentiality of EL4

In order to search for original structural motifs specific of mycobacteria, and thus targetable by future potential drugs, we compared the sequence of MtPMT to that of the different eukaryotic PMTs and other GT-Cs, by pairwise alignments. We noticed that a particular region of MtPMT, of 10 to ~40 amino acids, systematically caused the insertion of gaps into the sequence of non-PMT GT-Cs (Fig. 4A). This MtPMT sequence located at the C-terminal end of EL4, is very conserved in actinobacteria, but much less conserved in eukaryotic PMTs (Fig. 4A). Interestingly, the localization of the gaps generated into the GT-C sequences corresponds to the C-terminus of EL5 in AfAglB and ClPglB, which are both reported as mobile, and that associates intimately with the acceptor peptide co-crystallized within ClPglB 3D structure (Fig. 4A, box).

To confirm and understand the functional involvement of this region in the PMT activity, we constructed different punctual mutants in 8 distinct positions, 2 conserved in all PMTs (indicated by stars in bold, Fig. 4A), and 6 specific of actinobacteria (indicated by thin stars, Fig. 4A). For almost each residue, we designed a destructive mutation of the biochemical function of the side chain (mainly by substitution with an alanine), or a compensating one, (Fig. 4B), and then we carried out the PMT phenotypic assay.

Among the mutations that we designed to target specific positions of mycobacteria, none caused significant loss of activity (Fig. 4B).

On the other hand, the two positions conserved in all PMTs (bacterial and eukaryotic), provided evidence for the involvement of the EL4 PMT-specific region in protein mannosylation activity. Indeed, the mutation into alanines of tryptophan W362 and tyrosine Y371, prevented respectively totally or by half MtPMT activity. Interestingly, both are aromatic residues capable of generating hydrophobic interactions. However their substitution in phenylalanine gave distinct results. The substitution of the essential W362 into phenylalanine only restored 6±2% of activity (Fig. 4B), attesting to the need for aromaticity but still the fine

structure of this residue. Besides, the total rescuing observed for Y371F (Fig. 4B), as well as its replacement by a leucine in ScPMT2, or a phenylalanine in HsPomt2 (Fig. 4A) showed the importance of the hydrophobicity of the interactions generated at this position, without strong structural requirement.

In order to consolidate these conclusions, the 4 mutants commented above were subjected to a periplasmic expression control (Fig. 4C). Three of the four enzymatic mutants were correctly expressed, but for Y371A, the halving activity seems to correlate with a decrease in expression of the same proportion.

Discussion

Based on a model of the MtPMT 3D structure, we constructed several mycobacterial reporter strains, that permitted us to develop a new quantitative high-throughput MtPMT phenotypic assay, using ConA-recognition of mannosidic patterns of a highly mannosylated reporter protein substrate. The results of this assay illustrated well the glycosylation patterns that we also detected by LC-MS, for several punctual mutants of MtPMT. Especially, we designed mutations of the MtPMT active site that permitted us not only to validate our method, but also to confirm the catalytic mechanism of protein-*O*-mannosylation, previously proposed by Bai 2019 based on structural observations of ScPMT1/2. Indeed, our results allow unprecedented functional conclusions concerning the structure/function relationship of a single PMT enzyme. We propose thus here a general catalytic mechanism of the PMT family that would be analogous to the one of inverting-glycosidases (ref) (Fig. 3F). The reaction involves at least 4 essential residues, with two aspartic acids facing each other and catalyzing an acid/base reaction. The first acidic residue (corresponding to D74 in MtPMT), belonging to EL1, would activate the hydroxyl of the acceptor threonine by deprotonation, acting thus as a base, as already proposed for mycobacterial and yeast PMTs (VanderVen, 2005; Lommel, 2011). The second acidic residue, belonging to EL2 (and corresponding to D176 in MtPMT) would act as "proton donor", to promote the departure of the leaving lipid-P group, as reported for other GT-Cs (Arabinosyltf, autre?), by inducing the inversion of the configuration of the anomeric carbon of the saccharidic unit (Fig. 3A and 3F). Besides, residues R441 and Y444 would stabilize, respectively, the phosphate of the lipid-P through ionic linkage (refs GT-C), and the acceptor threonine trough hydrophobic interaction (Bai, 2019).

In our work, we also identified a region of MtPMT EL4, harbouring mycobacteria-specific motifs, essential for PMT activity. This region seems to be involved in the recognition of substrates, according to similarities of localization with such regions of GT-Cs (ref PglB). Interestingly, this region shares 30% identity and 52% similarity with the WW domain of dystrophin, the structure of which was resolved by NMR, in complex with a peptide derived from its natural ligand (Fig. 5A). Thus, this similarity also consolidates the hypothesis of the involvement of this region in the recognition of the peptide substrate. Indeed the sequence of the peptide recognized by dystrophin (RSPPPYVP) shows stong similarities with the sequons surrounding the target mannosylation sites of MtPMT (S/T, surrounded by 4 amino acids are preferentially proline or alanine, fig. 5B, supplementary data ?). In the WW domains, the

consensus amino acids Y and W_2 (boxed in Fig. 5A) are described to form, together, a proline cradle, by hydrophobic interactions involving their carbon rings, what allows to expect that the corresponding residues of MtPMT, Y371 and L381 would act in the same way. In agreement with this analogy, the region of EL4 is located in close proximity to the peptide substrate that we have modeled into the structure of MtPMT from that of ScPMT2, (Fig. 5C and 5D), with the surface of Y371 (represented in cyan Fig. 5D), facing the proline of the peptide substrate. In addition, our functional results supported especially the role of residue Y371 in the hydrophobic binding of the peptide, since its substitution in A reduced by half activity and the one in F restored it entirely. As alternative but compatible interpretation, the substitution into alanine could also result in expression defect (Fig. 4 C), by missing stabilization through peptide binding, or even by preventing the anti-parallel fold of the two beta-strands.

In addition to an area dedicated to weakly specific binding to the peptide substrate, rather hydrophobic (shown in intense pink, Fig. 5), the EL4 region of interest also seems to include two other portions, positioned either in the membrane moiety (shown in purple, Fig. 5) or into the periplasmic compartment (shown in light pink, Fig. 5). The close-up view of the 3D representation of MtPMT in the membrane moiety allows the visualization of especially the amino acids P359 and W362, that seem to provide, together, a membrane anchoring to EL4, by double stacking against TMH10 (Fig. 5E). Remarkably, the two residues align with the N-terminal consensus amino acids of the WW domains (P and W_1, Fig 5A), described to participate in the hydrophobic support of the domain (ref). This analogy with the WW domain architecture would therefore explain why W362 is essential to anchor EL4, by likely supporting the domain of MtPMT dedicated to peptide recognition.

During this work, we were also able to identify and localize specific patterns of mycobacteria, situated especially in the loops that connect the secondary structures of the EL4 PMT-specific region (including the portion represented in light pink). However, we were not able to assign any important role to these residues in PMT activity, at least when using FasC-6xHis as reporter substrate, and alanine mutations of MtPMT. This can be explained, as for the WW domains, by a fine contribution of the intercalary loops to the peptide recognition, while these would constitute the support of the peptide-binding specificity of such domains (refs).

As our data suggested that the PMT-specific EL4 region has a defined three-dimensional conformation during catalysis, and encouraged by the similarity with a WW domain, we have also performed solution NMR analyses of a corresponding synthetic peptide (with sequence TNSAGNYHPWESKPWTWPMSLRPVLYAIDQQDVAGSGAQSSVKAEM LVGTPAM). This sequence differs from the native one by the replacement of the two cysteins into serines (in bold in the sequence), in order to avoid *in vitro* aggregation due to intermolecular cystein linkage formation. This replacement has been authorized since we demonstrate that the C381S, C386S mutations did not impact MtPMT activity, neither individually, nor simultaneously (Fig. 4B)). But disappointly, the assignment of the corresponding two-dimensional COSY spectrum of this synthetic peptide did not reveal the presence of any secondary structure (supplemental Fig.). Finally, our work demonstrated the feasibility of targeting the active site of MtPMT with molecules added into the environment of

living mycobacteria. In particular, we have found a compound which weakly but significantly inhibits the activity of this enzyme, with a logarithmic dose effect. This promising result constitutes the first proof of concept for the use of our method to screen for inhibitors of the GT-C enzyme MtPMT, in living bacteria, for the search of new anti-tuberculosis drugs.

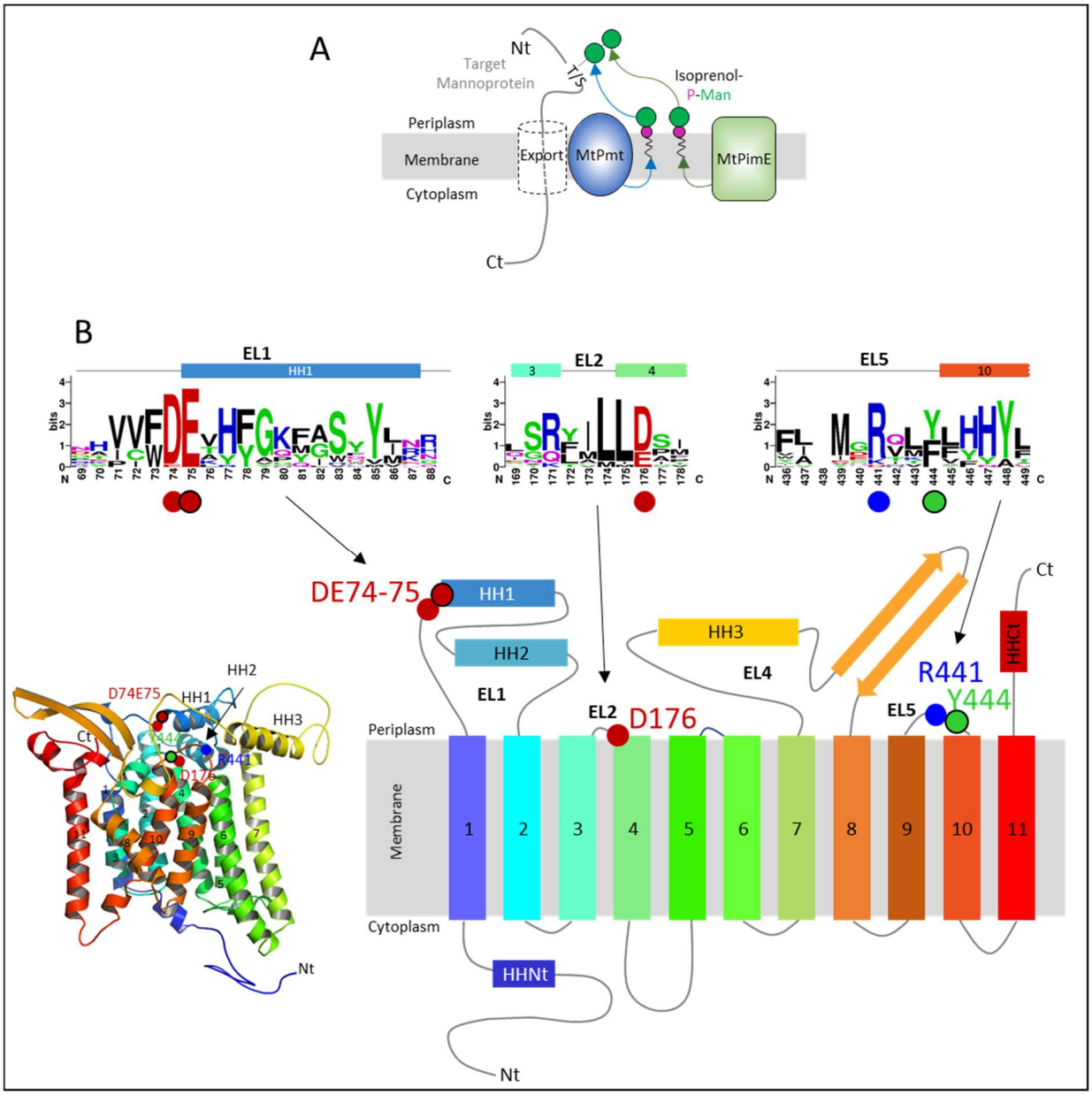


Figure 1. Molecular predictions of MtPmt structure and catalysis.

(A) MtPmt catalyzes the transfer of an initial mannose unit from isoprenol-phosphate-activated mannose (Isoprenol-P-Man) to the hydroxyl of a Threonine or a Serine residue of secreted or membrane mycobacterial mannoproteins (Tonini, 2020). Then, PimE mannosyltransferase would be responsible for the stepwise elongation of oligomannosyl appendages (Liu 2013). Analogies with eukaryote protein-O-mannosylation enables the hypothesis of a co-translocational process (VanderVen, 2005, Loibl, 2014).

(B) Structure predictions of MtPmt, based on homologies with ScPmt1/2 (Bai, 2019). Conserved areas in EL of the Pmt family are presented as logo sequences (Red, acidic; blue, basic; green, polar; black, hydrophobic and purple, neutral). Amino acids defining the active site are represented as circles (Bai, 2019) . Not-ringed circles, amino acids conserved among GT-C enzymes; ringed circles, conserved residues in the Pmt family only. Among these five conserved amino acids, to date, only the DE motif was functionally demonstrated as essential for Pmt catalysis, in mycobacteria (VanderVen, 2005) and yeast (Lommel, 2011). The 3D model and the corresponding topological diagram are presented. Numbering is based on the sequence of MtPmt. Horizontal helixes (HH), N- and C-terminal helixes (HHNt, and HHCt respectively), extracytoplasmic loops (EL). The 11 transmembrane helixes are numbered and colored in a rainbow scheme.

Mat et Met : The 3D structure prediction of MtPmt was carried out by using the Phyre2 web portal, that relied mainly on the homology with Pmt1 and Pmt2 from *Saccharomyces cerevisiae (Bai, 2019),* showing 26% and and 23% of identity with MtPmt respectively.

23 prokaryote and eukaryote members from the 3 main subfamilies Pmt1, 2, and 4 were used for the alignment, all of them recorded in Swiss-Prot, and as “characterized” in Cazy (Liu 2013, VanderVen 2005, Willer 2005, Prill 2004, Girrbach 2003, Immervoll 1995, Ichimiya 2004, Avşar-Ban 2010, Manya 2004, Lommel 2010, Lommel 2008, Manya 2006) (PMT1_SCHPO, PMT1_YEAST, PMT1_CANAL, PMT_CORGL, PMT_MYCSE, PMT_MYCTU, PMT4_CANAL, PMT4_YEAST, POMT1_MOUSE, POMT1_HUMAN, PMT4_SCHPO, POMT1_RAT, POMT1_DANRE, POMT1_DROME, POMT2_HUMAN, PMT2_YEAST, POMT2_DANRE, POMT2_DROME, PMT2_CANAL, PMT3_YEAST, PMT3_SCHPO, POMT2_DROPSPOMT2_MOUSE).

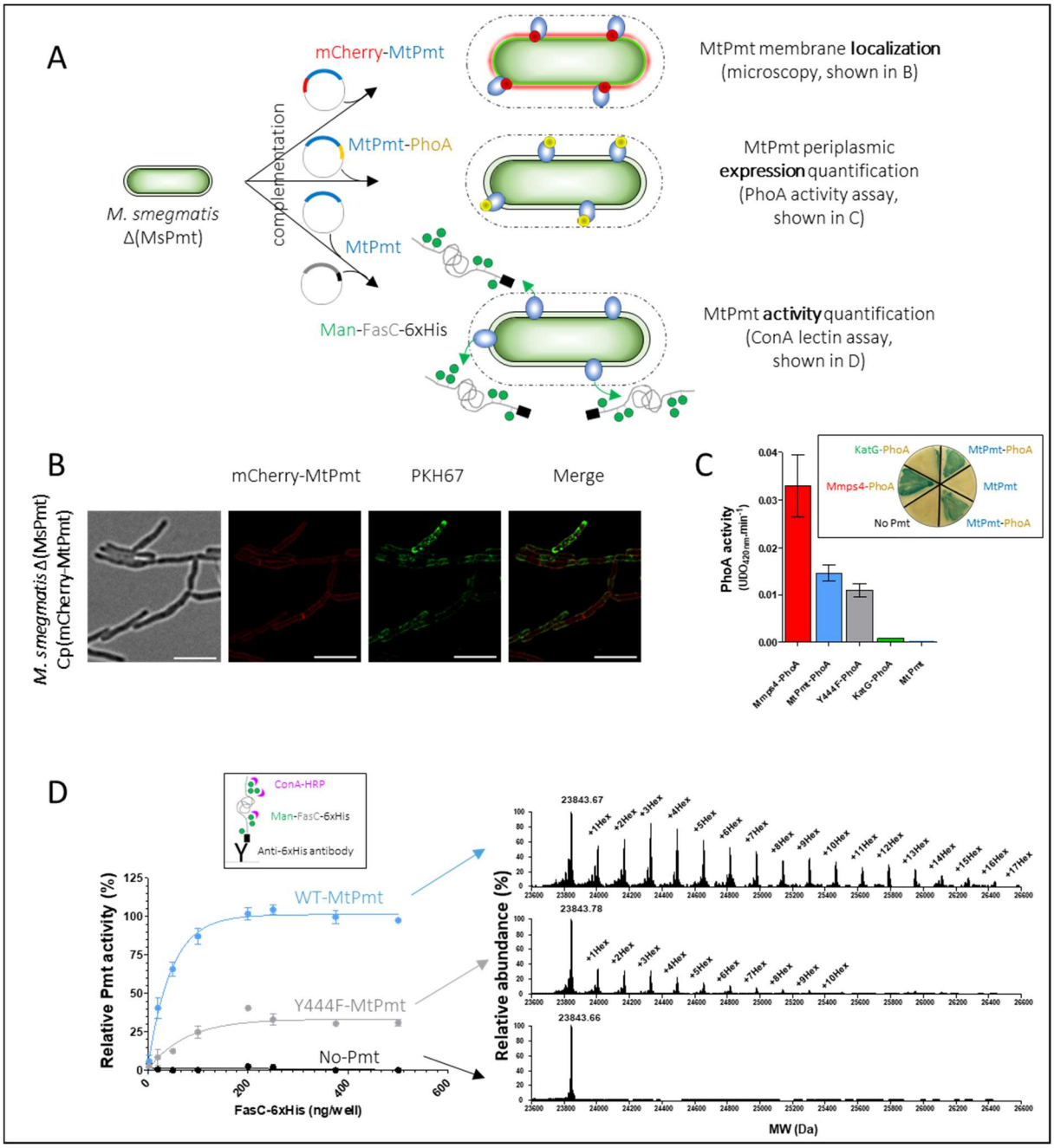


Figure 2. Validation of the MtPmt *M. smegmatis* repoter strains.

(A) *M. smegmatis* Δ(MsPmt) was complemented with single plasmids encoding mCherry-MtPmt or MtPmt-PhoA fusions, or two plasmids encoding respectively not-tagged MtPmt and FasC-6xHis. Bacteria were then cultivated during one week and processed for microscopy visualization, or the quantification of PhoA or Pmt activities. **(B)** Representative fluorescence microscopy images of *M. smegmatis* Δ(MsPmt) bacteria complemented with mCherry-MtPmt (red) and stained with PKH67 (membrane marker, green). Scale bars : 5 µm). **(C)** The histogram represents the periplasmic PhoA activity of *M. smegmatis* Δ(MsPmt) bacteria complemented withMmps4-PhoA (periplasmic control), MtPmt-PhoA, MtPmt/Y444F-PhoA (mutated on a Pmt-conserved residue from the active site), KatG-PhoA (cytoplasmic control) or not-tagged MtPmt. PhoA acitivity was measured as the increase of the 420nm absorbance over time (due to pNPP hydrolysis), in periplasmic fractions normalized to the bacterial concentration. Biological triplicates (analysed in technical triplicates) are presented. **(D)** MtPmt activity measured for three strains of *M. smegmatis* Δ(MsPmt), complemented with either WT-MtPmt, Y444F-MtPmt or No-Pmt plasmids, and the FasC-6xHis plasmid. **Left panel:** Dose-response curves obtained with the ConA lectin assay, for increasing concentrations of FasC-6xHis. Pmt activities were measured from culture supernatants as the intensity of ConA-HRP binding signals, normalized to the strongest signal obtained for MtPmt. Biological duplicates (analysed in technical triplicates) are presented. **Right panel:** Mass spectrometry analysis of the glycosylation pattern of purified entire FasC-6xHis, by LC-ESI-native-MS. Pics were normalized based on the intensity of the signal observed for the aglycone form.

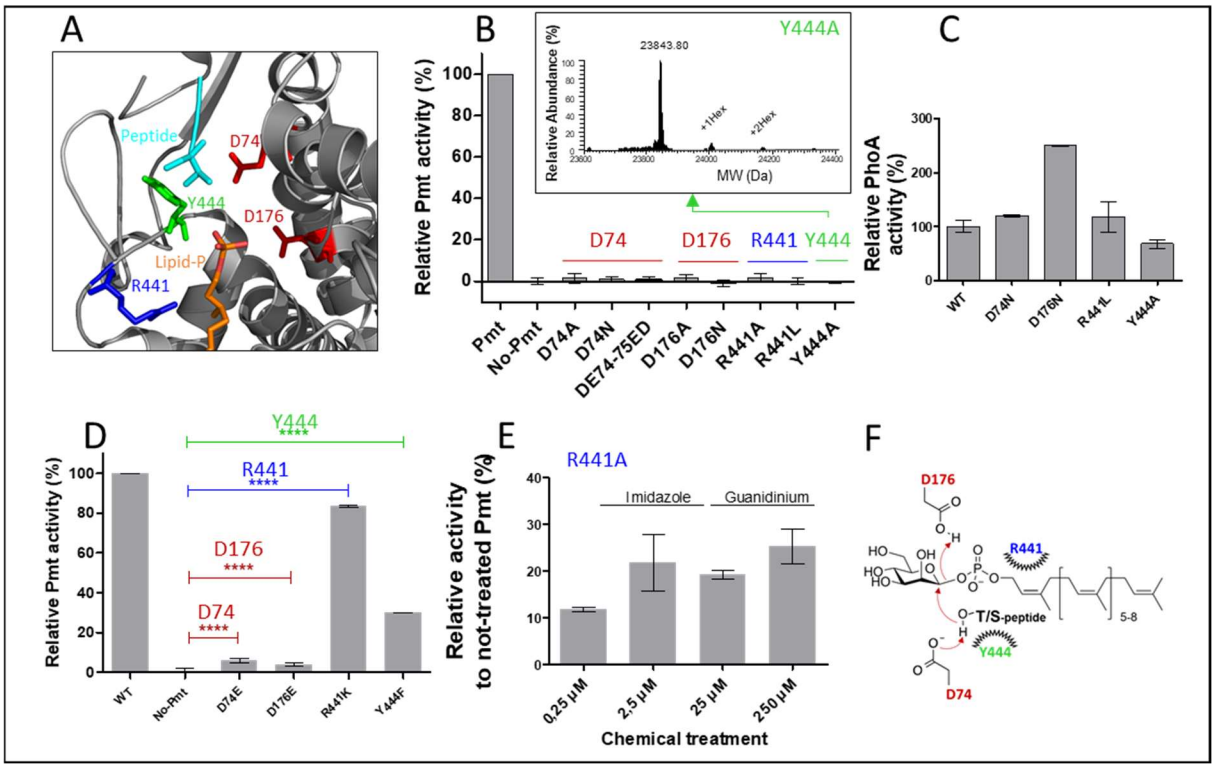


Figure 3. The active site of MtPmt is defined by residues D74, D176, R441 and Y441, essential for catalysis but no for the expression, and can be targeted by exogenous molecules in alive bacteria.

(A) Close-up view of the catalytic site of MtPmt, modeled with the Lipid-P (shown as orange sticks) and the acceptor peptide (shown as a cyan cartoon, with the Threonine-mannosylation-site represented as sticks), superimposed from the structures of ScPmt1/Pmt2 (Bai 2019). **(B)** The disruption, by site-directed mutagenesis, of biochemical functions of D74, D176, R441 and Y444 residues prevents totally Pmt catalysis in bacteria. Upper panel: the unsignificance of the glycosylation pattern of FasC-6xHis produced by the mutant Y444A (choosen as an example) was verified by mass spectrometry. **C)** Residues D74, D176, R441 and Y444 are not essential for MtPmt periplasmic expression, as their catalysis-disruptive mutation do not markedly affect the PhoA periplasmic activity in C-terminal taged fusions. **(D)** Substitutions by equivalent biochemical functions in mutants D74E, D176E, R441K and Y444F rescue partially the Pmt activity. **(E)** The chemical treatment of alive bacteria by basic exogenous compounds (imidazole or guanidinium), also rescues partially the Pmt activity of the R441A mutant. **(F)** From these structural and functional data, Pmt reaction seems to follow an analogous mechanism to inverting glycosidases, using a general acid/base catalysis to effect a direct displacement reaction, with D74 as base to activate the hydroxyl of the acceptor threonine (or serine), D176 as proton-donor, and R441 and Y444 stabilizing the phosphate of the lipid-P and the theronine (or serine), respectively. Biological duplicates, analysed in technical triplicates, are presented. The significance of the data was analysed by unpaired student's t test (**P<0,005; ***P<0,0005; ****P<0,00005). Relative Pmt activities were obtained from culture supernatants, with the ConA lectin assay (using saturating concentrations of FasC-6xHis), and normalized to the signal obtained for WT and not treated MtPmt.

Sup data :

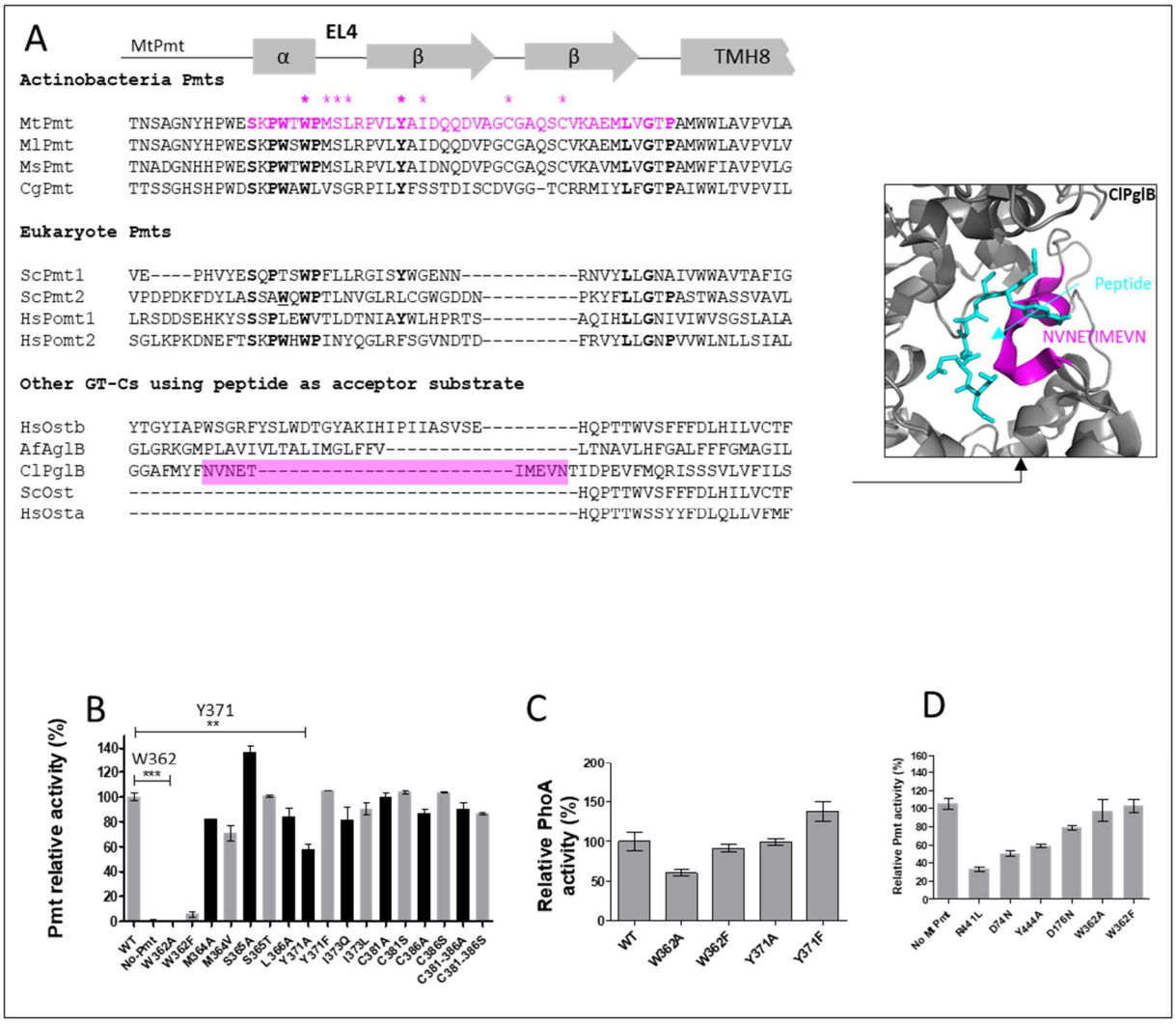


Figure 4. A region of EL4, conserved in Pmt sequences only, and possibly implicated in peptide binding, harbours some mycobacteria-specific motives and is essential for catalysis. (A) Left panel: EL4 sequence alignments of MtPmt with other Pmts and peptide-substrate-interacting GT-C enzymes. Pink letters indicate the EL4-Pmt specific region. Amino acids highly conserved among the Pmt family are represented in bold; Stars indicate the mutagenesis targets, choosen at positions conserved within the Pmt family (shown with bold stars), or specific of mycobacterial Pmts only (shown with light stars). Right panel : the corresponding region in ClPglB interacts titghtly with the substrate-peptide (Napiorkowska, 2017). **(B)** Scanning of the EL4 Pmt-specific region, by disruptive/rescueing mutagenesis in alive bacteria, for the identification of functional amino acids, important for MtPmt catalysis. Histogram bars : black, MtPmt with a catalysis-disruptive predicted mutation; grey, MtPmt WT or with a catalysis-rescueing predicted mutation. W362 and Y371 were detected as esssential or important, respectively, for Pmt activity. **(C)** Corresponding periplasmic expression in C-terminal PhoA fusions. The mutation Y371A significantly affects the periplasmic expression. **(D)** *In bacteria* attenuation of the endogenous MsPmt activity into the WT *M. smegmatis* strain, by competition with orthologous inactive mutated-MtPmt enzymes. The mutant W362F-MtPmt do not prevent MsPmt activity, in contrast to D74N-MtPmt or Y444A-MtPmt, suggesting that W362F can not compete with MsPmt for substrate binding. Biological duplicates, analysed in technical triplicates are presented.

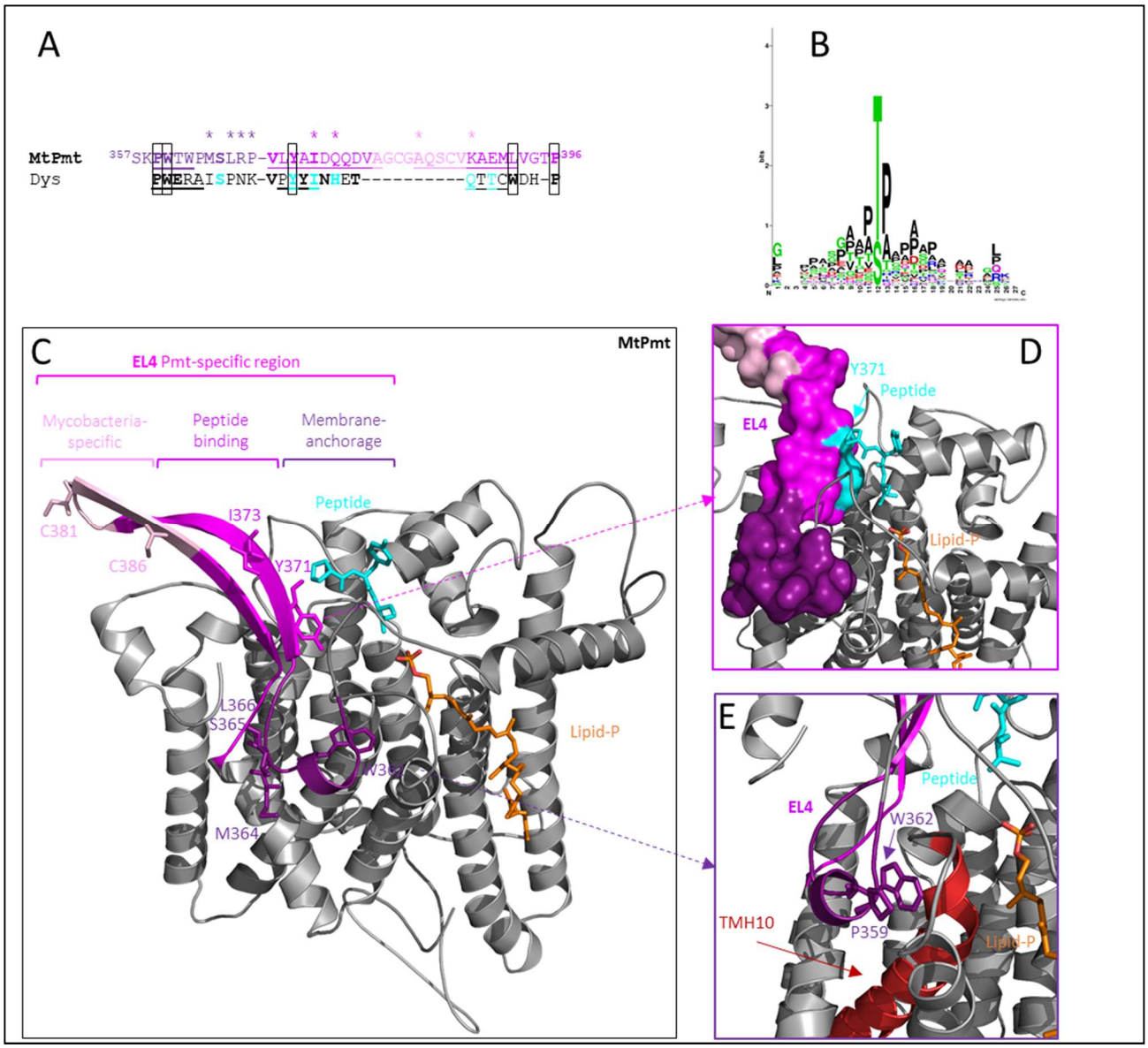
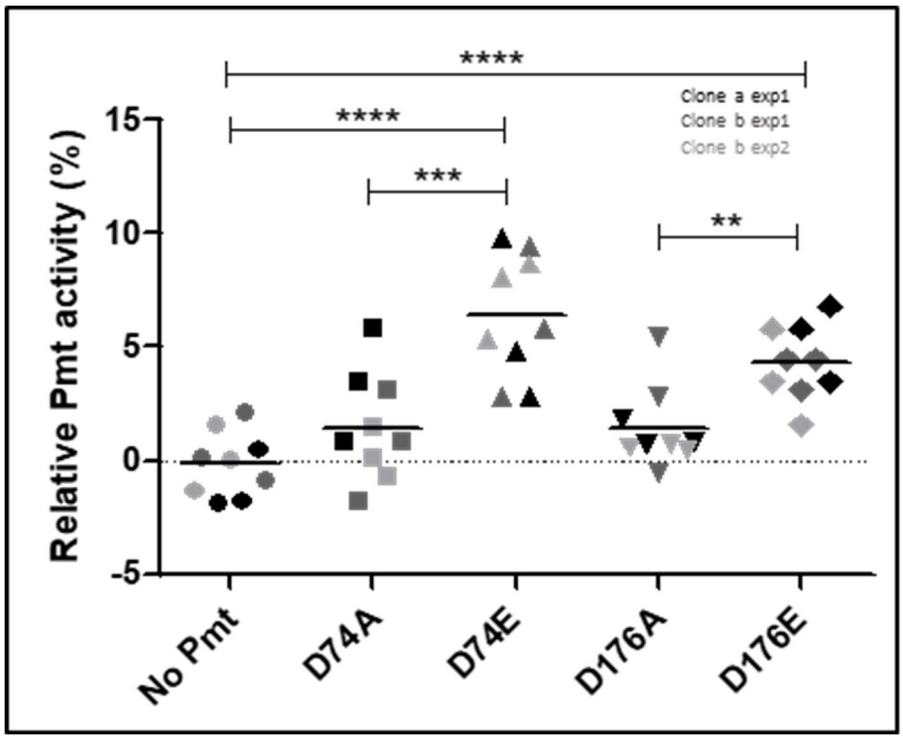


Figure 5. The mycobacteria-Pmt-specific region of EL4 is probably involved in peptide binding, and would require strong membrane anchorage. Three subsites are illustrated, predicted to be involved in membrane anchorage (purple portion), peptide binding (intense pink portion), or specificity of recognition in mycobacteria (light pink portion). **(A)** The MtPmt EL4-Pmt specific region shares significant similarity with the WW domain of Dystrophin (Dys). Stars indicate the mutagenesis targets, as in Fig. 4. Letters in bold highlight identity with dystrophin; Lines represent secondary structues; letters in cyan show dystrophin amino acids reported to interact with the peptide RSPPPYVP, derived from its natural ligand; boxes show the consensus amino acids from the WW domain family, conserved in EL4. **(B)** Logo sequence obtained from the alignment of 40 different glycopeptides of Mt. A placer éventuellement en sup data. **(C)** 3D model of MtPmt, with the Lipid-P (shown as orange sticks) and the acceptor peptide (shown as cyan sticks) superimposed from the structures of ScPmt1/Pmt2 (Bai 2019). Amino acids choosen for the mutagenesis experiments are represented in sticks and labeled. **(D)** Close-up view of the peptide-binding region of MtPmt, with the EL4 Pmt-specific region surface shown, and Y371 (well positioned for peptide interaction) represented in cyan. **(E)** Close-up view of the EL4 membrane-anchorage region, with W362, stacked by P359 to the TMH10, represented in purple sticks.

Supplemental material

Sup Data to Fig3

The significance of the data was analysed by unpaired student's t test (**P<0,005; ***P<0,0005; ****P<0,00005). Relative Pmt activities were obtained from culture supernatants, with the ConA lectin assay (using saturating concentrations of FasC-6xHis), and normalized to the signal obtained for WT and not treated MtPmt.



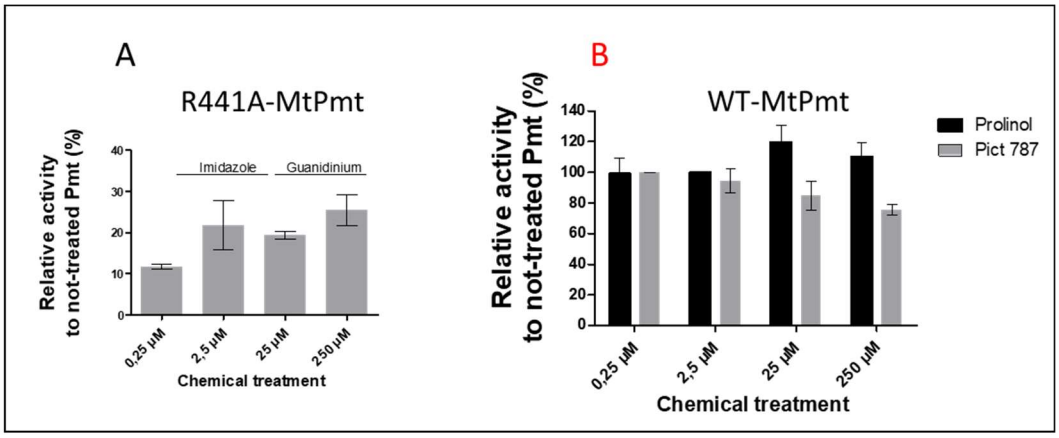
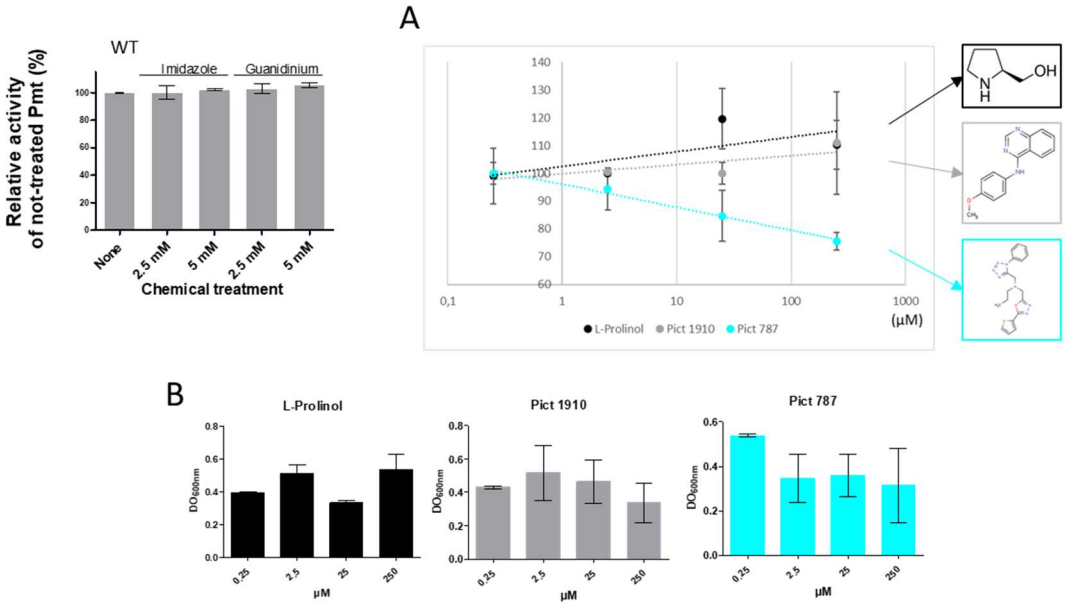


Figure 3'. The active site of MtPmt can be targeted by exogenous molecules in alive bacteria.
(A) The chemical treatment of alive bacteria by basic exogenous compounds (imidazole or guanidinium), rescued partially the activity of the R441A-MtPmt mutant. **(B)** The chemical treatment of alive bacteria by the compound Pict 787, inhibited the WT-MtPmt activity up to 25%, with dose-effect. Cette manip doit être confirmée. Biological duplicates, analysed in technical triplicates, are presented. Relative Pmt activities were obtained from culture supernatants, with the ConA lectin assay (using saturating concentrations of FasC-6xHis), and normalized to the signal obtained for WT and not treated MtPmt.

Sup data :



Supplementary figure ?. The chemical compound Pict 787 can inhibit MtPmt activity with dose effect. (A) Phenotypic assay carried on for WT-MtPmt, in the presence of various concentrations of L-Prolinol, Pict 1910, or Pict 787. Biological duplicates, analysed in technical triplicates are presented. **(B)** Corresponding cell growths, measured at DO_{600nm}.

Alignment with ClPglB :

```
MtPmt_____      41 LRGWIVTGLITLLATVTRFLNLGSLTDAGTPIF---------DEKHYAPQAWQVLNNHGV   91 (522)
                   ++..+..+++.+++...|++-......-+...+         |+-+|..-+++..++...
ClPglB_____     13 KYTCILILIAFAFSVLCRLYWVAWASEFYEFFFNDQLMITTNDGYAFAEGARDMIAGFHQ   72 (683)

MtPmt_____      92 EDNPGYGLVVHPPVGKQLIAIGEAIFGYN-GFGWRFTGALLGVVLVALVVRIVRRISRST  150 (522)
                   .+....    ++|....+.+...+++|.+ +...++.++++|+++++.+|.++|+++ ++
ClPglB_____     73 PNDLSY----FGSSLSTLTYWLYSILPFSFESIILYMSTFFASLIVVPIILIAREYK-LT  127 (683)

MtPmt_____     151 LVGAIAGVLLICDGVSFVTART--ALLDGFLTFFVVAAFGALIVDRDQVRERMHIALLAG  208 (522)
                   .+|++|++++++.|.++..++.  ..+|....+|.+++++++..++++
ClPglB_____    128 TYGFIAALLGSIANSYYNRTMSGYYDTDMLVLVLPMLILLTFIRLTINKD----------  177 (683)

MtPmt_____     209 RSAATVWGPRVGVRWWRFGAGVLLGLACATKWSGVYFVLFFGAMALAFDVAARRQYQVQR  268 (522)
                               +.+..+++++.+++...++.+......+.....+......++.....
ClPglB_____    178 -------------IFTLLLSPIFIMIYLWWYPSSYSLNFAMIGLFGLYTLVFHRKEKIFY  224 (683)

MtPmt_____     269 PWLGTVRRDVLPSGYALGLIPFAVYLATYAPWFASETAIDRHAVGQAVGRNSVVPLPDAV  328 (522)
                   ..............................                        ......
ClPglB_____    225 LAIALMIIALSMLAWQYKLALIVLLFAIFAFK----------------------EEKINF  262 (683)

MtPmt_____     329 RSLWHYTAKAFHFHAGLTNSAGNYHPWESKPWTWPMSLRPVLYAIDQQDVAGCGAQSCVK  388 (522)
                   ......................+..++|                        ..+-...
ClPglB___      263 YMIWALIFISISILHLSGGAFMYFNVNET----------------------IMEVNTI  298 (683)

MtPmt_____     389 AEMLVGTPAMWWLAVPVLAYAGWRMFVRRDWRYAVVLVGYCAGWLPWFADIDRQMYFFYA  448 (522)
                   ................+.........+|+......+.-.+...+.... ..|  |.++.
ClPglB_____    299 DPEVFMQRISSSVLVFILSFIGFILLCKDHKSMLLALPMLALGFMALRA-GLR--FTIYA  355 (683)

MtPmt_____     449 ATMAPFLVMGISLVLGDILYHPGQGSERRTLGLIVVCCYVALVVTNFAWLYPVLTGLPIS  508 (522)
                   .|.+..++-.....+....++.......+.....+....+.++........+...+-.++
ClPglB_____    356 VPVMALGFGYFLYAFFNFLEKKQIKLSLRNKNILLILIAFFSISPALMHIYYYKSSTVFT  415 (683)

MtPmt_____     509 QQTWNLEIWLPS  520 (522)
                   ++..+++.|+++
ClPglB_____    416 SYEASILNDLKN  427 (683)
```

Alignment with ScPmt1 :

```
MtPmt_____      14 DRPEEPVVPVVSPGPLVPVADFGPLDRLRGWIVTGLITLLATVTRFLNLGSLTDAGTPIF   73 (522)
                   |..|+|.||+...+|..+....+..++++.|+...+|+++++++|+++++   .|++++|
ScPmt1_____      6 DIKQGPVRPFIVTDPSAELASLRTMVTLKEKLLVACLAVFTAVIRLHGLA---WPDSVVF   62 (731)

MtPmt_____      74 DEKHYAPQAWQVLNNHGVEDNPGYGLVVHPPVGKQLIAIGEAIFGYNG------------  121 (522)
                   ||+||++++++|+++.++.|       +||||+||++|++.++||+++
ScPmt1_____      63 DEVHFGGFASQYIRGTYFMD-------VHPPLAKMLYAGVASLGGFQGDFDFENIGDSFP  115 (731)

MtPmt_____     122 -----FGWRFTGALLGVVLVALVVRIVRRISRSTLVGAIAGVLLICDGVSFVTARTALLD  176 (522)
                        +++|++++++|+++++++|.++|++++|+.+|++|++++++|+.++++||++++|
ScPmt1_____     116 STTPYVLMRFFSASLGALTVILMYMTLRYSGVRMWVALMSAICFAVENSYVTISRYILLD  175 (731)

MtPmt_____     177 GFLTFFVVAAFGALIVDRDQVRERMHIALLAGRSAATVWGPRVGVRWWRFGAGVLLGLAC  236 (522)
                   ++++||++++++++.++.+.++.                  ..++++|++++|+++|+++
ScPmt1_____     176 APLMFFIAAAVYSFKKYEMYPAN------------------SLNAYKSLLATGIALGMAS  217 (731)

MtPmt_____     237 ATKWSGVYFVLFFGAMALAFDVAARRQYQVQRPWLGTVRRDVLPSGYALGLIPFAVYLAT  296 (522)
                   ++||+|++.++.+++..+.+.+   ...+..+.......++.+.+.+.++++|+.+|+.+
ScPmt1_____     218 SSKWVGLFTVTWVGLLCIWRLW---FMIGDLTKSSKSIFKVAFAKLAFLLGVPFALYLVF  274 (731)

MtPmt_____     297 YAPWFAS-------------------------------------------ETAIDRHAV  312 (522)
                   |..|+..                                          ..+.+.|..
ScPmt1_____     275 FYIHFQSLTLDGDGASFFSPEFRSTLKNNKIPQNVVADVGIGSIISLRHLSTMGGYLHSH  334 (731)

MtPmt_____     313 GQAVGRNSVVP----------------------------------------------  323 (522)
                   ..+...+....
ScPmt1_____     335 SHNYPAGSEQQQSTLYPHMDANNDWLLELYNSLTTFQNLTDGTKVRLFHTVTRCRLHSHD  394 (731)

MtPmt_____     324 ----------------------------------------------------------  323 (522)

ScPmt1_____     395 HKPPVSESSDWQKEVSCYGYSGFDGDANDDWVVEIDKKNSAPGVAQERVIALDTKFRLRH  454 (731)

MtPmt_____     324 ----------------------------------------------------------LPD  326 (522)
                                                                              ...
ScPmt1_____     455 AMTGCYLFSHEVKLPAWGFEQQEVTCASSGRHDLTLWYVENNSNPLLPEDTKRISYKPAS  514 (731)

MtPmt_____     327 AVRSLWHYTAKAFHFHAGLTNSAGNYHPWESKPWTWPMSLRPVLYAIDQQDVAGCGAQSC  386 (522)
                   .+++++++|..|+.+|.++..    +|+|+|+||+||++.+|+.+|++..
ScPmt1_____     515 FISKFIESHKKMWHINKNLVE----PHVYESQPTSWPFLLRGISYWGENN----------  560 (731)

MtPmt_____     387 VKAEMLVGTPAMWWLAVPVLAYAGWR------------------MFVRRDWRYAVVLVGY  428 (522)
                    +.|.++|||++||.+.++++.++..                  ...+++++...+++||
ScPmt1_____     561 -RNVYLLGNAIVWWAVTAFIGIFGLIVITELFSWQLGKPILKDSKVVNFHVQVIHYLLGF  619 (731)

MtPmt_____     429 CAGWLPWFADIDRQMYFFYAATMAPFLVMGISLVLGDILYHPGQGSERRTLGLIVVCCYV  488 (522)
                   +++|+||++ +.|++|.|||+|+++|+++++|..++.+.++..+.  ++..+..++++++
ScPmt1_____     620 AVHYAPSFL-MQRQMFLHHYLPAYYFGILALGHALDIIVSYVFRS--KRQMGYAVVITFL  676 (731)

MtPmt_____     489 ALVVTNFAWLYPVLTGLPISQQTWNLEIWLPSWR  522 (522)
                   ++++++|++|+|++||+|+|+++|++|||++||.
ScPmt1_____     677 AASVYFFKSFSPIIYGTPWTQELCQKSQWLSGWD  710 (731)
```

Bibliographie

1. Global Tuberculosis Report 2020. (2020).
2. Paulson, T. Epidemiology: A mortal foe. *Nature* **502**, S2–S3 (2013).
3. WHO. *Global Tuberculosis Report 2017*. (2017). doi:WHO/HTM/TB/2017.23.
4. WHO. COVID-19 : considérations relatives à la prise en charge de la tuberculose.
5. Nunes-Alves, C. *et al.* In search of a new paradigm for protective immunity to TB. *Nat. Rev. Microbiol. 2014 124* **12**, 289–299 (2014).
6. Hershkovitz, I. *et al.* Detection and molecular characterization of 9000-year-old Mycobacterium tuberculosis from a neolithic settlement in the Eastern mediterranean. *PLoS One* (2008) doi:10.1371/journal.pone.0003426.
7. Brites, D. & Gagneux, S. Co-evolution of M ycobacterium tuberculosis and Homo sapiens. *Immunol. Rev.* **264**, 6–24 (2015).
8. Dheda, K., Barry, C. E. & Maartens, G. Tuberculosis. *Lancet* **387**, 1211–1226 (2016).
9. Vynnycki, E. & Fine, P. E. M. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol. Infect.* **119**, 183–201 (1997).
10. Cooper, A. M., Mayer-Barber, K. D. & Sher, A. Role of innate cytokines in mycobacterial infection. *Mucosal Immunol. 2011 43* **4**, 252–260 (2011).
11. Cooper, A. M. T cells in mycobacterial infection and disease. *Curr. Opin. Immunol.* **21**, 378–384 (2009).
12. Urdahl, K. B., Shafiani, S. & Ernst, J. D. Initiation and regulation of T-cell responses in tuberculosis. *Mucosal Immunol.* **4**, 288–293 (2011).
13. Gallegos, A. M., Pamer, E. G. & Glickman, M. S. Delayed protection by ESAT-6–specific effector CD4+ T cells after airborne M. tuberculosis infection. *J. Exp. Med.* **205**, 2359–2368 (2008).
14. Russell, D. G. Who puts the tubercle in tuberculosis? *Nat. Rev. Microbiol.* **5**, 39–47 (2007).
15. Pawlowski, A., Jansson, M., Sköld, M., Rottenberg, M. E. & Källenius, G. Tuberculosis and HIV Co-Infection. *PLoS Pathog.* **8**, e1002464 (2012).
16. Getahun, H., Gunneberg, C., Granich, R. & Nunn, P. HIV Infection–Associated Tuberculosis: The Epidemiology and the Response. *Clin. Infect. Dis.* **50**, S201–S207 (2010).
17. Souriant, S. *et al.* Tuberculosis Exacerbates HIV-1 Infection through IL-10/STAT3-Dependent Tunneling Nanotube Formation in Macrophages. *Cell Rep.* **26**, 3586-3599.e7 (2019).
18. Ahmed, A. *et al.* A century of BCG: Impact on tuberculosis control and beyond. *Immunol. Rev.* **301**, 98–121 (2021).
19. Pai, M. *et al.* Tuberculosis. *Nat. Rev. Dis. Prim.* **2**, 16076 (2016).
20. Justesen, U. S. *et al.* Pharmacokinetic interaction between rifampin and the combination of indinavir and low-dose ritonavir in HIV-infected patients. *Clin. Infect. Dis.* **38**, 426–9 (2004).
21. Gries, R., Sala, C. & Rybniker, J. Host-Directed Therapies and Anti-Virulence Compounds to Address Anti-Microbial Resistant Tuberculosis Infection. *Appl. Sci.* **10**, 2688 (2020).
22. Arnon, S. S., Schechter, R., Maslanka, S. E., Jewell, N. P. & Hatheway, C. L. Human botulism immune globulin for the treatment of infant botulism. *N. Engl. J. Med.* **354**, 462–71 (2006).
23. Mazumdar, S. Raxibacumab. *MAbs* **1**, 531–8 (2009).

24. Greig, S. L. Obiltoxaximab: First Global Approval. *Drugs* **76**, 823–830 (2016).
25. Forrellad, M. A. *et al.* Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* **4**, 3–66 (2013).
26. Dickey, S. W., Cheung, G. Y. C. & Otto, M. Different drugs for bad bugs: Antivirulence strategies in the age of antibiotic resistance. *Nat. Rev. Drug Discov.* **16**, 457–471 (2017).
27. Hamasur, B. *et al.* A mycobacterial lipoarabinomannan specific monoclonal antibody and its F(ab') fragment prolong survival of mice infected with Mycobacterium tuberculosis. *Clin. Exp. Immunol.* **138**, 30–8 (2004).
28. Daffé, M., Quémard, A. & Marrakchi, H. Mycolic Acids: From Chemistry to Biology. in *Biogenesis of Fatty Acids, Lipids and Membranes* (ed. Geiger, O.) 1–36 (Springer International Publishing, 2017). doi:10.1007/978-3-319-43676-0_18-1.
29. Sani, M. *et al.* Direct Visualization by Cryo-EM of the Mycobacterial Capsular Layer: A Labile Structure Containing ESX-1-Secreted Proteins. *PLoS Pathog.* **6**, e1000794 (2010).
30. Zuber, B. *et al.* Direct Visualization of the Outer Membrane of Mycobacteria and Corynebacteria in Their Native State. *J. Bacteriol.* **190**, 5672–5680 (2008).
31. Dubnau, E. *et al.* Oxygenated mycolic acids are necessary for virulence of Mycobacterium tuberculosis in mice. *Mol. Microbiol.* (2000) doi:10.1046/j.1365-2958.2000.01882.x.
32. 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).
33. Strohmeier, G. R. & Fenton, M. J. Roles of lipoarabinomannan in the pathogenesis of tuberculosis. *Microbes Infect.* **1**, 709–717 (1999).
34. Hunter, S. W. & Brennan, P. J. Evidence for the presence of a phosphatidylinositol anchor on the lipoarabinomannan and lipomannan of Mycobacterium tuberculosis. *J. Biol. Chem.* **265**, 9272–9279 (1990).
35. Briken, V., Porcelli, S. A., Besra, G. S. & Kremer, L. Mycobacterial lipoarabinomannan and related lipoglycans: From biogenesis to modulation of the immune response. *Mol. Microbiol.* **53**, 391–403 (2004).
36. Gaur, R. L. *et al.* LprG-Mediated Surface Expression of Lipoarabinomannan Is Essential for Virulence of Mycobacterium tuberculosis. *PLoS Pathog.* (2014) doi:10.1371/journal.ppat.1004376.
37. Tonini, L. *et al.* Potential Plasticity of the Mannoprotein Repertoire Associated to Mycobacterium tuberculosis Virulence Unveiled by Mass Spectrometry-Based Glycoproteomics. *Molecules* **25**, 2348 (2020).
38. Smith, G. T., Sweredoski, M. J. & Hess, S. O-linked glycosylation sites profiling in Mycobacterium tuberculosis culture filtrate proteins. *J. Proteomics* **97**, 296–306 (2014).
39. Romain, F., Augier, J., Pescher, P. & Marchal, G. Isolation of a proline-rich mycobacterial protein eliciting delayed-type hypersensitivity reactions only in guinea pigs immunized with living mycobacteria. *Proc. Natl. Acad. Sci.* **90**, 5322–5326 (1993).
40. Espitia, C. *et al.* Antigenic and structural similarities between Mycobacterium tuberculosis 50- to 55-kilodalton and Mycobacterium bovis BCG 45- to 47-kilodalton antigens. *Infect. Immun.* **63**, 580–584 (1995).
41. Romain, F. *et al.* Deglycosylation of the 45/47-kilodalton antigen complex of Mycobacterium tuberculosis decreases its capacity to elicit in vivo or in vitro cellular immune responses. *Infect. Immun.* (1999).
42. Horn, C. *et al.* Decreased capacity of recombinant 45/47-kDa molecules (Apa) of Mycobacterium tuberculosis to stimulate T lymphocyte responses related to changes

in their mannosylation pattern. *J. Biol. Chem.* **274**, 32023–32030 (1999).
43. Pitarque, S. *et al.* Deciphering the molecular bases of *Mycobacterium tuberculosis* binding to the lectin DC-SIGN reveals an underestimated complexity. *Biochem. J.* (2005) doi:10.1042/BJ20050709.
44. Ragas, A., Roussel, L., Puzo, G. & Rivière, M. The Mycobacterium tuberculosis cell-surface glycoprotein apa as a potential adhesin to colonize target cells via the innate immune system pulmonary C-type lectin surfactant protein A. *J. Biol. Chem.* **282**, 5133–5142 (2007).
45. Hewinson, R. G., Michell, S. L., Russell, W. P., McAdam, R. A. & Jacobs, W. R. Molecular characterization of MPT83: a seroreactive antigen of Mycobacterium tuberculosis with homology to MPT70. *Scand. J. Immunol.* **43**, 490–9 (1996).
46. Diaz-Silvestre, H. *et al.* The 19-kDa antigen of Mycobacterium tuberculosis is a major adhesin that binds the mannose receptor of THP-1 monocytic cells and promotes phagocytosis of mycobacteria. *Microb. Pathog.* (2005) doi:10.1016/j.micpath.2005.06.002.
47. Parra, J. *et al.* Scrutiny of Mycobacterium tuberculosis 19 kDa antigen proteoforms provides new insights in the lipoglycoprotein biogenesis paradigm. *Sci. Rep.* **7**, 43682 (2017).
48. Sartain, M. J. & Belisle, J. T. N-Terminal clustering of the O-glycosylation sites in the Mycobacterium tuberculosis lipoprotein SodC. *Glycobiology* **19**, 38 (2009).
49. Herrmann, J. L., Delahay, R., Gallagher, A., Robertson, B. & Young, D. Analysis of post-translational modification of mycobacterial proteins using a cassette expression system. *FEBS Lett.* **473**, 358–362 (2000).
50. Alonso, H. *et al.* Protein O-mannosylation deficiency increases LprG-associated lipoarabinomannan release by Mycobacterium tuberculosis and enhances the TLR2-associated inflammatory response. *Sci. Rep.* (2017) doi:10.1038/s41598-017-08489-7.
51. Esparza, M. *et al.* PstS-1, the 38-kDa Mycobacterium tuberculosis glycoprotein, is an adhesin, which binds the macrophage mannose receptor and promotes phagocytosis. *Scand. J. Immunol.* **81**, 46–55 (2015).
52. Liu, C.-F. *et al.* Bacterial protein-O-mannosylating enzyme is crucial for virulence of Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6560–5 (2013).
53. Girrbach, V., Zeller, T., Priesmeier, M. & Strahl-Bolsinger, S. Structure-function analysis of the dolichyl phosphate-mannose: Protein O-mannosyltransferase ScPmt1p. *J. Biol. Chem.* **275**, 19288–19296 (2000).
54. Lommel, M., Schott, A., Jank, T., Hofmann, V. & Strahl, S. A Conserved Acidic Motif Is Crucial for Enzymatic Activity of Protein O -Mannosyltransferases. *J. Biol. Chem.* **286**, 39768–39775 (2011).
55. VanderVen, B. C., Harder, J. D., Crick, D. C. & Belisle, J. T. Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways. *Science* **309**, 941–3 (2005).
56. Van Reeuwijk, J. *et al.* The Expanding Phenotype of POMT1 Mutations: From Walker-Warburg Syndrome to Congenital Muscular Dystrophy, Microcephaly, and Mental Retardation. *Hum. Mutat.* **27**, 453–459 (2006).
57. Richards, J. B. & Hemming, F. W. *The Transfer of Mannose from Guanosine Diphosphate Mannose to Dolichol Phosphate and Protein by Pig Liver Endoplasmic Reticulum*. *Biochem. J* vol. 130 (1972).
58. Babczinski, P. & Tanner, W. Involvement of dolicholmonophosphate in the formation of specific mannosyl-linkages in yeast glycoproteins. *Biochem. Biophys. Res. Commun.* **54**, 1119–24 (1973).

59. Mahne, M., Tauch, A., Pühler, A. & Kalinowski, J. The Corynebacterium glutamicum gene pmt encoding a glycosyltransferase related to eukaryotic protein-O-mannosyltransferases is essential for glycosylation of the resuscitation promoting factor (Rpf2) and other secreted proteins. *FEMS Microbiol. Lett.* (2006) doi:10.1111/j.1574-6968.2006.00269.x.
60. Howlett, R. *et al.* Streptomyces coelicolor strains lacking polyprenol phosphate mannose synthase and protein O-mannosyl transferase are hyper-susceptible to multiple antibiotics. *Microbiol. (United Kingdom)* **164**, 369–382 (2018).
61. Halim, A. *et al.* Discovery of a nucleocytoplasmic O-mannose glycoproteome in yeast. *Proc. Natl. Acad. Sci.* (2015) doi:10.1073/pnas.1511743112.
62. Larsen, I. S. B. *et al.* Mammalian O-mannosylation of cadherins and plexins is independent of protein O-mannosyltransferases 1 and 2. *J. Biol. Chem.* (2017) doi:10.1074/jbc.M117.794487.
63. Servier Medical Art.
64. Joshi, H. J. *et al.* SnapShot: O-Glycosylation Pathways across Kingdoms. *Cell* **172**, 632-632.e2 (2018).
65. Chiba, A. *et al.* Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve α-dystroglycan. The role of a novel O-mannosyl-type oligosaccharide in the binding of α-dystroglycan with laminin. *J. Biol. Chem.* **272**, 2156–2162 (1997).
66. Endo, T. & Manya, H. O-mannosylation in mammalian cells. *Methods Mol. Biol.* **347**, 43–56 (2006).
67. Lussier, M., Sdicu, A.-M. & Bussey, H. The KTR and MNN1 mannosyltransferase families of Saccharomyces cerevisiae. *Biochim. Biophys. Acta - Gen. Subj.* **1426**, 323–334 (1999).
68. Romero, P. A. *et al.* Mnt2p and Mnt3p of Saccharomyces cerevisiae are members of the Mnn1p family of -1,3-mannosyltransferases responsible for adding the terminal mannose residues of O-linked oligosaccharides. *Glycobiology* **9**, 1045–1051 (1999).
69. Graham, T. R. & Emr, S. D. Compartmental organization of Golgi-specific protein modification and vacuolar protein sorting events defined in a yeast sec18 (NSF) mutant. *J. Cell Biol.* **114**, 207–218 (1991).
70. Kanagawa, M. & Toda, T. Ribitol-phosphate—a newly identified posttranslational glycosylation unit in mammals: structure, modification enzymes and relationship to human diseases. *J. Biochem.* **163**, 359–369 (2018).
71. Moore, C. J. & Hewitt, J. E. Dystroglycan glycosylation and muscular dystrophy. *Glycoconj. J.* **26**, 349–357 (2009).
72. Goto, M. Protein O -Glycosylation in Fungi: Diverse Structures and Multiple Functions. *Biosci. Biotechnol. Biochem.* **71**, 1415–1427 (2007).
73. Ecker, M. *et al.* O -Mannosylation precedes and potentially controls the N -glycosylation of a yeast cell wall glycoprotein. *EMBO Rep.* **4**, 628–632 (2003).
74. González, M., Brito, N. & González, C. High abundance of Serine/Threonine-rich regions predicted to be hyper-O-glycosylated in the secretory proteins coded by eight fungal genomes. *BMC Microbiol.* **12**, 213 (2012).
75. Cooper, H. N. *et al.* Characterization of mycobacterial protein glycosyltransferase activity using synthetic peptide acceptors in a cell-free assay. *Glycobiology* **12**, 427–434 (2002).
76. Mealer, R. G. *et al.* Glycobiology and schizophrenia: a biological hypothesis emerging from genomic research. *Mol. Psychiatry* **25**, 3129–3139 (2020).
77. Jerber, J. *et al.* Biallelic Mutations in TMTC3, Encoding a Transmembrane and TPR-Containing Protein, Lead to Cobblestone Lissencephaly. *Am. J. Hum. Genet.* **99**, 1181–

1189 (2016).
78. Farhan, S. M. K. *et al.* Identification of a novel synaptic protein, TMTC3, involved in periventricular nodular heterotopia with intellectual disability and epilepsy. *Hum. Mol. Genet.* **26**, 4278–4289 (2017).
79. Willer, T. *et al.* Targeted disruption of the Walker-Warburg syndrome gene Pomt1 in mouse results in embryonic lethality. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 14126–14131 (2004).
80. Uribe, M. L. *et al.* Retinal Proteomics of a Mouse Model of Dystroglycanopathies Reveals Molecular Alterations in Photoreceptors. *J. Proteome Res.* (2021) doi:10.1021/acs.jproteome.1c00126.
81. Ichimiya, T. *et al.* The twisted abdomen phenotype of Drosophila POMT1 and POMT2 mutants coincides with their heterophilic protein O-mannosyltransferase activity. *J. Biol. Chem.* **279**, 42638–42647 (2004).
82. Strahl-Bolsinger, S., Immervoll, T., Deutzmann, R. & Tanner, W. PMT1, the gene for a key enzyme of protein O-glycosylation in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8164–8168 (1993).
83. Loibl, M. & Strahl, S. Protein O-mannosylation: What we have learned from baker's yeast. *Biochimica et Biophysica Acta - Molecular Cell Research* vol. 1833 (2013).
84. Rouabhia, M. *et al.* Virulence of the Fungal Pathogen Candida albicans Requires the Five Isoforms of Protein Mannosyltransferases. *Infect. Immun.* **73**, 4571–4580 (2005).
85. Olson, G. M., Fox, D. S., Wang, P., Alspaugh, J. A. & Buchanan, K. L. Role of protein O-mannosyltransferase Pmt4 in the morphogenesis and virulence of Cryptococcus neoformans. *Eukaryot. Cell* (2007) doi:10.1128/EC.00182-06.
86. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
87. Chiapparino, A. *et al.* Functional implications of MIR domains in protein O-mannosylation. *Elife* **9**, (2020).
88. Bai, L., Kovach, A., You, Q., Kenny, A. & Li, H. Structure of the eukaryotic protein O-mannosyltransferase Pmt1−Pmt2 complex. *Nat. Struct. Mol. Biol.* **26**, 704–711 (2019).
89. Paschinger, K., Yan, S. & Wilson, I. B. H. N-glycomic Complexity in Anatomical Simplicity: Caenorhabditis elegans as a Non-model Nematode? *Front. Mol. Biosci.* **0**, 9 (2019).
90. Jurado, L. A. P., Coloma, A. & Cruces, J. Identification of a Human Homolog of the Drosophila rotated abdomen Gene (POMT1) Encoding a Putative Protein O-Mannosyl-Transferase, and Assignment to Human Chromosome 9q34.1. *Genomics* **58**, 171–180 (1999).
91. Immervoll, T., Gentzsch, M. & Tanner, W. PMT3 andPMT4, two new members of the protein-O-mannosyltransferase gene family ofSaccharomyces cerevisiae. *Yeast* **11**, 1345–1351 (1995).
92. Gentzsch, M. & Tanner, W. Protein-O-glycosylation in yeast: protein-specific mannosyltransferases. **7**, (1997).
93. Strahl-Bolsinger, S. & Tanner, W. Protein O-glycosylation in Saccharomyces cerevisiae. *Eur. J. Biochem.* **190**, 185–190 (1991).
94. Prill, S. K.-H. *et al.* PMT family of Candida albicans: five protein mannosyltransferase isoforms affect growth, morphogenesis and antifungal resistance. *Mol. Microbiol.* **55**, 546–560 (2004).
95. Lommel, M., Willer, T., Cruces, J. & Strahl, S. POMT1 is Essential for Protein O-Mannosylation in Mammals. in 323–342 (2010). doi:10.1016/S0076-6879(10)79018-2.
96. Lommel, M., Willer, T. & Strahl, S. POMT2, a key enzyme in Walker–Warburg syndrome:

somatic sPOMT2, but not testis-specific tPOMT2, is crucial for mannosyltransferase activity in vivo. *Glycobiology* **18**, 615–625 (2008).
97. Manya, H. *et al.* Demonstration of mammalian protein O-mannosyltransferase activity: Coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc. Natl. Acad. Sci.* **101**, 500–505 (2004).
98. Sanders, S. L., Gentzsch, M., Tanner, W. & Herskowitz, I. O-glycosylation of Ax12/Bud10p by Pmt4p is required for its stability, localization, and function in daughter cells. *J. Cell Biol.* **145**, 1177–1188 (1999).
99. Pejenaute-Ochoa, M. D., Santana-Molina, C., Devos, D. P., Ibeas, J. I. & Fernández-álvarez, A. Structural, evolutionary, and functional analysis of the protein o-mannosyltransferase family in pathogenic fungi. *J. Fungi* **7**, (2021).
100. Willer, T., Amselgruber, W., Deutzmann, R. & Strahl, S. Characterization of POMT2, a novel member of the PMT protein O-mannosyltransferase family specifically localized to the acrosome of mammalian spermatids. *Glycobiology* **12**, 771–783 (2002).
101. Girrbach, V. & Strahl, S. Members of the evolutionarily conserved PMT family of protein O-mannosyltransferases form distinct protein complexes among themselves. *J. Biol. Chem.* **278**, 12554–62 (2003).
102. Kriangkripipat, T. & Momany, M. Aspergillus nidulans Pmts form heterodimers in all pairwise combinations. *FEBS Open Bio* **4**, 335–341 (2014).
103. Schultz, J. C. & Takayama, K. The role of mannosylphosphorylpolyisoprenol in glycoprotein biosynthesis in Mycobacterium smegmatis. *BBA - Gen. Subj.* **381**, 175–184 (1975).
104. Jia, L., Sha, S., Yang, S., Taj, A. & Ma, Y. Effect of Protein O-Mannosyltransferase (MSMEG_5447) on M. smegmatis and Its Survival in Macrophages. *Front. Microbiol.* **12**, (2021).
105. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 (2003).
106. Willer, T., Brandl, M., Sipiczki, M. & Strahl, S. Protein O-mannosylation is crucial for cell wall integrity, septation and viability in fission yeast. *Mol. Microbiol.* **57**, 156–170 (2005).
107. Manya, H., Chiba, A., Margolis, R. U. & Endo, T. Molecular cloning and characterization of rat Pomt1 and Pomt2. *Glycobiology* **16**, 863–873 (2006).
108. Avsar-Ban, E. *et al.* Protein O-mannosylation is necessary for normal embryonic development in zebrafish. *Glycobiology* **20**, 1089–1102 (2010).
109. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
110. Dereeper, A. *et al.* Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469 (2008).
111. Becker, K. *et al.* Lipoprotein glycosylation by protein-O-mannosyltransferase (MAB_1122c) contributes to low cell envelope permeability and antibiotic resistance of Mycobacterium abscessus. *Front. Microbiol.* **8**, 2123 (2017).
112. Wehmeier, S. *et al.* Glycosylation of the phosphate binding protein, PstS, in Streptomyces coelicolor by a pathway that resembles protein O-mannosylation in eukaryotes. *Mol. Microbiol.* **71**, 421–433 (2009).
113. Garbe, T. *et al.* Expression of the Mycobacterium tuberculosis 19-kilodalton antigen in Mycobacterium smegmatis: immunological analysis and evidence of glycosylation. *Infect. Immun.* **61**, 260–7 (1993).
114. Pesnot, T., Jørgensen, R., Palcic, M. M. & Wagner, G. K. Structural and mechanistic basis

for a new mode of glycosyltransferase inhibition. *Nat. Chem. Biol.* **6**, 321–3 (2010).
115. Vrielink, A., Rüger, W., Driessen, H. P. C. & Freemont, P. S. Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* **13**, 3413–22 (1994).
116. Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350–356 (2011).
117. Lommel, M. *et al.* Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proc. Natl. Acad. Sci.* **110**, 21024–21029 (2013).
118. Orchard, M. G. *et al.* Rhodanine-3-acetic acid derivatives as inhibitors of fungal protein mannosyl transferase 1 (PMT1). *Bioorg. Med. Chem. Lett.* **14**, 3975–3978 (2004).
119. Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **326 (Pt 3**, 929–39 (1997).
120. Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328**, 307–317 (2003).
121. Zhang, L. *et al.* Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol. *Science (80-.).* **9102**, eaba9102 (2020).
122. Tan, Y. Z. *et al.* Cryo-EM Structures and Regulation of Arabinofuranosyltransferase AftD from Mycobacteria. *Mol. Cell* **78**, 683-699.e11 (2020).
123. Bloch, J. S. *et al.* Structure and mechanism of the ER-based glucosyltransferase ALG6. *Nature* **579**, 443–447 (2020).
124. Matsumoto, S. *et al.* Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17868–73 (2013).
125. Wild, R. *et al.* Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* 1–12 (2018) doi:10.1126/science.aar5140.
126. Braunger, K. *et al.* Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science (80-.).* **360**, 215–219 (2018).
127. Ramírez, A. S., Kowal, J. & Locher, K. P. Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science (80-.).* **366**, 1372–1375 (2019).
128. Petrou, V. I. *et al.* Structures of aminoarabinose transferase ArnT suggest a molecular basis for lipid A glycosylation. *Science (80-.).* **351**, 608–612 (2016).
129. Lovering, A. L., de Castro, L. H., Lim, D. & Strynadka, N. C. J. Structural Insight into the Transglycosylation Step of Bacterial Cell-Wall Biosynthesis. *Science (80-.).* **315**, 1402–1405 (2007).
130. Kim, H., Von Heijne, G. & Nilsson, I. M. Membrane topology of the STT3 subunit of the oligosaccharyl transferase complex. *J. Biol. Chem.* **280**, 20261–20267 (2005).
131. Strahl-Bolsinger, S. & Scheinost, A. Transmembrane Topology of Pmt1p, a Member of an Evolutionarily Conserved Family of Protein O-Mannosyltransferases. *J. Biol. Chem.* **274**, 9068–9075 (1999).
132. Bai, L. & Li, H. Cryo-EM is uncovering the mechanism of eukaryotic protein N-glycosylation. *FEBS J.* **286**, 1638–1644 (2019).
133. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
134. Napiórkowska, M. *et al.* Molecular basis of lipid-linked oligosaccharide recognition and processing by bacterial oligosaccharyltransferase. *Nat. Struct. Mol. Biol.* **24**, 1100–1106 (2017).

135. Gerber, S. *et al.* Mechanism of Bacterial Oligosaccharyltransferase. *J. Biol. Chem.* **288**, 8849–8861 (2013).
136. Steentoft, C. *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488 (2013).
137. Tonini, L. Développement de nouvelles stratégies protéomiques pour l'analyse de glycoprotéines mycobacteriennes. (UT3 Paul Sabatier, 2014).
138. Morita, Y. S. *et al.* PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria. *J. Biol. Chem.* **281**, 25143–25155 (2006).
139. Zatorska, E. *et al.* Cellular Consequences of Diminished Protein O-Mannosyltransferase Activity in Baker's Yeast. *Int. J. Mol. Sci.* **18**, 1226 (2017).
140. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1760–5 (2009).
141. Kelkar, D. S. *et al.* Proteogenomic analysis of Mycobacterium tuberculosis by high resolution mass spectrometry. *Mol. Cell. Proteomics* **10**, M111.011627 (2011).
142. Schubert, O. T. *et al.* The Mtb Proteome Library: A Resource of Assays to Quantify the Complete Proteome of Mycobacterium tuberculosis. *Cell Host Microbe* **13**, 602–612 (2013).
143. Desiere, F. The PeptideAtlas project. *Nucleic Acids Res.* **34**, D655–D658 (2006).
144. Belardinelli, J. M. & Jackson, M. Green Fluorescent Protein as a protein localization and topological reporter in mycobacteria. *Tuberculosis* **105**, 13–17 (2017).
145. Aronson, D. E., Costantini, L. M. & Snapp, E. L. Superfolder GFP Is Fluorescent in Oxidizing Environments When Targeted via the Sec Translocon. *Traffic* **12**, 543–548 (2011).
146. Meiresonne, N. Y. *et al.* Superfolder mTurquoise2 ox optimized for the bacterial periplasm allows high efficiency in vivo FRET of cell division antibiotic targets. *Mol. Microbiol.* **111**, 1025–1038 (2019).
147. Dammeyer, T. & Tinnefeld, P. Engineered fluorescent proteins illuminate the bacterial periplasm. *Comput. Struct. Biotechnol. J.* **3**, e201210013 (2012).
148. Wells, R. M. *et al.* Discovery of a Siderophore Export System Essential for Virulence of Mycobacterium tuberculosis. *PLoS Pathog.* **9**, (2013).
149. Braunstein, M. *et al.* Identification of genes encoding exported Mycobacterium tuberculosis proteins using a Tn552'phoA in vitro transposition system. *J. Bacteriol.* **182**, 2732–2740 (2000).
150. Drew, D. *et al.* Rapid topology mapping of Escherichia coli inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 2690–2695 (2002).
151. Sone, M., Kishigami, S., Yoshihisa, T. & Ito, K. Roles of disulfide bonds in bacterial alkaline phosphatase. *J. Biol. Chem.* **272**, 6174–6178 (1997).
152. Wang, X., Dutton, R. J., Beckwith, J. & Boyd, D. Membrane topology and mutational analysis of mycobacterium tuberculosis VKOR, a protein involved in disulfide bond formation and a homologue of human vitamin K epoxide reductase. *Antioxidants Redox Signal.* **14**, 1413–1420 (2011).
153. Sigrist, C. J. A. *et al.* ProRule: a new database containing functional and structural information on PROSITE profiles. *Bioinformatics* **21**, 4060–4066 (2005).
154. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

155. Yang, J. *et al.* The I-TASSER suite: Protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2014).
156. Yang, J. & Zhang, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).
157. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
158. Song, Y. *et al.* High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013).
159. Raman, S. *et al.* Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins Struct. Funct. Bioinforma.* **77**, 89–99 (2009).
160. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
161. Adasme, M. F. *et al.* PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.* **49**, W530–W534 (2021).
162. Liu, Y. *et al.* CB-Dock: a web server for cavity detection-guided protein–ligand blind docking. *Acta Pharmacol. Sin.* **41**, 138–144 (2020).
163. Hudson, K. L. *et al.* Carbohydrate–Aromatic Interactions in Proteins. *J. Am. Chem. Soc.* **137**, 15152–15160 (2015).
164. Spiwok, V. CH/π Interactions in Carbohydrate Recognition. *Molecules* **22**, 1038 (2017).
165. Cao, Y. & Li, L. Improved protein-ligand binding affinity prediction by using a curvature-dependent surface-area model. *Bioinformatics* **30**, 1674–80 (2014).
166. London, N., Raveh, B., Cohen, E., Fathi, G. & Schueler-furman, O. Rosetta FlexPepDock web server — high resolution modeling of peptide – protein interactions. **39**, 249–253 (2011).
167. Raveh, B., London, N. & Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* **78**, 2029–40 (2010).
168. Marshall, R. D. Glycoproteins. *Annu. Rev. Biochem.* **41**, 673–702 (1972).
169. Krieg, J. *et al.* Recognition signal for C-mannosylation of Trp-7 in RNase 2 consists of sequence Trp-x-x-Trp. *Mol. Biol. Cell* **9**, 301–309 (1998).
170. Pedebos, C., Arantes, P. R., Giesel, G. M. & Verli, H. In silico Investigation of the PglB Active Site Reveals Transient Catalytic States and Octahedral Metal Ion Coordination. *Glycobiology* **25**, 1183–1195 (2015).
171. Perona, J. J. *et al.* Exogenous Acetate Reconstitutes the Enzymatic Activity of Trypsin Asp189Ser. *Biochemistry* **33**, 3252–3259 (1994).
172. Monegal, A. & Planas, A. Chemical Rescue of α3-Galactosyltransferase. Implications in the Mechanism of Retaining Glycosyltransferases. *J. Am. Chem. Soc.* **128**, 16030–16031 (2006).
173. Qiao, Y., Molina, H., Pandey, A., Zhang, J. & Cole, P. A. Chemical rescue of a mutant enzyme in living cells. *Science (80-.).* **311**, 1293–1297 (2006).
174. Tu, C., Rowlett, R. S., Tripp, B. C., Ferry, J. G. & Silverman, D. N. Chemical rescue of proton transfer in catalysis by carbonic anhydrases in the ??- and ??-class. *Biochemistry* **41**, 15429–15435 (2002).
175. Newport, T. D., Sansom, M. S. P. & Stansfeld, P. J. The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res.* **47**, D390–D397 (2019).
176. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–34 (1995).
177. Lizak, C. *et al.* A catalytically essential motif in external loop 5 of the bacterial

oligosaccharyltransferase PglB. *J. Biol. Chem.* (2014) doi:10.1074/jbc.M113.524751.
178. Schulz, B. L. *et al.* Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc. Natl. Acad. Sci.* **106**, 11061–11066 (2009).
179. Huang, X. *et al.* Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat. Struct. Biol.* **7**, 634–8 (2000).
180. Aragón, E. *et al.* A Smad action turnover switch operated by WW domain readers of a phosphoserine code. *Genes Dev.* **25**, 1275–88 (2011).
181. Sudol, M., Chen, H. I., Bougeret, C., Einbond, A. & Bork, P. Characterization of a novel protein-binding module - the WW domain. *FEBS Lett.* **369**, 67–71 (1995).
182. Macias, M. J., Gervais, V., Civera, C. & Oschkinat, H. Structural analysis of WW domains and design of a WW prototype. *Nat. Struct. Biol.* **7**, 375–379 (2000).
183. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L. & Serrano, L. Thermodynamic and Kinetic Analysis of the SH3 Domain of Spectrin Shows a Two-State Folding Transition. *Biochemistry* **33**, 2142–2150 (1994).
184. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).
185. Lengeler, K. B., Tielker, D. & Ernst, J. F. Protein-O-mannosyltransferases in virulence and development. *Cell. Mol. Life Sci.* **65**, 528–544 (2008).
186. Arroyo, J. *et al.* Functional and genomic analyses of blocked protein O-mannosylation in baker's yeast. *Mol. Microbiol.* **79**, 1529–1546 (2011).
187. Szafranski-Schneider, E. *et al.* Msb2 shedding protects Candida albicans against antimicrobial peptides. *PLoS Pathog.* **8**, (2012).
188. Argyros, R. *et al.* A Phenylalanine to Serine Substitution within an O-Protein Mannosyltransferase Led to Strong Resistance to PMT-Inhibitors in Pichia pastoris. *PLoS One* (2013) doi:10.1371/journal.pone.0062229.
189. Dorman, G. & Prestwich, G. D. Benzophenone Photophores in Biochemistry. *Biochemistry* **33**, 5661–5673 (1994).
190. Kauer, J. C., Erickson-Viitanen, S., Wolfe, H. R. & DeGrado, W. F. p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. *J. Biol. Chem.* **261**, 10695–700 (1986).
191. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–44 (2010).
192. Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the Genetic Code of Escherichia coli. *Science (80-.).* **292**, 498–500 (2001).
193. Chin, J. W. *et al.* Addition of p -Azido- l -phenylalanine to the Genetic Code of Escherichia c oli. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
194. Wang, L. & Schultz, P. G. A general approach for the generation of orthogonal tRNAs. *Chem. Biol.* **8**, 883–890 (2001).
195. Santoro, S. W., Wang, L., Herberich, B., King, D. S. & Schultz, P. G. An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol.* **20**, 1044–1048 (2002).
196. Deiters, A. *et al.* Adding Amino Acids with Novel Reactivity to the Genetic Code of Saccharomyces Cerevisiae. *J. Am. Chem. Soc.* **125**, 11782–11783 (2003).
197. Liu, W., Brock, A., Chen, S., Chen, S. & Schultz, P. G. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods* **4**, 239–244 (2007).
198. Lee, H. S., Guo, J., Lemke, E. A., Dimla, R. D. & Schultz, P. G. Genetic Incorporation of a Small, Environmentally Sensitive, Fluorescent Probe into Proteins in Saccharomyces

cerevisiae. *J. Am. Chem. Soc.* **131**, 12921–12923 (2009).
199. Cole, S. T. *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537–544 (1998).
200. Blattner, F. R. *et al.* The complete genome sequence of Escherichia coli K-12. *Science (80-.).* **277**, 1453–1462 (1997).
201. Wang, F., Robbins, S., Guo, J., Shen, W. & Schultz, P. G. Genetic incorporation of unnatural amino acids into proteins in Mycobacterium tuberculosis. *PLoS One* **5**, (2010).
202. Touchette, M. H. *et al.* A Screen for Protein-Protein Interactions in Live Mycobacteria Reveals a Functional Link between the Virulence-Associated Lipid Transporter LprG and the Mycolyltransferase Antigen 85A. *ACS Infect. Dis.* **3**, 336–348 (2017).
203. Friehs, K. Plasmid Copy Number and Plasmid Stability. in *Advances in biochemical engineering/biotechnology* vol. 86 47–82 (Springer, Berlin, Heidelberg, 2004).
204. Wójcik, K., Wieckiewicz, J., Kuczma, M. & Porwit-Bóbr, Z. Instability of hybrid plasmids in Bacillus subtilis. *Acta Microbiol. Pol.* **42**, 127–36 (1993).
205. Timm, J. *et al.* Transcription and expression analysis, using lacZ and phoA gene fusions, of Mycobacterium fortuitum ?-lactamase genes cloned from a natural isolate and a high-level ?-lactamase producer. *Mol. Microbiol.* **12**, 491–504 (1994).
206. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
207. Pereira, G. C. *et al.* A High-Resolution Luminescent Assay for Rapid and Continuous Monitoring of Protein Translocation across Biological Membranes. *J. Mol. Biol.* **431**, 1689–1699 (2019).
208. Ke, N., Landgraf, D., Paulsson, J. & Berkmen, M. Visualization of periplasmic and cytoplasmic proteins with a self-labeling protein tag. *J. Bacteriol.* **198**, 1035–1043 (2016).
209. Taylor, R. G., Walker, D. C. & Mclnnes, R. R. E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* **21**, 1677–1678 (1993).
210. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **27**, 129–134 (2018).
211. Carroll, P. *et al.* Sensitive Detection of Gene Expression in Mycobacteria under Replicating and Non-Replicating Conditions Using Optimized Far-Red Reporters. *PLoS One* **5**, e9823 (2010).
212. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci.* **99**, 11020–11024 (2002).