



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

**Análisis mutacional de sitios con alta probabilidad
de glicosilación en la proteína LpqH de
Mycobacterium tuberculosis expresada en
*Streptomyces lividans***

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGA

P R E S E N T A:

KAREN VANESSA HUITRÓN SÁNCHEZ



TUTOR:

DR. LUIS SERVÍN GONZÁLEZ

Ciudad Universitaria, CD. MX., 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de datos del jurado

1. Datos del alumno

Huitrón

Sánchez

Karen Vanessa

58 40 50 10

Universidad Nacional Autónoma de México

Facultad de Ciencias

Biología

310181169

2. Datos del tutor

Dr.

Luis

Servín

González

3. Datos del sinodal 1

Dra.

Gloria

Soberón

Chávez

4. Datos del sinodal 2

Dra.

Claudia Andrea

Segal

Kischinevzky

5. Datos del sinodal 3

M. en C.

Gabriela

González

Cerón

6. Datos del sinodal 4

M. en C.

Marco Antonio

Carballo

Ontiveros

7. Datos del trabajo escrito

Análisis mutacional de sitios con alta probabilidad de glicosilación en la proteína LpqH de *Mycobacterium tuberculosis* expresada en *Streptomyces lividans*

46 p

2018

A mis amados padres

Agradecimientos

A la Universidad Nacional Autónoma de México que me abrió sus puertas desde el bachillerato, a mis profesores de la Facultad de Ciencias y a todos aquellos que han formado parte de mi desarrollo académico.

Al Dr. Luis Servín por recibirme en el laboratorio, por su paciencia y su inigualable forma de transmitir el conocimiento.

A la M. en C. Gabriela González Cerón por apoyarme desde el primer día hasta el último en el trabajo experimental y en el escrito.

A la Dra. Gloria Soberón Chávez por sus valiosos comentarios en los seminarios y por sus correcciones a este trabajo.

A la Dra. Claudia Andrea Segal Kischinevzky y al M. en C. Marco Antonio Carballo Ontiveros por su dedicación como profesores, por sus observaciones y correcciones a este trabajo.

A la Biol. María Guadalupe Códiz Huerta y a la Unidad de Secuenciación del Instituto de Fisiología Celular, U.N.A.M.

A mis amigos que siempre estuvieron ahí. A Fa con la que compartí momentos inolvidables en las prácticas de campo y en las aulas. A mis amigos de coro: Mariana, Sandra, DK, Jesús y Miriam porque con la música y con ellos pude sobrellevar los tiempos difíciles.

A mis amigos y compañeros del laboratorio: Angie, Erick, Lalo, Mauricio, Kevin, Katya, Bere, en especial a Isaí y a Gaby por enseñarme lo necesario cuando entré al laboratorio, por su paciencia y comprensión, a Yersaín por ayudarme a mejorar con sus comentarios dentro y fuera de los seminarios y por tanto tiempo divertido en el laboratorio.

A Dios y a mi hermosa familia. A mis papás por tanto amor y comprensión indescriptible, por enseñarme a ser una mejor persona, por su tiempo y dedicación. A mis hermanos, a mis abuelitas, tíos y primos que hacen este camino más fácil.

Índice

1. Resumen	6
2. Introducción	7
2.1 <i>Mycobacterium tuberculosis</i>	8
2.2 Características de la tuberculosis	9
2.3 Biosíntesis de lipoproteínas de <i>M. tuberculosis</i>	5
2.4 Glicosilación de lipoproteínas	11
2.5 Motivo de glicosilación	12
2.6 Glicoproteínas de <i>M. tuberculosis</i>	14
2.6.1 Proteína de 19 kDa, LpqH	14
2.7 <i>Streptomyces</i>	15
3. Justificación	18
4. Hipótesis	19
5. Objetivos	19
5.1 Objetivo general	19
5.2 Objetivos particulares	19
6. Materiales y métodos	20
6.1 Cepas, plásmidos y oligonucleótidos	20
6.2 Predicción de sitios de glicosilación en LpqH	23
6.3 Mutagénesis dirigida del sitio en Rv3763 que codifica para la Serina 43 para su sustitución por Alanina	23
6.4 Construcción del gen Rv3763 con las mutaciones T34-36A, T40-41A y T34-36,40-41A y la etiqueta de seis histidinas	24
6.5 Clonación del gen Rv3763 con las mutaciones S43A, 3A2T, 3T2A y 3A2A en el vector multicopia pIJ6021	27
6.6 Transformación de la cepa 1326 de <i>S. lividans</i> con los plásmidos pIJS43AH, pIJ3A2TH, pIJ3T2AH y pIJ3A2AH	27
6.7 Expresión de la proteína LpqH	28
6.8 Fraccionamiento celular	28
6.9 Purificación de proteínas por afinidad a níquel	29
6.10 Electroforesis de proteínas en gel de poliacrilamida-SDS (SDS-PAGE)	29
6.11 Transferencia de proteínas a membrana PVDF	30
6.12 Inmunodetección de proteínas	30
6.13 Determinación del estado de glicosilación de LpqH por reconocimiento con concanavalina A	30
7. Resultados y discusión	31
7.1 Identificación y selección de probables sitios de glicosilación en LpqH	31
7.2 Mutagénesis múltiple del gen Rv3763, para introducir los cambios S43A, T34-36A, T40-41A y T34-36,40-41A en la proteína LpqH	35
7.3 Expresión y glicosilación de la proteína LpqH silvestre y sus variantes	36
8. Conclusiones	42
9. Referencias	43

1. Resumen

En algunos estudios de la proteína LpqH de *Mycobacterium tuberculosis* se ha argumentado que la mutagénesis de cinco residuos de treonina (Thr₃₄₋₃₆ y Thr₄₀₋₄₁) localizados dentro de una secuencia de ocho aminoácidos en el extremo N-terminal de la proteína madura abate la glicosilación. Sin embargo, sigue indefinido si los residuos que coinciden con el motivo de glicosilación son específicamente los que permiten la glicosilación de la proteína y si existe algún otro sitio de glicosilación independiente. Para profundizar en esto se seleccionaron los residuos con base en los puntajes de glicosilación generados a partir de NetOGlyc, con la paridad que existe entre estos y el motivo de glicosilación. Se realizó la sustitución de los cinco residuos de treonina y del residuo Ser₄₃ por alanina de la proteína LpqH de *M. tuberculosis*, siendo únicamente Ser₄₃ el residuo que coincide con el motivo de glicosilación. Las versiones modificadas de esta proteína fueron expresadas y purificadas en *Streptomyces lividans* para ser posteriormente analizadas mediante un ensayo de Western blot con concanavalina A. Este ensayo reveló que la proteína LpqH con la mutación de Thr₃₄, Thr₃₅ y Thr₃₆ no tiene un efecto negativo en la glicosilación al igual que LpqH con la mutación de Thr₄₀ y Thr₄₁ y con la mutación individual de Ser₄₃, mientras que la proteína modificada en las cinco treoninas resulta en el abatimiento de la glicosilación de la proteína.

2. Introducción

2.1 *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (Mtb) es una especie bacteriana que pertenece al grupo de los actinomicetos. Es un aerobio obligado, sin movilidad y con un crecimiento muy lento en medios ricos; por ejemplo, los cultivos de aislados clínicos tardan de 4 a 6 semanas para crecer (McMurray, 1996). A nivel celular, se trata de bacilos Gram-positivos, con una envoltura celular constituida por cápsula, pared celular y membrana plasmática. La cápsula sirve de protección y tiene una interacción directa con los elementos de la respuesta inmune. La membrana citoplasmática está cubierta por una capa densa de peptidoglicanos unidos a polisacáridos esterificados con los ácidos micólicos formados por lípidos libres, peptidoglicolípidos y glicolípidos, lo que le confiere alta hidrofobicidad, resistencia a detergentes y a antibióticos. Los glicolípidos y algunas proteínas son responsables de las características antigénicas de la bacteria (Brennan, 1989) (Fig. 1).

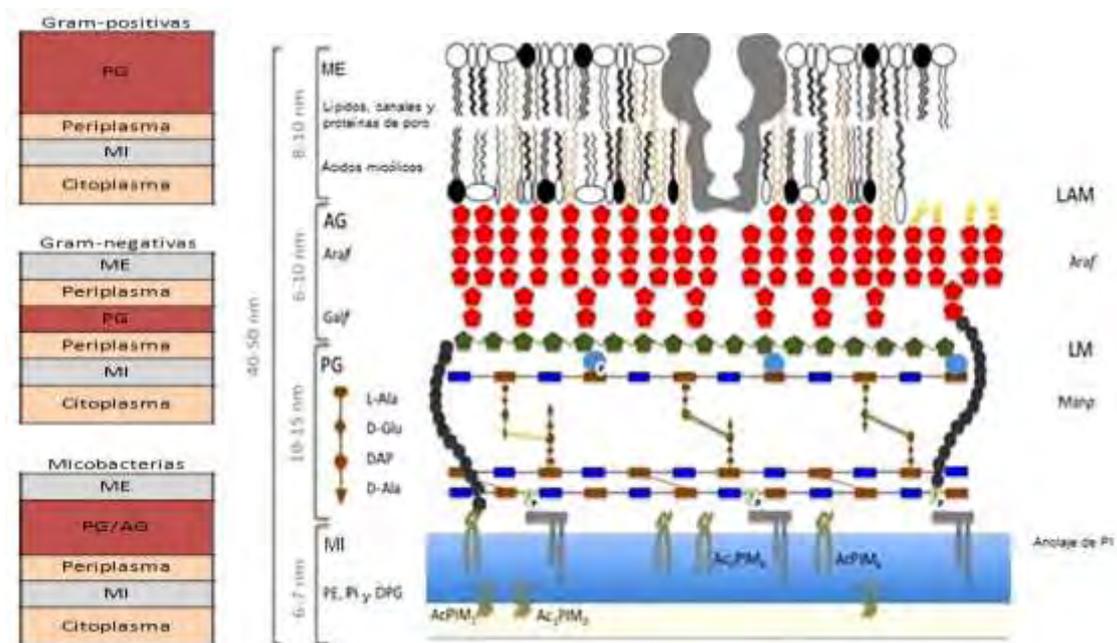


Figura 1. Estructura general de la pared celular de micobacterias. (Panel izquierdo) Composición de la envoltura celular de gram-positivas, gram-negativas y micobacterias. (Panel

derecho) Capas de la envoltura celular de micobacterias. La membrana interna (MI) con los fosfolípidos más abundantes representados fosfatidilinositol (PI), fosfatidiletanolamina (PE), difosfatidilglicerol (DPG) y los fosfatidilinositol manósidos (PIMs). Lipoarabinomanano (LAM) y lipomanano (LM) están unidos covalentemente a la IM y las unidades de manosa están en forma de piranosa (Manp). Los peptidoglicanos, arabinogalactano (AG) y los ácidos micólicos (MA) que también contribuyen a la membrana externa forman el centro de la pared celular. Los arabinanos y galactános construyen bloques de furanosa (Araf y Galf). La capa de AG está unida covalentemente a los motivos de ácido murámico. Modificado de: Rebrova, 2017.

2.2 Características de la tuberculosis

La tuberculosis, causada por *M. tuberculosis*, es la segunda enfermedad infecciosa más letal a nivel mundial, pues ocasiona tres millones de muertes anuales e infecta de manera latente a entre mil a dos mil millones de personas. Aunque la mayoría de las infecciones activas de tuberculosis son curables con un diagnóstico temprano, existen en la actualidad cepas de *M. tuberculosis* que han desarrollado resistencia a los medicamentos, además de que es difícil y costosa de tratar en individuos inmunocomprometidos (World Health Organization, 2014).

La tuberculosis pulmonar se transmite por la inhalación de bacilos exhalados de pacientes con la enfermedad. Los síntomas incluyen tos, debilidad, fiebre, sudores nocturnos, pérdida de peso y producción de esputo con sangre (Jiansong *et al.*, 2016).

La vacuna contra la tuberculosis utilizada en la actualidad es la vacuna BCG (bacilo de Calmette-Guérin) que fue aislada originalmente de una cepa viva atenuada de *M. bovis* por Albert Calmette y Camille Guérin. Se ha utilizado ampliamente, aunque no es eficaz contra la tuberculosis pulmonar en adultos por dos razones: no es capaz de inducir respuesta inmune de larga duración después de la vacunación y no evita la reactivación de la tuberculosis latente que es la principal propagación de bacilos. Además, la exposición previa a micobacterias ambientales interfiere con su eficacia que varía según la zona geográfica. (Andersen, 2005). En términos de eficacia, ninguna cepa de BCG es claramente mejor que las otras y no existe un consenso mundial sobre cuál es idónea para uso general. Con el objetivo de hacer más eficientes candidatas a

vacunas, se ha obtenido la secuencia completa del genoma de *M. tuberculosis*, lo cual ha permitido el desarrollo de investigaciones enfocadas al estudio de moléculas antigénicas indispensables para la virulencia de la bacteria como proteínas, lipoglicanos y glicoproteínas (Corbière *et al.*, 2014)

2.3 Biosíntesis de lipoproteínas de *M. tuberculosis*

En años recientes se han identificado genes de micobacterias esenciales para la virulencia. La mayoría codifican enzimas de vías de síntesis de lípidos, proteínas de la superficie celular y proteínas de los sistemas de transducción de señales (Forrellad *et al.*, 2013). Las lipoproteínas (lpp) se encuentran asociadas a la membrana y participan en el transporte de nutrientes a través de la pared celular y algunas en la inducción de respuestas humoral y celular, especialmente de la respuesta de las células T de memoria en bacterias Gram-positivas (Young, y Garbe, 1991).

Aparentemente todas las bacterias localizan proteínas específicas en sus envolturas celulares mediante modificaciones postraduccionales lipídicas (lipoproteínas), que facilitan el anclaje de proteínas hidrofílicas a superficies hidrofóbicas por la interacción de los grupos acilo con los fosfolípidos de la membrana celular, permitiendo a las proteínas funcionar de manera eficaz en medio acuoso. Asimismo, estas moléculas están involucradas en la biogénesis de la envoltura celular y plegamiento de proteínas (Kamalakkannan *et al.*, 2004). La biogénesis de lpp es esencial para la supervivencia de las bacterias Gram-positivas y depende de la presencia del péptido señal que dirige a la prelipoproteína a exportarse a través de la membrana plasmática. El péptido señal contiene tres regiones: la región N se encuentra cargada de forma positiva, seguida por una región altamente hidrofóbica (región H) y finalizando con la región de corte C, que contendrá el lipobox en el extremo C-terminal.

Anteriormente se aceptaba que las lipoproteínas bacterianas se translocaban por la vía Sec, sin embargo, ahora se sabe que existen lipoproteínas que también se translocan por la vía Tat, por la cual se secretan las proteínas plegadas (Berks, B., 2000).

Para que las proteínas sean secretadas por Sec, deben permanecer en un estado no plegado. Este sistema se compone por la proteína SecA que es una ATPasa y que junto con el canal SecYEG secretan las proteínas a través de la membrana citoplasmática. El sistema de translocación de argininas gemelas (twin-arginine transport) o Tat a diferencia del sistema Sec, no se encuentra en todas las bacterias y requiere que todas las proteínas transportadas se encuentren completamente plegadas. Este sistema secreta proteínas de manera postranscripcional cuyo péptido señal tiene un motivo Tat que contiene dos residuos de arginina conservados entre la región N y la región H. El sistema de secreción Tat está conformado por el complejo protéico TatBC que induce la formación del poro del translocón conformado por la interacción de proteínas TatA. (Goosens, *et al.* 2014) (Fig. 2).

El lipobox es un motivo peptídico constituido por cuatro aminoácidos ([LVI], [ASTVI], [GAS], [C]), tiene en su secuencia una cisteína conservada en el extremo C-terminal que funciona para ser reconocida por la maquinaria de lipidación de proteínas. La enzima que se encarga de la lipidación es la lipoproteína diacilglicerol transferasa (Lgt), que puede ser indispensable en Gram-positivas y que añade un diacilglicerol por un enlace tioéster al grupo sulfhidrilo de la cisteína del lipobox, quedando así anclada la proteína a la membrana. Posteriormente a la lipidación, el péptido señal es eliminado por medio de un corte que lleva a cabo la peptidasa señal de lipoproteínas (Lsp), mediante el corte del enlace peptídico inmediatamente anterior a la cisteína lipidada, convirtiendo a este aminoácido en el extremo N-terminal de la proteína madura. En Gram-negativas, y en algunas Gram-positivas como *M. tuberculosis* pueden ocurrir pasos adicionales como la adición de ácidos grasos en el grupo amino de la cisteína por una N-aciltransferasa (Lnt) (Sutcliffe, y Harrington, 2004) (Fig. 2).

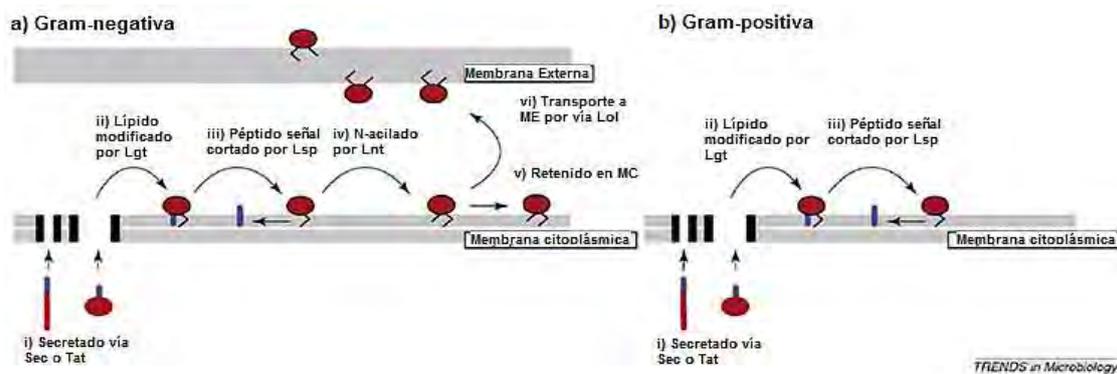


Figura 2. Vía de biogénesis de lipoproteínas. a) En bacterias Gram-negativas: 1) la lpp plegada (círculo rojo) o no plegada (línea roja) es directamente translocada a través de la membrana citoplásmica por la vía Sec o Tat por sus péptidos o secuencias señal (en azul), 2) un grupo lipídico consistente en un diacilglicerol (línea negra angulada) es covalentemente unido al grupo sulfhidrilo de la cisteína del lipobox por la enzima Lgt, 3) el péptido señal es cortado por la peptidasa Lsp, 4) un segundo grupo lipídico consistente en una cadena de acilo (línea negra) es enlazado al grupo amino de la cisteína del lipobox por Lnt, 5) esto ocurre en estricto orden y son retenidos en la membrana citoplásmica, o 6) son transportados a la membrana externa por la vía Lol (responsable de la localización de lipoproteínas de membrana externa). b) en bacterias Gram-positivas: 1) la vía es conservada pero no necesariamente ocurre en estricto orden. Hay evidencia de la N-acilación en bacterias Gram-positiva con bajo contenido GC como *Bacillus subtilis* y *Staphylococcus aureus* a pesar de la ausencia de homólogos de Lnt. Por el contrario, algunas bacterias Gram-positivas con alto contenido GC como *Mycobacterium* y *Streptomyces* llevan a cabo la N-acilación por homólogos de Lnt. Modificado de: Sutcliffe y Harrington, 2004.

2.4 Glicosilación de lipoproteínas

La adición de carbohidratos o glicanos a glicoproteínas extracelulares y de membrana es una modificación postraduccional. Arqueas, procariontes y eucariontes comparten precursores biosintéticos, lo cual sugiere que la glicosilación es una modificación universal de proteínas que tiene un papel biológico similar en todos los organismos como la regulación de la estructura, protección contra la degradación, estabilidad y la función de proteínas membranales y secretadas como la adhesión celular y el reconocimiento inmunológico específico (González-Zamorano *et al.*, 2009).

En la N-glicosilación de proteínas los precursores de los oligosacáridos se ensamblan primero en una molécula acarreadora antes de ser transferidos en bloque al nitrógeno del grupo amido de un Asn de la proteína aceptora, la cual

se encuentra dentro del motivo Asn-X-Ser/Thr, donde X puede ser cualquier aminoácido excepto prolina. Este proceso es común en eucariontes y en algunas bacterias Gram-negativas, pero no se ha observado en bacterias Gram-positivas ni en micobacterias patógenas. En las bacterias Gram-positivas, y en especial en los actinomicetos, es común la O-glicosilación de proteínas por medio de residuos de manosa, proceso conocido como O-manosilación. A diferencia de la N-glicosilación, en el proceso de O-manosilación el residuo inicial de manosa es transferido al grupo hidroxilo de un residuo aceptor de treonina o serina por la enzima O-manosiltransferasa de proteínas (Pmt). Posterior a la adición de este primer residuo de manosa pueden ser añadidos otros a la manosa inicial por medio de otras manosiltransferasas. Aunque el papel de la O-glicosilación aún se está dilucidando, se sabe que es esencial para la virulencia en micobacterias patógenas, pues mutantes carentes de Pmt no son virulentas (Corbière *et al.*, 2014). El reconocimiento célula-célula entre el patógeno (Mtb) y el hospedero está mediado parcialmente por la glicosilación de proteínas (Smith *et al.*, 2014). Se ha propuesto que los receptores de manosa en las células hospederas pueden interactuar directamente con proteínas manosiladas de Mtb para entrar a los macrófagos y contribuir a su supervivencia (Peter-Katalinić, 2005).

2.5 Motivo de glicosilación

Aunque el conocimiento acerca de las glicoproteínas de Mtb es limitado, mediante un estudio glicoproteómico se han obtenido los sitios de glicosilación de 33 glicoproteínas obtenidas de un filtrado de cultivo de Mtb utilizando técnicas de cromatografía líquida con espectrometría de masas (Smith *et al.*, 2014). A partir de los datos obtenidos de dicho análisis, mediante la aplicación WebLogo se generó el logo de la secuencia, que es la representación gráfica de la secuencia alineada de aminoácidos. Éste consiste en símbolos apilados, en donde una pila corresponde a cada posición en la secuencia donde la altura de los caracteres es proporcional a la frecuencia relativa de cada aminoácido en la posición en la que se encuentran. A partir de este logo se puede determinar no sólo una secuencia consenso, sino también la frecuencia relativa de las bases y la información del contenido en cada posición en el sitio o secuencia (Schneider,

y Stephens, 1990). En general, un logo provee una descripción más rica y precisa de un sitio de unión a comparación de una secuencia consenso (Crooks *et al.*, 2004).

La Figura 3 presenta el logo de la secuencia que determinaron Smith *et al.* (2014) a partir de sus datos glicoproteómicos, donde se muestra que aproximadamente el 60% de los sitios de glicosilación correspondieron a residuos de treonina y el 40% en residuos de serina. Estos resultados también revelaron la alta frecuencia de los sitios de glicosilación en los extremos C-terminal de las proteínas. También se observa que los sitios de O-glicosilación se encuentran en regiones con alta presencia de residuos de prolina y de alanina como ya se había demostrado (Smith *et al.*, 2014).



Figura 3. Motivo de glicosilación. El motivo de O-glicosilación generado de las identificaciones de sitios de glicosilación en un estudio glicoproteómico muestra una alta probabilidad de prolina y alanina hidrofóbicas intercaladas con algunas treoninas hidrofílicas alrededor del sitio de manosilación, correspondiente a la treonina o serina de la posición 16. El motivo de glicosilación también indica que muchos sitios se encuentran cerca del C-terminal de las proteínas, lo cual está representado por los recuadros negros a la derecha, que corresponden a la ausencia de aminoácidos. Tomado de: Smith, 2014.

Existen herramientas que permiten predecir los sitios de glicosilación en proteínas eucariontes como NetOGlyc y NetNGlyc. En el caso de NetNGlyc la predicción se basa en el motivo Asn-X-Ser/Thr, donde X no puede ser una prolina, pero en caso de NetOGlyc la predicción se basa en un algoritmo derivado de los sitios de glicosilación determinados a partir de glicoproteínas conocidas, el cual se va modificando conforme más proteínas son estudiadas. La versión

4.0 con respecto a la anterior versión 3.1, representa una importante mejoría en cuanto a la sensibilidad y especificidad general ante los diferentes conjuntos de datos. La versión 3.1 a pesar de su menor sensibilidad, predice la mayoría de los sitios ortogonales al nuevo predictor. Además, se argumenta que aún existe la necesidad apremiante de integrar y mejorar la información de datos negativos (que indican no glicosilación) (Steentoft *et al.*, 2013).

2.6 Glicoproteínas de *M. tuberculosis*

Las micobacterias tienen una gran diversidad de biomoléculas glicosiladas complejas con enlace de tipo O–glicosídico.

Se estima que el 25% del peso seco de las micobacterias corresponde a lípidos o glicolípidos. Se sabe que la mayoría de las glicoproteínas, lipoproteínas y glicolipoproteínas secretadas por *M. tuberculosis* son antígenos que el sistema inmune humano reconoce y que han sido detectadas por lectinas en el sobrenadante de cultivos y en las membranas de *M. tuberculosis*. Las proteínas de 19, 38 y 45 kDa son algunas de las glicoproteínas más estudiadas de este organismo (Hermann *et al.*, 1996).

2.6.1 Proteína de 19 kDa, LpqH

La proteína de 19 kDa o LpqH es miembro de una familia altamente inmunogénica de lipoproteínas micobacterianas asociadas a la superficie. Es un antígeno secretado que se dirige a la ruta de exportación Sec de proteínas por un péptido señal de secreción de 21 aminoácidos (Parra *et al.*, 2017).

Dentro del complejo de Mtb el gen que codifica para el antígeno de 19 kDa (Rv3763) está muy conservado y tanto *M. avium* como *M. intracellulare* tienen genes que codifican proteínas con un alto grado de identidad, de aproximadamente un 75%, con el complejo antigénico de *M. tuberculosis*. Otras especies como *M. asiaticum*, *M. phlei*, *M. kansasii*, *M. paratuberculosis*, y *M. scrofulaceum* pueden tener un antígeno homólogo, pero no hay reportes de proteínas relacionadas con el antígeno de 19 kDa en otros géneros bacterianos

(Mahenthiralingam *et al.*, 1998). La proteína LpqH participa en la regulación de la inmunidad adaptativa induciendo la secreción de citocinas y moléculas coestimuladoras en células del sistema inmune innato y a través de la regulación directa de la activación de linfocitos T de memoria (Ocampo *et al.*, 2014). Se ha encontrado después de examinar las cinéticas de crecimiento de cepas carentes de este antígeno, que aislados de Mtb en los que no se expresa la proteína LpqH, se ven afectados en la virulencia (Henao-Tamayo *et al.*, 2007). A pesar del amplio estudio de la inmunobiología de LpqH, su función específica permanece desconocida.

La glicosilación puede ser parte de un mecanismo protector para retener al antígeno de 19 kDa en su ubicación asociada a la membrana y asegurar su papel inmunológico. La evidencia de glicosilación de esta lipoproteína inicialmente se basó en su capacidad de reconocimiento por la lectina concanavalina A en ensayos de Western blot (Sutcliffe, y Harrington, 2004).

En un estudio de mutagénesis se propusieron cinco sitios potenciales de glicosilación de la proteína a través del reemplazo de los residuos de treonina localizados en el extremo N-terminal entre las posiciones 33 y 41, por residuos de valina, que abate la unión de LpqH a ConA (Hermann *et al.*, 1996).

Recientemente, en un estudio donde se expresa la proteína LpqH en *M. smegmatis* se ha confirmado mediante análisis de espectrometría de masas que la glicosilación ocurre en cuatro de estos cinco residuos de treonina de manera progresiva en el siguiente orden: Thr₄₁, Thr₃₅, Thr₃₄ y Thr₃₆; esta progresión preferencial probablemente esté aunada a una diferente accesibilidad de la proteína manosa transferasa a los distintos sitios de glicosilación (Parra *et al.*, 2017).

2.7 Streptomyces

Es un género de actinobacterias que abunda en el suelo. Una de sus características más conocidas es la producción de antibióticos, además sintetizan fungicidas, citostáticos, moduladores de la respuesta inmune y efectores del crecimiento de las plantas. *Streptomyces* tiene una adaptación al crecimiento que ayuda a adherirse y penetrar en los restos orgánicos insolubles

de hongos, plantas y otros organismos, que es el crecimiento miceliar (Chater *et al.*, 2010). De forma parecida a los hongos, los estreptomicetos crecen como hifas ramificadas que forman un micelio vegetativo y se dispersan a través de las esporas que forman estructuras reproductivas especializadas llamadas hifas aéreas, las cuales emergen de la superficie de la colonia hacia el aire. Estas similitudes podrían ser el resultado de adaptaciones a nichos ecológicos similares, aunque en realidad tienen diferentes orígenes evolutivos.

Cuando una spora de *Streptomyces* se encuentra en condiciones favorables con los nutrientes necesarios, germina. Es decir, uno o dos tubos germinales emergen y crecen para formar hifas; estos tubos germinales crecen por extensión y ramificación de la punta para formar un micelio vegetativo que se dispersa en el sustrato.

En condiciones de estrés se producen metabolitos secundarios y ocurre la diferenciación morfológica. Posteriormente, las hifas aéreas rompen la tensión superficial escapando de la capa acuosa del micelio vegetativo y crecen en el aire. Las hifas aéreas se dividen por una forma de división celular controlada por el desarrollo, en largas cadenas de compartimentos de preesporas, que luego desarrollan paredes gruesas, sintetizan un pigmento gris y adquieren otras características de esporas maduras. Las esporas de *Streptomyces* son menos resistentes a las condiciones adversas, aunque pueden sobrevivir largo tiempo en un estado de desecado, lo cual indica que están bien adaptadas para la dispersión en el medio ambiente (Flärdh y Buttner, 2009) (Fig. 4). En un ambiente acuoso, *Streptomyces* crece como redes de micelio que típicamente forman *pellets* grandes o grupos de *pellets*. El ambiente impacta la morfología profundamente. Entre otros, la composición del medio, la viscosidad, la hidrodinámica (si se trata de un medio líquido) y el pH afectan la fisiología del cultivo (van Dissel y van Wezel, 2018).

Streptomyces coelicolor es el modelo más estudiado de los estreptomicetos, el cual tiene un cromosoma lineal, alto contenido de G+C (72 %) en su DNA, más de 6,200 marcos de lectura abierta (ORFs), y gran cantidad de proteínas de secreción (Chater *et al.*, 2010). Un análisis proteómico ha confirmado que los estreptomicetos secretan numerosas proteínas, de las cuales un alto porcentaje está involucrado en transporte y en la degradación de nutrientes extracelulares.

S. lividans tiene muchas similitudes con *S. coelicolor*, sin embargo, esta especie se utiliza en experimentos para la producción heteróloga de proteínas por su menor actividad de proteasas en el medio y por la ausencia de un sistema de restricción del DNA dependiente de metilación. Por lo tanto, no degrada DNA metilado de otras bacterias como *Escherichia coli*, facilitando su manipulación en el laboratorio (González-Cerón *et al.*, 2009).

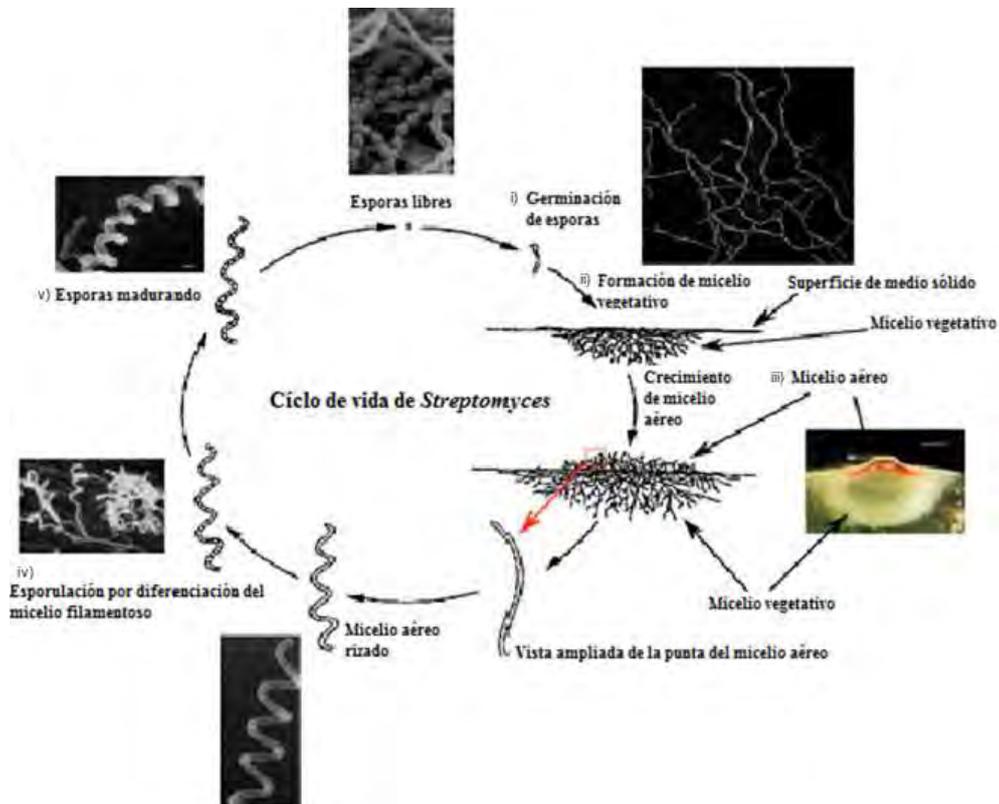


Figura 4. Ciclo de vida de *Streptomyces*. i) Una espora germina, los tubos germinales emergen, ii) crecen por extensión y ramificación formando el micelio vegetativo que se dispersa en el medio sólido. iii) El micelio vegetativo forma el micelio aéreo orientado verticalmente hacia la superficie. iv) El micelio aéreo en espiral y luego los filamentos aéreos se dividen en largas cadenas de compartimentos de preesporas que v) maduran hasta convertirse en esporas. Los metabolitos secundarios, incluidos los antibióticos, se secretan durante la generación de hifas del micelio vegetativo. Modificado de: Laboratorio de bioquímica celular, Universidad de Hiroshima. (2011).

3. Justificación

La tuberculosis, ocasionada por *M. tuberculosis*, es la segunda causa de mortalidad por agentes infecciosos a nivel mundial, provocando aproximadamente 1.7 millones de muertes al año. Aunque existe una vacuna contra la infección por el agente causal de esta enfermedad (la vacuna BCG), se ha demostrado que no es completamente efectiva. Por lo que, con el objetivo de hacer vacunas más eficientes, se han desarrollado investigaciones enfocadas al estudio de moléculas antigénicas indispensables para la virulencia de la bacteria como proteínas, lipoglicanos y glicoproteínas; para así entender los mecanismos de acción del patógeno, sus interacciones y modificaciones.

En la actualidad los estudios proteómicos han recibido mayor atención ya que muchas proteínas están involucradas en la respuesta inmune, principalmente en el reconocimiento célula-célula entre el patógeno (*Mtb*) y su hospedero mediado por glicoproteínas.

La expresión de proteínas heterólogas en organismos no patógenos como *S. lividans* minimiza el riesgo de infección y facilita el trabajo experimental, aumentando la velocidad de producción y la cantidad de proteínas obtenidas sin afectar el patrón de glicosilación que ocurre en el organismo original.

En *M. tuberculosis*, se ha propuesto que la proteína de 19 kDa se glicosila en el extremo N-terminal en cinco residuos de treonina. Sin embargo, existen otros residuos que podrían tener una mejor coincidencia con el motivo de glicosilación, determinado mediante estudios proteómicos y que podrían estar involucrados en la glicosilación de la lipoproteína expresada en *S. lividans*.

4. Hipótesis

La expresión de la proteína de 19 kDa (LpqH) en *S. lividans* resultará en su glicosilación en los mismos residuos de treonina que han sido sugeridos como los sitios donde ocurre la glicosilación. Por lo tanto, mutaciones que sustituyan estos cinco residuos de treonina (Thr₃₄, Thr₃₅, Thr₃₆, Thr₄₀ y Thr₄₁) por alanina resultarán en el abatimiento de la glicosilación. Aunque es probable que otro residuo de treonina o serina de LpqH se ajuste al motivo de glicosilación y esté glicosilado, evento que puede ser a su vez predicho por la herramienta NetOGlyc.

5. Objetivos

5.1 Objetivo general

Evaluar el efecto de mutaciones en sitios con alta probabilidad de glicosilación en la lipoproteína LpqH de *M. tuberculosis* expresada en *S. lividans*, utilizando el motivo de glicosilación y el *software* de predicción de glicosilación NetOGlyc.

5.2 Objetivos particulares

- Identificar y mutar posibles sitios de glicosilación en la proteína LpqH considerando el motivo de glicosilación y la herramienta NetOGlyc.
- Expresar y purificar mediante cromatografía de afinidad a níquel la proteína LpqH con las distintas mutaciones en *S. lividans*.
- Detectar si se encuentra glicosilación en la proteína de 19 kDa con las mutaciones realizadas, mediante un ensayo de inmunodetección con concanavalina A.

6. Materiales y métodos

6.1 Cepas, plásmidos y oligonucleótidos

Se utilizaron cepas de *E. coli* para la manipulación del gen Rv3763 y cepas de *S. lividans* para la expresión de las diferentes formas de la proteína recombinante LpqH. En la Tabla 1 se resumen las cepas utilizadas, en la Tabla 2 los plásmidos utilizados y construidos, y en la Tabla 3 los oligonucleótidos utilizados en este trabajo.

Tabla 1. Cepas utilizadas para clonar y expresar Rv3763.

Cepa	Descripción	Referencia
<i>E.coli</i>		
DH5 α	F ⁻ <i>endA1glnV44thi-1recA1relA1gyrA96deoRnupG</i> Φ 80d <i>lacZ</i> Δ M15 Δ (<i>lacZYA argF</i>)U169, <i>hsdR17</i> (rK ⁻ mK ⁺), λ -	Grant <i>et al.</i> , 1990
IBEC58	Δ <i>dam</i> Δ <i>dcm</i> Δ <i>hsdRMS</i> derivado de BW25113. Para obtener DNA no metilado.	González-Cerón <i>et al.</i> , 2009
<i>Streptomyces</i>		
<i>S. lividans</i> 1326	Silvestre.	John Innes Centre, cepa 1326
Δ <i>pmt</i>	Derivada de <i>S. lividans</i> 1326 con el gen SI3509 deletado, el cual codifica para la O-manosiltransferasa <i>pmt</i> .	Arista, 2016

Tabla 2. Plásmidos utilizados para clonar los genes en las cepas.

Plásmido	Descripción	Referencia
pUC18	Plásmido que contiene un sitio múltiple de clonación y un gen que confiere resistencia a ampicilina. Produce el péptido de la β -galactosidasa que complementa la delección de <i>lac</i> en cepas de <i>E. coli</i> como DH5 α .	Yanisch-Perron <i>et al.</i> , 1985
pUC18Rv3763H	Gen Rv3763 modificado para la adición de 6His a LpqH, clonado en pUC18.	Este trabajo
pUCS43AH	Gen Rv3763 con la mutación S43A+6His clonado en pUC18.	Este trabajo
pQ3A2T	Fragmento sintético del gen Rv3763 con la sustitución T34-36A, clonado en el vector pMA-RQ en <i>E. coli</i> (amp ^R) flanqueado por sitios <i>Bam</i> HI y <i>Nru</i> I.	Este trabajo
pQ3T2A	Fragmento sintético del gen Rv3763 con la sustitución T40-41A clonado en el vector pMA-RQ en <i>E. coli</i> (amp ^R) flanqueado por sitios <i>Bam</i> HI y <i>Nru</i> I.	Este trabajo
pQ3A2A	Fragmento sintético del gen Rv3763 con la sustitución del <i>cluster</i> de cinco treoninas por alaninas (T34-36,40-41A) clonado en el vector pMA-RQ en <i>E. coli</i> (amp ^R) flanqueado por los sitios <i>Bam</i> HI y <i>Nru</i> I.	Este trabajo
pTZ18r	Vector de clonación en <i>E. coli</i> con marcador <i>lacZ</i> , amp ^R .	Mead <i>et al.</i> , 1986
pZ3A2T	Gen Rv3763 con la sustitución T34-36A clonado en pTZ18r.	Este trabajo
pZ3T2A	Gen Rv3763 con la sustitución T40-41A clonado en pTZ18r.	Este trabajo
pZ3A2A	Gen Rv3763 con la sustitución del <i>cluster</i> de 5 treoninas por alaninas (T34-36,40-41A) clonado en pTZ18r.	Este trabajo

pZ3A2TH	Gen Rv3763 completo con la sustitución T34-36A + 6His, clonado en pTZ18r.	Este trabajo
pZ3T2AH	Gen Rv3763 completo con la sustitución T40-41A + 6His, clonado en pTZ18r.	Este trabajo
pZ3A2AH	Gen Rv3763 completo con la sustitución del <i>cluster</i> de cinco treoninas por alaninas (T34-36,40-41A) + 6His clonado en pTZ18r.	Este trabajo
pIJ6021	Vector de expresión Kan ^R inducible con tioestreptón, alto número de copias.	Takano <i>et al.</i> , 1995
pIJ6021+3763	Gen Rv3763+6His clonado en pIJ6021.	Este trabajo
pIJS43AH	Gen Rv3763 con la sustitución S43A + 6His clonado en pIJ6021.	Este trabajo
pIJ3A2TH	Gen Rv3763 con la sustitución T34-36A + 6His clonado en pIJ6021.	Este trabajo
pIJ3T2AH	Gen Rv3763 con la sustitución T40-41A + 6His clonado en pIJ6021.	Este trabajo
pIJ3A2AH	Gen Rv3763 con la sustitución del <i>cluster</i> de cinco treoninas por alaninas (T34-36,40-41A) + 6His clonado en pIJ6021.	Este trabajo

Tabla 3. Oligonucleótidos utilizados para la mutagénesis dirigida y para secuenciación.

Oligonucleótido	Secuencia	
Rv3763HQCHSAFor	FWD 5'- GGCAGGCACGACGGCAGCCCCCGGCGCCGC- 3'	Producto de QuikChange pUCS43AH
Rv3763HQCHSARev	REV 5'- GCGGCGCCGGGGGCTGCCGTCGTGCCTGCC- 3'	
6021NdeI	FWD 5'- GCTTGCACCTCACGTCAC-3'.	Para secuenciar genes clonados en pIJ6021
6021EcoRI	REV 5'-CACTCCGCTGAAACTGTTGA-3'	

6.2 Predicción de sitios de glicosilación en LpqH

La secuencia de la proteína LpqH disponible en la base de datos TubercuList (<http://tuberculist.epfl.ch/quicksearch.php?gene+name=rv3763>) fue sometida a una búsqueda de motivos con el programa MOTIF Search (disponible en línea: <http://www.genome.jp/tools/motif/>) a partir del motivo consenso de glicosilación generado por Smith *et al.* (2014), pero acotado a solamente cinco aminoácidos a cada lado del sitio de glicosilación, es decir:

[AGSDPTEIV]-[APTDEGIKLSV]-[TAGPVIMQR]-[TAPSDKLMV]-[APTDHKQ]-
[TS]-[PDTAMN]-[TDPQRSFAFGIV]-[PTDKVAENQRS]-[APSEGINQRT]-
[TAIPCDFGLQRSY].

Como se observa, este motivo es más corto que el original ya que el uso del motivo completo no permite la predicción de la proteína de 19 kDa. Se encontró el motivo AGTTASPGAAS en la posición 38-48 y con este resultado, se sometió a análisis a la proteína por el *software* de predicción de glicosilación NetOGlyc 3.1 y NetOGlyc 4.0 *model for prediction of O-glycosylation* (disponible en línea: <http://www.cbs.dtu.dk/services/NetOGlyc/>) (Steentoft *et al.*, 2013), con el objetivo de seleccionar los residuos de aminoácido que tuviesen una alta probabilidad de estar glicosilados para posteriormente sustituirlos.

6.3 Mutagénesis dirigida del sitio en Rv3763 que codifica para la Serina 43 para su sustitución por Alanina

Para realizar la mutagénesis dirigida *in vitro* de la serina 43 (que tiene un alto puntaje como probable sitio de glicosilación y que fue detectada mediante la búsqueda con el programa MOTIF search) se utilizó el protocolo QuikChange para el cual se diseñaron los oligos de la Tabla 3.

Se utilizó como templado para la síntesis pUC18Rv3763H que contiene la secuencia del gen Rv3763 modificada para añadir una etiqueta de seis histidinas para posteriormente purificar la proteína por cromatografía de afinidad a Ni.

El protocolo QuikChange se realizó bajo las siguientes condiciones: 95°C/ 30 segundos, (95°C/ 30 segundos, 55° C/ 1 minuto, 68° C/ 5 minutos) x12 ciclos (Agilent Technologies, 2015). El producto de la síntesis se digirió con *DpnI* a

37°C durante 2 horas para eliminar el DNA templado metilado sin mutaciones y se transformó en la cepa de *E. coli* DH5α mediante electroporación de células electrocompetentes a 1.6 kV. De las colonias transformantes obtenidas se realizaron minipreparaciones de DNA de plásmido (Arista, 2016). Este plásmido se analizó por restricción para identificar las construcciones correctas de acuerdo con el patrón esperado, y confirmarlas mediante secuenciación en la Unidad de Biología Molecular del Instituto de Fisiología Celular, UNAM. A partir de la secuenciación, se confirmó que la construcción tenía el gen y la mutación correcta.

6.4 Construcción del gen Rv3763 con las mutaciones T34-36A, T40-41A y T34-36,40-41A y la etiqueta de seis histidinas

Inicialmente se intentaron introducir las mutaciones 3A2T, 3T2A y 3A2A mediante el método de QuikChange o mediante PCR. Sin embargo, no se pudieron obtener las mutaciones, a pesar de purificar los oligonucleótidos mediante electroforesis en geles de poliacrilamida y de realizar múltiples intentos modificando las temperaturas de alineamiento de estos, ya que la secuencia al sustituir los codones consecutivos de las treoninas por codones para alanina (GCC y GCG) resultó muy rica en G+C, lo que seguramente aumentaba el alineamiento inespecífico entre oligonucleótidos y templado. Por lo anterior, se decidió mandar sintetizar el fragmento del gen que contiene las mutaciones. Para esto, se utilizó la herramienta de síntesis *in vitro* de los fragmentos de la compañía GeneArt™ de Invitrogen™ (disponible en línea: <https://www.thermofisher.com/mx/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis.html>), que supera a la clonación tradicional por ser más rápida, precisa y efectiva ya que permite obtener la secuencia de interés de manera exacta. Con el fin de ahorrar costos, se mandó sintetizar solamente la secuencia desde el inicio del gen, el cual fue precedido de un sitio para la enzima *Bam*HI, hasta un sitio interno único para la enzima *Nru*I (Fig. 5).

BamHI NdeI
 1 GGATCCATATGAAGCGTGGACTGACGGTCGCGGTAGCCGGAGCCGCCATTCTGGTCGCAGGTCCTTT
 1 M K R G L T V A V A G A A I L V A G L S

 66 CCGGATGTTCAAGCAACAAGTCGACTACAGGAAGCGGTGAGACCACGACCGCGGCAGGCACGACGG
 21 G C S S N K S T T G S G E T T T A A G T T A

 133 CAAGCCCCGGCGCCGCTCCGGGCCGAAGGTCGTCATCGACGGTAAGGACCAGAACGTCACCGGCT
 43 S P G A A S G P K V V I D G K D Q N V T G S
NruI Asp718I
 181 CCGTGGTGTGCACAACCGCGGCCGGCAATGTCAACATCGCGAGGTACC
 65 V V C T T A A G N V N I A

Figura 5. Secuencia del fragmento del gen Rv3763 que se utilizó para la síntesis del gen con las sustituciones de treoninas por alaninas 3A2T, 3T2A y 3A2A donde se resaltan en verde los sitios de restricción *BamHI* y *Asp718I* que se utilizaron para la clonación en pTZ18r, en morado el sitio de *NdeI* utilizado para la clonación en el vector multicopia pIJ6021, en anaranjado el sitio interno único para la enzima *NruI* utilizado para la complementación del gen que incluye la etiqueta de histidinas. En recuadros rojos se marcan las sustituciones de treonina por alanina realizadas en los genes sintéticos, en un recuadro azul la serina que tiene un alto puntaje de glicosilación y que coincide con el motivo de glicosilación.

Posteriormente, se reconstruyó el gen completo con las diferentes mutaciones mediante la ligación de los fragmentos obtenidos a partir de la digestión con las enzimas *BamHI* y *Asp718I* de los genes sintéticos y el vector pTZ18r, y su subsecuente transformación en la cepa DH5 α y la selección de plásmidos recombinantes por alfa-complementación utilizando IPTG y X-gal.

Con el fin de añadir la etiqueta de seis histidinas en el extremo C-terminal, se realizó una doble digestión con *NruI* y *EcoRI* de las construcciones del gen Rv3763 con las diferentes mutaciones clonado en el vector pTZ18r y del plásmido pUC18 con el gen silvestre que contiene la etiqueta de seis histidinas proveniente de la cepa de *E. coli* IBEC58 debido a que la enzima *NruI* es sensible a metilación, los fragmentos obtenidos se ligaron con T4 DNA ligasa (Fig. 6).

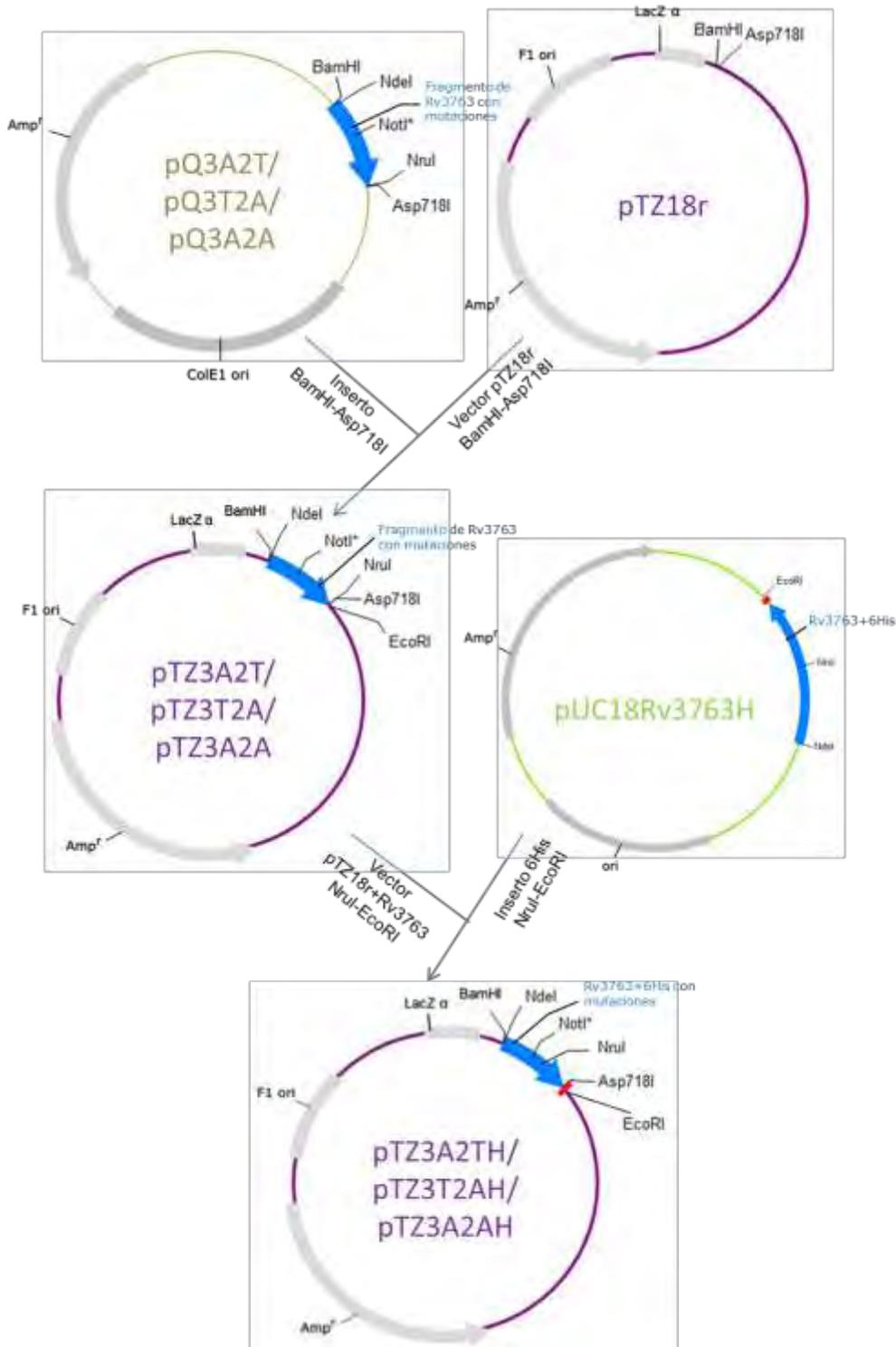


Figura 6. Construcciones realizadas para la complementación del gen Rv3763 con las mutaciones 3A2T, 3T2A y 3A2A con la etiqueta de seis histidinas. Se muestra en cada plásmido únicamente los sitios de restricción importantes para este trabajo; en verde oscuro el plásmido que contiene los genes sintéticos con las sustituciones de Thr por Ala, en morado el vector Ptz18r, en verde claro el plásmido pUC18 con el gen Rv3763 y la etiqueta de 6 His en rojo y en azul el gen Rv3763.

6.5 Clonación del gen Rv3763 con las mutaciones S43A, 3A2T, 3T2A y 3A2A en el vector multicopia pIJ6021

Debido a que *S. lividans* 1326 acepta con baja frecuencia el DNA metilado, se transformó por electroporación la cepa de *E. coli* IBEC58, que carece de los sistemas de metilación de DNA (González-Cerón *et al.*, 2009), con los distintos plásmidos que llevan al gen Rv3763 modificado por las diferentes mutaciones. Los plásmidos no metilados se purificaron y digirieron con *Nde*I y *Eco*RI, al igual que el vector pIJ6021. Los fragmentos correspondientes a las distintas versiones del gen Rv3763, así como el vector, se purificaron después de realizar electroforesis en geles de agarosa de bajo punto de fusión (LMP) al 0.6% y se realizó la reacción de ligación con T4 DNA ligasa toda la noche a temperatura ambiente.

6.6 Transformación de la cepa 1326 de *S. lividans* con los plásmidos pIJS43AH, pIJ3A2TH, pIJ3T2AH y pIJ3A2AH

Se transformaron protoplastos de *S. lividans* 1326 con los plásmidos pIJS43AH, pIJ3A2TH, pIJ3T2AH y pIJ3A2AH (Arista, 2016). De las colonias transformantes aisladas se tomaron seis y se homogenizaron por separado en 100 µl de agua para sembrarlas en sectores en cajas con medio sólido MS (manitol-soya) + kanamicina (50 µg/ml) que se incubaron a 30°C. Una vez esporuladas las cepas, se cosecharon y se inocularon en medio LB+azúcar al 25% con kanamicina y se realizaron minipreparaciones de DNA plasmídico de cada sector (Arista, 2016), el cual se analizó por restricción con la enzima *Kpn*I, con la que se esperaba un patrón de dos bandas, una de 2,269 pb y otra de 6,391 pb, mientras que el vector sin inserto contiene tres sitios de corte para la enzima, que dan tres bandas de 561 pb, 1,263 pb y 6,391 pb. Las construcciones correctas se reintrodujeron por transformación a *S. lividans* 1326 y se sembraron en cajas de MS con kanamicina para obtener crecimiento confluyente, dando lugar a un tapete de micelio esporulado después de una semana de incubación a 30°C. Posteriormente, las esporas se cosecharon con ayuda de agua destilada estéril,

se filtraron por algodón para eliminar los fragmentos de micelio, se centrifugaron a 3,500 rpm por 10 minutos y se lavaron con agua para ser resuspendidas en un volumen pequeño de glicerol al 20% y almacenadas a -20°C.

6.7 Expresión de la proteína LpqH

Se tomaron de 100 a 500 µl de una preparación densa de esporas para inocular 10 ml de medio YT2X, el cual fue incubado a 37° C en agitación durante 4 a 6 horas para lograr la germinación. Con las esporas germinadas se inocularon 100 ml de medio LB+sacarosa al 25% con kanamicina [50µg/ml], en un matraz de 500 ml y se incubaron durante 16 h a 30°C y en agitación constante. Posteriormente, para inducir la expresión de la proteína cuyo gen se encuentra bajo el control del promotor P_{tipA} en el vector pIJ6021, se añadió tioestreptón a una concentración final de 5 µg/ml y se continuó la incubación durante 36 a 40 h a 30°C con agitación.

Al terminar la incubación se centrifugó el cultivo a 8,000 rpm por 10 minutos y se almacenó el *pellet* a -70°C.

6.8 Fraccionamiento celular

La proteína LpqH es una lipoproteína que se encuentra anclada a la membrana, por lo que fue necesario separar las células del medio de cultivo, sonicar las células para lisarlas y fraccionar para separar los elementos de la membrana del citoplasma, pared celular y otros componentes celulares. Se lavó el *pellet* de micelio con 100 ml de Tris 20 mM, pH 8.0 y se incubó a -70°C durante 2 horas para facilitar el posterior rompimiento de las células. Se resuspendió en buffer para sonicar (Tris 50 mM, MgCl₂ 5 mM, DTT 4mM, pH 8.0, al que se añadió una pastilla inhibidora de proteasas (*complete* de Roche™) y 0.5 µl de RNAsa, y se sonicó con 30 pulsos de 20 segundos cada uno.

El extracto total se centrifugó a baja velocidad (7,000 rpm) por una hora a 4°C y el sobrenadante se ultracentrifugó a 33,000 rpm por una hora a 4°C. El *pellet* se lavó con Tris 20 mM, NaCl 50 mM, pH 8.0 con una pastilla inhibidora de

proteasas y se recentrifugó a 33,000 rpm por 30 minutos a 4°C. Finalmente el *pellet* se resuspendió en 1 ml de buffer para sonicar (Arista, 2016).

6.9 Purificación de proteínas por afinidad a níquel

Las diferentes versiones de la proteína LpqH estudiadas en este trabajo contenían una etiqueta de seis histidinas en el extremo C-terminal, lo que permitió su purificación por cromatografía de afinidad a níquel. Esto se realizó en un equipo de FPLC ÄktaPrime Plus, con una columna HisTrap HP de 1ml acoplada (GE Healthcare™) (López-Portillo, 2016).

Para inyectar la muestra en el equipo, la fracción membranal se solubilizó mediante tratamiento con el detergente LDAO al 2% durante 1 h en agitación constante a temperatura ambiente, seguido de centrifugación a 14,000 rpm por 30 min. El sobrenadante se colectó y se filtró a través de un filtro de 0.45 µm previo a la inyección en el equipo de FPLC (Arista, 2016). El buffer utilizado para la cromatografía fue Tris 20 mM, pH 8.0 con 20mM imidazol, y la proteína LpqH fue eluida con un gradiente de 20mM a 500 mM de imidazol en el mismo buffer.

6.10 Electroforesis de proteínas en gel de poliacrilamida-SDS (SDS-PAGE)

Las proteínas de las fracciones colectadas a partir de la purificación por FPLC se analizaron mediante electroforesis en geles desnaturalizantes de poliacrilamida con SDS (SDS-PAGE). La concentración de poliacrilamida fue 12% con un grosor de 1.5 mm. Se utilizó un marcador de peso molecular (Precision Plus Protein™ Dual Color Standards; BioRad Cat. #161-0374), un control positivo que es la proteína LpqH silvestre (purificada y proporcionada por el M. en C. Isaí Arista) y un control negativo que es la proteína LpqH silvestre purificada de una cepa de *S. lividans* mutante que carece de la O-manosiltransferasa de proteínas (Δpmt) y que, por lo tanto, es incapaz de glicosilar proteínas. Para su visualización las proteínas se tiñeron con azul brillante de Coomassie R-250 (López-Portillo, 2016).

6.11 Transferencia de proteínas a membrana PVDF

Para los ensayos de Western blot, las proteínas se transfirieron de los geles de poliacrilamida a membranas de PVDF (Immobilon P), en una cámara húmeda (Mini Trans-Blot® Electrophoretic Transfer Cell; BioRad™). La transferencia se realizó durante 14 horas a 12 mA en buffer Tris-Glicina (glicina 1.4%, Tris 0.3% y metanol 20%) (López-Portillo, 2016).

6.12 Inmunodetección de proteínas

Las membranas de PVDF con las proteínas inmovilizadas se bloquearon con TBS-Tween 20 al 0.05% en leche descremada en polvo al 5% por un mínimo de dos horas. Posteriormente se incubaron con anticuerpo primario anti-his (Pierce™ 6x-His Epitope Tag Antibody. Cat. #MAI-21315) en dilución 1:4,000 por 2 horas. Se lavó después la membrana con agua y con TBS-Tween 20 y se adicionó buffer de bloqueo con anticuerpo secundario anti-IgG de ratón acoplado con peroxidasa (Sigma™ Anti- Mouse IgG- (whole molecule)- Peroxidase Antibody. Cat. #A4416) en dilución 1:1,000. Se concluyó el ensayo utilizando el Kit de quimioluminiscencia (BM Chemiluminescence Blotting Substrate (POD) de Roche™) para revelar y detectar las proteínas (Arista, 2016).

6.13 Determinación del estado de glicosilación de LpqH por reconocimiento con concanavalina A

Con el fin de detectar la glicosilación de las proteínas, se bloquearon las membranas de PVDF con TBS-BSA (albúmina sérica bovina, fracción V) al 1% durante 2 horas. Posteriormente, se incubó con TBS adicionado con concanavalina A acoplada con peroxidasa (Sigma™ Concanavalina A-Peroxidasa. Cat. #L6397), en una dilución 1:1,000, con 1 mM de MgCl₂, 1 mM de CaCl₂ y 1 mM de MnCl₂ durante 14 horas y se lavó con TBS 1X. Finalmente, la detección se realizó con el kit de quimioluminiscencia (BM

Chemiluminescence Blotting Substrate (POD) de Roche™) para revelar y detectar las proteínas (Arista, 2016).

7. Resultados y discusión

7.1 Identificación y selección de probables sitios de glicosilación en LpqH

Previamente, López-Portillo (2016) realizó una búsqueda con el motivo completo proveniente del análisis de los sitios de O-manosilación de glicoproteínas en *M. tuberculosis* (Smith *et al.*, 2014) sobre todas las lipoproteínas del genoma de *Streptomyces coelicolor*, con el fin de intentar predecir las proteínas glicosiladas de este organismo. Sin embargo, no se encontraron proteínas que mostraran coincidencias con el motivo completo, aunque, sí se encontraron coincidencias cuando el motivo se recortaba a solamente seis aminoácidos a cada lado del sitio probable de glicosilación. Con el fin de determinar los sitios con mayor probabilidad de ser glicosilados en la proteína LpqH de *M. tuberculosis*, se realizó un análisis de la secuencia de aminoácidos de esta proteína utilizando el *software* “MOTIF search” (disponible en línea: <http://www.genome.jp/tools/motif/>). Para hacer la búsqueda se utilizó el mismo motivo previamente recortado:

[AIPSVDEGKLM]-[AGSDPTEIV]-[APTDEGIKLSV]-[TAGPVIMQR]-
[TAPSDKLMV]-[APTDHKQ]-[TS]-[PDTAMN]-[TDPQRS AFGIV]-
[PTDKVAENQRS]-[APSEGINQRT]-[TAIPCDFGLQRSY]-[PRTADEFHMN]

Al igual que en el trabajo de López Portillo (2016), este motivo está basado en el reportado por Smith *et al.*, (2014), limitándolo solamente a seis aminoácidos a cada lado del probable sitio de glicosilación (marcado con negritas). MOTIF search no fue capaz de detectar con el motivo recortado un probable sitio de glicosilación sobre la secuencia de la proteína LpqH, por lo que tuvo que recortarse todavía más, a solamente cinco posiciones a cada lado de la posición

central de serina o treonina. Esto permitió encontrar probables sitios de glicosilación en la secuencia de LpqH (Fig. 7), quedando de la siguiente manera:

[AGSDPTEIV]-[APTDEGIKLNVS]-[TAGPVIMQR]-[TAPSDKLMV]-[APTDHKQ]-
 [TS]-[PDTAMN]-[TDPQRSAFGIV]-[PTDKVAENQRS]-[APSEGINQRT]-
 [TAIPCDFGLQRSY]



Figura 7. Motivo de glicosilación recortado. El motivo de O-glicosilación indica una alta probabilidad de prolina y alanina hidrofóbicas intercaladas con algunas treoninas hidrofílicas alrededor del sitio de glicosilación. En el recuadro se observan las 11 posiciones utilizadas para generar el motivo con el cual se detectan posibles sitios de glicosilación en LpqH. Modificado de Smith *et al.* (2014).

Con este motivo de once aminoácidos se encontró la secuencia AGTTASPGAAS, correspondiente a las posiciones 38-48 cercanas al extremo N-terminal de la proteína madura, colocando a la serina central de la secuencia (Ser₄₃) como el sitio de glicosilación más probable en la proteína LpqH. Es importante mencionar que el uso del motivo recortado no encontró a los aminoácidos que habían sido reportados con anterioridad como probables sitios de glicosilación. Es decir, las treoninas localizadas entre las posiciones 33 y 41. Adicionalmente al análisis con el programa MOTIF, se analizó la secuencia de la proteína madura utilizando el *software* de predicción de glicosilación NetOGlyc 3.1, posteriormente, con el *software* NetOGlyc 4.0. En NetOGlyc 3.1 el puntaje-G pertenece al mejor predictor general mientras que el puntaje-I es el del mejor predictor del sitio aislado. Si el puntaje-G es >0.5, el residuo se predice como

glicosilado, cuanto más alto sea, más confiable será la predicción. En el caso de las treoninas, se debe considerar adicionalmente el puntaje-I, si el puntaje-G es <0.5 , si el puntaje-I es >0.5 y no hay sitios aledaños predichos (con una distancia de menos de 10 residuos), también se predice como un residuo glicosilado. En NetOGlyc 4.0 sólo los sitios con puntajes superiores a 0.5 se predicen como glicosilados y se anotan como "#POSITIVO" en los comentarios (Steentoft *et al.*, 2013).

Con las dos herramientas se encontraron 13 probables sitios de glicosilación diferentes, con distintos puntajes en ambos servidores (Fig. 8). Se encontró paridad entre algunos de los puntajes más altos con la probable secuencia de glicosilación AGTTASPGAAS. En NetOGlyc 3.1 se encontró paridad entre el motivo y los residuos Thr₄₀ y Thr₄₁, cuyos puntajes-G van de 0.694 a 0.697; Ser₄₃ y Ser₄₈ con puntajes de 0.565 y 0.522, respectivamente. En NetOGlyc 4.0 se encontró paridad con los mismos residuos, pero con puntajes distintos, Thr₄₀ con 0.936417, Thr₄₁ con 0.952175, Ser₄₃ con 0.90923 y Ser₄₈ con 0.5 (Fig. 8). Es importante mencionar que, aunque tanto NetOGlyc 3.1 como NetOGlyc 4.0 predicen a la Ser₄₃ como sitio de glicosilación, existen otros aminoácidos cercanos con un puntaje mayor, como son las Thr₄₀ y Thr₄₁, aunque estos residuos no hubieran sido predichos directamente por el programa MOTIF como el sitio donde ocurre la glicosilación.

El análisis con los predictores NetOGlyc mostró que también existen otros sitios probables de glicosilación. Los puntajes más altos corresponden a Thr₃₄ (0.703), Thr₃₅ (0.741) y Thr₃₆ (0.728) y son incluso mayores a los de Ser₄₃. Adicionalmente, hay otros residuos con puntajes mayores a 0.5 de acuerdo con NetOGlyc 3.1, como son Ser₂₇ (0.501), Thr₂₈ (0.623), Thr₂₉ (0.661), Ser₃₁ (0.593), Thr₁₅₇ (0.593) y Ser₁₅₉ (0.541). Por otro lado, con NetOGlyc 4.0 tienen puntajes notablemente altos los residuos Ser₂₇ (0.801881), Thr₂₈ (0.928446), Thr₂₉ (0.832214), Ser₃₁ (0.937942), Thr₃₄ (0.937693), Thr₃₅ (0.985754) y Thr₃₆ (0.979309), mientras que Thr₁₅₇ y Ser₁₅₉ ya no se predicen como sitios de glicosilación con esta versión del *software*. Adicionalmente, se encuentran los residuos Ser₂₃ (0.521305) y Ser₂₄ (0.6097221) con puntajes relativamente bajos (Fig. 8)

En el estudio de mutagénesis de la proteína LpqH de Hermann *et al.* (1996) se describieron como sitios probables de glicosilación a las cinco treoninas que se encuentran en las posiciones 34-36 y 40-41. Dentro del motivo de glicosilación únicamente se encuentran los residuos Thr₄₀ y Thr₄₁ que tienen una alta probabilidad de glicosilación. Sin embargo, el llamado triplete de treoninas (Thr₃₄₋₃₆) que tiene puntajes-G mayores a 0.7 en NetOGlyc 3.1 y en NetOGlyc 4.0 de 0.924358 (Thr₃₄), 0.985754 (Thr₃₅) y 0.979309 (Thr₃₆), notablemente más altos que los del doblete de treoninas (Thr₄₀₋₄₁), no se encuentra dentro del motivo de glicosilación acotado. El análisis mutacional que se realiza en este trabajo podría ayudar a vislumbrar si el motivo de glicosilación descrito por Smith *et al.* (2014) que se propone para describir los sitios de glicosilación en proteínas de Mtb es realmente eficiente para predecir sitios de glicosilación, ya que los puntajes de NetOGlyc parecen indicar que la glicosilación puede darse con mayor probabilidad en otros sitios fuera del logo de secuencia, y si estos sitios con alta probabilidad de glicosilación que se ajustan al motivo realmente participan en la O-manosilación de la proteína.

S/T	Pos	G-score	I-score	Y/N	feature	start	end	score	strand	frame	comment
T	6	0.414	0.065	.	CARBOHYD	6	6	0.0660013	.	.	
S	20	0.393	0.040	.	CARBOHYD	20	20	0.0733455	.	.	
S	23	0.441	0.028	.	CARBOHYD	23	23	0.521305	.	.	#POSITIVE
S	24	0.439	0.045	.	CARBOHYD	24	24	0.609721	.	.	#POSITIVE
S	27	0.501	0.019	5	CARBOHYD	27	27	0.801881	.	.	#POSITIVE
T	28	0.623	0.052	T	CARBOHYD	28	28	0.928446	.	.	#POSITIVE
T	29	0.661	0.084	T	CARBOHYD	29	29	0.832214	.	.	#POSITIVE
S	31	0.593	0.039	5	CARBOHYD	31	31	0.937942	.	.	#POSITIVE
T	34	0.703	0.023	T	CARBOHYD	34	34	0.937693	.	.	#POSITIVE
T	35	0.741	0.039	T	CARBOHYD	35	35	0.988373	.	.	#POSITIVE
T	36	0.718	0.147	T	CARBOHYD	36	36	0.982438	.	.	#POSITIVE
T	40	0.694	0.079	T	CARBOHYD	40	40	0.04712	.	.	#POSITIVE
T	41	0.697	0.511	T	CARBOHYD	41	41	0.962862	.	.	#POSITIVE
S	43	0.565	0.095	5	CARBOHYD	43	43	0.925178	.	.	#POSITIVE
S	48	0.522	0.043	5	CARBOHYD	48	48	0.533723	.	.	#POSITIVE
T	62	0.423	0.372	.	CARBOHYD	62	62	0.0686362	.	.	
S	64	0.270	0.040	.	CARBOHYD	64	64	0.0888868	.	.	
T	68	0.383	0.170	.	CARBOHYD	68	68	0.192009	.	.	
T	69	0.371	0.316	.	CARBOHYD	69	69	0.138639	.	.	
T	83	0.465	0.036	.	CARBOHYD	83	83	0.0908244	.	.	
T	90	0.439	0.200	.	CARBOHYD	90	90	0.0507093	.	.	
S	99	0.321	0.079	.	CARBOHYD	99	99	0.0431421	.	.	
T	109	0.473	0.051	.	CARBOHYD	109	109	0.263416	.	.	
T	113	0.387	0.072	.	CARBOHYD	113	113	0.218720	.	.	
S	114	0.318	0.051	.	CARBOHYD	114	114	0.299001	.	.	
T	116	0.362	0.117	.	CARBOHYD	116	116	0.216983	.	.	
S	122	0.389	0.032	.	CARBOHYD	122	122	0.154172	.	.	
T	124	0.497	0.077	.	CARBOHYD	124	124	0.158797	.	.	
S	128	0.383	0.017	.	CARBOHYD	128	128	0.121531	.	.	
T	133	0.498	0.046	.	CARBOHYD	133	133	0.112528	.	.	
T	135	0.467	0.099	.	CARBOHYD	135	135	0.0807845	.	.	
T	137	0.486	0.030	.	CARBOHYD	137	137	0.147182	.	.	
S	146	0.363	0.045	.	CARBOHYD	146	146	0.0880212	.	.	
S	151	0.419	0.069	.	CARBOHYD	151	151	0.0142571	.	.	
T	157	0.593	0.053	T	CARBOHYD	157	157	0.0216825	.	.	
S	159	0.541	0.042	5	CARBOHYD	159	159	0.00520219	.	.	

Figura 8. Análisis de la proteína LpqH con la herramienta NetOGlyc. A la izquierda la versión 3.1; a la derecha la versión 4.0. En un cuadro verde se encierra el triplete de treoninas (Thr₃₄₋₃₆),

en un cuadro azul el doblete (Thr₄₀₋₄₁), en rojo la serina Ser₄₃; todas éstas con altos puntajes. Entre corchetes se encierra el *cluster* de 5 treoninas con alta probabilidad de glicosilación.

7.2 Mutagénesis múltiple del gen Rv3763, para introducir los cambios S43A, T34-36A, T40-41A y T34-36,40-41A en la proteína LpqH

Como se mencionó en la sección anterior, el único sitio predicho por el motivo de glicosilación modificado de Smith *et al.* (2014) corresponde a la Ser₄₃. Adicionalmente se tiene el antecedente de que cinco treoninas (Thr₃₄₋₃₆, Thr_{40,41}) podrían participar en la glicosilación de LpqH (Hermann *et al.*, 1996), aunque a la fecha en que se inició el presente trabajo no se había demostrado que estuvieran glicosiladas. Una posibilidad es que estas cinco treoninas fueran importantes sólo como parte del dominio de glicosilación, pero que el aminoácido glicosilado fuera únicamente la Ser₄₃. Otra posibilidad es que efectivamente una o varias de las treoninas estuvieran glicosiladas. En este sentido es importante mencionar que el estudio proteómico de Smith *et al.* (2014) mostró que la mayoría de los sitios de glicosilación corresponden a treoninas. Por ejemplo, en el único estudio previo donde se analizó una glicoproteína en detalle (Dobos *et al.*, 1995), sólo se identificaron treoninas como sitios de glicosilación en el antígeno de *M. tuberculosis* de 45 kDa. Con el fin de resolver si solamente la Ser₄₃, que es predicha por el motivo de glicosilación y tiene un puntaje alto en NetOGlyc, es el único sitio de glicosilación; se decidió realizar mutagénesis de sitio dirigido dentro del gen Rv3763 mediante QuikChange, con el fin de cambiar este residuo por una alanina que tiene una estructura similar que no afecta el plegamiento de la proteína, es decir, un cambio LpqHS43A.

Adicionalmente, y con el fin de resolver si las treoninas se encuentran también glicosiladas, se realizaron mutaciones múltiples en el gen Rv3763, para cambiar el doblete de treoninas Thr₄₀₋₄₁ por alaninas (a esta le llamamos LpqH3T2A), el triplete de treoninas Thr₃₄₋₃₆ por alaninas (LpqH3A2T) y las cinco treoninas por alaninas (LpqH3A2A). Para este proceso se mandaron sintetizar los fragmentos que contienen las mutaciones, que comprenden desde el codón de inicio del gen Rv3763 hasta el sitio único interno de *NruI*, posteriormente, se regeneró el gen completo en todas las versiones mutantes (ver Materiales y Métodos). Todas las

versiones, tanto la versión silvestre de la proteína LpqH como las versiones mutantes, poseen una etiqueta de seis histidinas en el extremo carboxilo terminal (Fig. 9).

Todas estas construcciones fueron clonadas en el vector de expresión pIJ6021, y se secuenciaron con los oligonucleótidos 6021*Nde*I y 6021*Eco*RI (Tabla 3) para confirmar que la secuencia fuera correcta.



Figura 9. Secuencia del gen Rv3763 con el péptido señal subrayado en verde, el lipobox encerrado en morado, los residuos predichos por ambas versiones de NetOGlyc encerrados en azul, los residuos de treonina y serina seleccionados para la mutagénesis señalados con flechas y la etiqueta de histidinas subrayada en naranja.

7.3 Expresión y glicosilación de la proteína LpqH silvestre y sus variantes

Para los experimentos de expresión de las diferentes formas de la proteína LpqH, se utilizó la cepa silvestre 1326 de *S. lividans*, la cual se transformó con los plásmidos recombinantes. Se obtuvieron preparaciones densas de esporas que se utilizaron para inocular cultivos de 400 ml de medio LB líquido con sacarosa al 25%; la expresión de las diferentes versiones del gen Rv3763 se indujo con tioestreptón, y después de 36 horas se colectó el micelio por centrifugación,

posteriormente, éste se lisó por sonicación, para obtener el extracto total. A partir de este se obtuvo la fracción membranal por ultracentrifugación como se describe en materiales y métodos, ya que en dicha fracción es donde se encuentra LpqH, al tratarse de una lipoproteína. A partir de la fracción membranal, las diferentes versiones de la proteína LpqH fueron purificadas por cromatografía de afinidad a níquel en un equipo de FPLC Äkta Prime plus. En todos los casos se seleccionó la fracción que se encontraba más enriquecida en proteína LpqH. Las diferentes versiones puras de la proteína LpqH (S43A, 3A2T, 3T2A, 3A2A) se cuantificaron y se analizaron 2 µg de cada muestra en un gel desnaturante de poliacrilamida al 12% con SDS, el cual se tiñó con azul de Coomassie. Adicionalmente, se incluyeron en el gel tanto la proteína silvestre purificada de la cepa *S. lividans* 1326, como control positivo, y la proteína silvestre pero purificada de una mutante Δpmt de *S. lividans* como control negativo, ya que dicha mutante es incapaz de glicosilar proteínas (Fig. 10). Ambas proteínas fueron proporcionadas por el M. en C. Isaí Arista.

Se puede observar que la proteína LpqH expresada a partir de la cepa de *S. lividans* Δpmt , que es incapaz de glicosilar proteínas y tiene un peso molecular aparente menor al de las otras versiones de la proteína (Fig. 10, carril 6). Ésta corresponde a la masa teórica de la proteína sin péptido señal (13.2 kDa), la cual aumenta aproximadamente 800 Da adicionales por la presencia de los lípidos: un diacilglicerol y una cadena de acilo adicional (Arista, 2016). Por otro lado, la proteína LpqH silvestre muestra un peso molecular aparente mayor a la versión no glicosilada (Fig. 10, carril 1), lo que se puede atribuir a la presencia de azúcares (Arista, 2016). Es importante observar que la proteína LpqH silvestre y la LpqHS43A comigran, lo que indica que S43A se glicosila de manera prácticamente igual a la proteína silvestre (Fig. 10, carril 2). La proteína LpqH3T2A (Fig. 10, carril 4) se observa con un peso molecular ligeramente menor que el de LpqH silvestre pero mayor a la proteína sin glicosilar. Esto indica que, aunque existe glicosilación en la proteína, ésta es menor que la silvestre. A su vez, esto sugiere un menor número de residuos de manosa. En el caso de la proteína LpqH3A2T (Fig. 10, carril 3) la migración es intermedia entre LpqH3T2A y LpqH sin glicosilar, lo que indica que sí está glicosilada, pero con un menor número de manosas que la LpqH3T2, por lo tanto, que la proteína LpqH silvestre. La proteína LpqH con las cinco treoninas sustituidas por alaninas, es decir la

versión LpqH3A2A (Fig. 10 carril 5) tiene un corrimiento similar a la proteína sin glicosilar. Lo anterior sugiere que la proteína LpqH se glicosila principalmente en las treoninas, y que la ausencia de la Ser₄₃ prácticamente no afecta la glicosilación. Sin embargo, para poder demostrar las diferencias en glicosilación directamente, se optó por realizar un ensayo de Western blot detectando la presencia de manosas con la lectina concanavalina A (ConA).

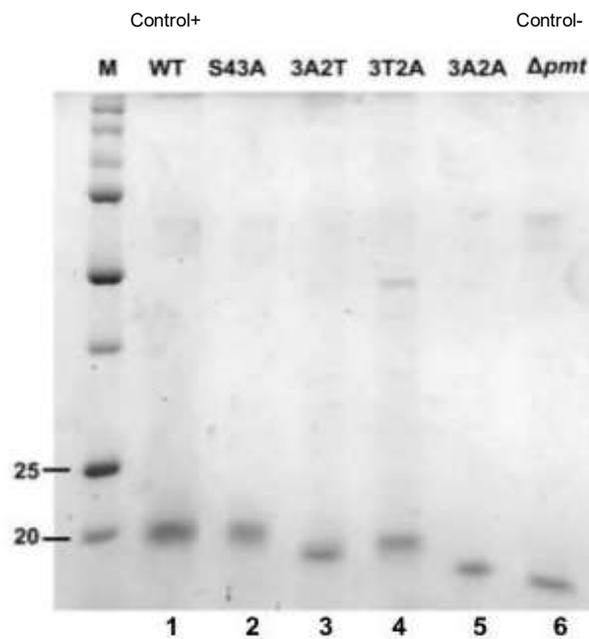


Figura 10. SDS-PAGE teñido con azul de Coomassie de la proteína LpqH, donde se muestra la diferencia de peso molecular de las proteínas LpqH que contienen la etiqueta de 6 histidinas expresadas en *S. lividans*: LpqH silvestre control positivo (WT) (carril 1), LpqHS43A (carril 2), LpqH3A2T (carril 3), LpqH3T2A (carril 4), LpqH3A2A (carril 5) y LpqH purificada de la cepa mutante Δpmt control negativo (carril 6). "M" indica el marcador de peso molecular.

Adicionalmente, para confirmar que las proteínas purificadas por cromatografía de afinidad a níquel (que se muestran en la Fig. 10) efectivamente corresponden a la proteína LpqH, se realizó un inmunoblot usando un anticuerpo comercial anti-histidinas (Fig. 11). Como se observa, este experimento confirmó que las bandas efectivamente corresponden a las diferentes versiones de la proteína LpqH.

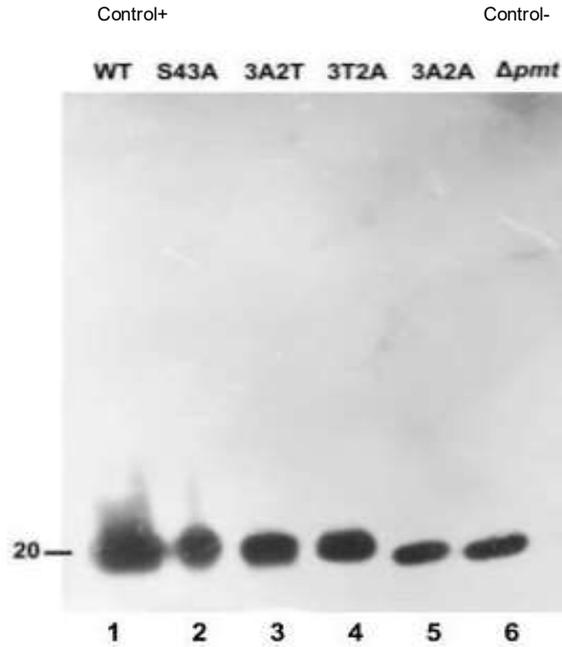


Figura 11. Western blot con anticuerpos anti-Histidinas de la proteína LpqH, donde se muestran las proteínas purificadas que contienen la etiqueta de 6 histidinas: LpqH silvestre control positivo (WT) (carril 1), LpqHS43A (carril 2), LpqH3A2T (carril 3), LpqH3T2A (carril 4), LpqH3A2A (carril 5) y LpqH purificada de la cepa Δpmt control negativo (carril 6).

Finalmente, en la Figura 12 se presenta un Western blot donde las proteínas fueron detectadas mediante la lectina ConA, donde es posible observar que las sustituciones individuales S43A, 3A2T y 3T2A (Fig. 12, carriles 2, 3 y 4; respectivamente) no eliminan la unión de la ConA. Esto indica que dichas versiones de la proteína LpqH se siguen glicosilando, aunque aparentemente las proteínas LpqH3A2T y LpqH3T2A lo hagan en un menor grado, debido a las diferencias observadas en el corrimiento electroforético en comparación con LpqH silvestre (Fig. 10, carriles 3 y 4).

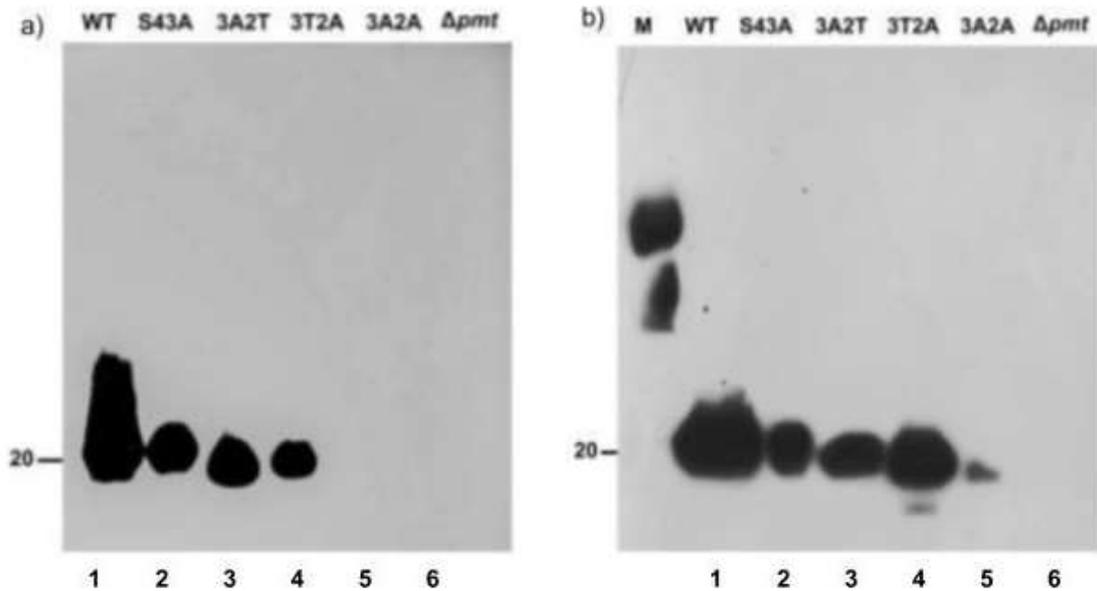


Figura 12. Western blot con ConA de la proteína LpqH, a) Placa con un tiempo de exposición de aproximadamente 2 minutos. b) Placa sobreexpuesta con un tiempo de exposición de aproximadamente 10 minutos. Se muestran las diferentes versiones de la proteína LpqH: LpqH silvestre (WT) (carril 1), LpqHS43A (carril 2), LpqH3A2T (carril 3), LpqH3T2A (carril 4), LpqH3A2A (carril 5), LpqH purificada de la cepa Δpmt (carril 6). “M” indica el marcador de peso molecular, en el cual la ovoalbúmina es detectada también por la ConA.

La proteína LpqH3A2A, que carece de las cinco treoninas no fue detectada en el Western blot con ConA, cuando la placa se expuso el tiempo de exposición necesario para la detección de las otras versiones de la proteína (Fig. 12a, carril 5). Esto evidencia que prácticamente todos los azúcares son unidos a las treoninas que fueron eliminadas en la versión 3A2A. Sin embargo, en la placa con un tiempo de exposición prolongado sí se detectó una muy leve señal (Fig. 12b, carril 5) que no fue detectada en la proteína purificada de la cepa Δpmt . Esto podría indicar que sí hay sitios de glicosilación presentes que no corresponden a las cinco treoninas que fueron eliminadas. Pudiera tratarse de la Ser₄₃, o bien, de otros sitios fuera del motivo de glicosilación que podrían ser aquellos que tuvieron un puntaje alto en NetOGlyc, tal como S₃₁ que está cerca del motivo de glicosilación y tiene una probabilidad muy alta de estar glicosilada, aún mayor a la de T₃₄ o a la de S₄₈ o S₄₃, que, a pesar de estar dentro del motivo, tiene un puntaje bajo. Sin embargo, es evidente que la glicosilación ocurre con mucho menor probabilidad en dicho sitio, ya que la reacción es muy débil.

Al iniciar este trabajo, la única información sobre los sitios de glicosilación en la proteína LpqH correspondían al trabajo de Hermann *et al.* (1996), donde no se hicieron diferentes mutaciones a los *clusters* de treoninas cerca del extremo N-terminal, que corresponden a los analizados en el presente trabajo. Sin embargo, no era claro si el efecto observado se debía a que dichos aminoácidos eran glicosilados directamente, o si más bien alteraban la glicosilación indirectamente. Esta última posibilidad se hacía más relevante en función del estudio glicoproteómico de Smith *et al.* (2014) en donde se obtuvo un consenso de los aminoácidos que rodean al sitio de glicosilación, muchos de los cuales son a su vez treoninas. Recientemente, Parra *et al.* (2017) publicó un estudio muy detallado en el que se analiza mediante espectrometría de masas de alta resolución a la proteína LpqH purificada y expresada a partir de *M. smegmatis*. En dicho trabajo se demuestra claramente que, la glicosilación de la lipoproteína LpqH en *M. smegmatis* ocurre secuencialmente en cuatro de los cinco sitios de glicosilación sugeridos previamente en el trabajo de Hermann *et al.* (1996) que son Thr₃₄, Thr₃₅, Thr₃₆ y Thr₄₁; y en dicho trabajo también se sugiere que la glicosilación no se distribuye al azar, sino que es progresiva debido a la accesibilidad desigual de los sitios de glicosilación para la Pmt. La disminución de peso molecular de 1-2 kDa observada en el SDS-PAGE de la proteína 3A2A (Fig. 10, carril 5) y la prácticamente nula reacción con la ConA de la misma (Fig. 12a, carril 5) es consistente con la pérdida de residuos de carbohidratos asociados a los *clusters* de treoninas mutadas. Asimismo, en el estudio de Parra *et al.* (2017) se observó que Thr₄₁ (que fue sustituida por alanina en el presente trabajo) tiene la mayor cantidad de manosas asociadas respecto a las otras tres treoninas confirmadas por espectrometría de masas y es un sitio preferencial para la glicosilación de LpqH. Es importante mencionar que, aunque Thr₄₀ tiene una alta probabilidad de ser glicosilada, de acuerdo con NetOGlyc y que en el trabajo de Hermann *et al.* (1996) se consideraba como un residuo que era importante para la glicosilación, no es glicosilado en *M. smegmatis* de acuerdo con el análisis con espectrometría de masas realizado por Parra *et al.* (2017). Lo anterior refleja la dificultad de predecir los sitios de O-glicosilación en las proteínas. En este sentido se puede argumentar, bajo estos fundamentos, que el motivo propuesto por Smith *et al.* (2014) no puede ser utilizado para la predicción efectiva de sitios de glicosilación en las proteínas de Mtb, a pesar de

que se deriva de un estudio glicoproteómico detallado de las glicoproteínas presentes en filtrados de cultivo de dicho organismo con afinidad a ConA. Como perspectiva, sería importante realizar un análisis detallado con espectrometría de masas de la proteína LpqH silvestre expresada en *S. lividans* para confirmar con precisión los residuos de aminoácidos en donde ocurre la glicosilación y observar la cantidad de manosas que hay en cada uno de ellos. Sin embargo, a pesar de no contar con dicha información todavía, sí es posible afirmar que la enzima Pmt de *S. lividans* al menos sí es capaz de reconocer los mismos *clusters* de treoninas dentro de la proteína LpqH que reconoce la Pmt de *M. tuberculosis* para llevar a cabo la glicosilación. Esto es consistente con la observación previa de Lara *et al.* (2004) sobre la glicosilación que presenta la proteína Apa de Mtb cuando es expresada en *S. lividans*.

8. Conclusiones

- El motivo propuesto por Smith *et al.* (2014) carece de poder predictivo para la identificación efectiva de sitios de glicosilación en las proteínas de Mtb.
- La glicosilación de la proteína LpqH al ser expresada en *S. lividans* ocurre en los *clusters* de treonina conformados por Thr₃₄, Thr₃₅, Thr₃₆, Thr₄₀ y Thr₄₁.
- El residuo Ser₄₃ no es el sitio principal de glicosilación de LpqH expresada en *S. lividans*.
- Hay uno o varios residuos de serina o treonina aparte de Thr₃₄, Thr₃₅, Thr₃₆, Thr₄₀ y Thr₄₁ que pueden glicosilarse, pero con muy baja eficiencia.

9. Referencias

- Agilent Technologies. (2015). *QuikChange Site-Directed Mutagenesis Kit*. Revision D.01. Agilent Technologies, Inc. 1-18.
- Andersen, T.M. (2005). *The success and failure of BCG-implications for a novel tuberculosis vaccine*. *Nat Rev Microbiol*, 3 (8), 656–662.
- Arista, I. (2016). *Caracterización de lipoproteínas de Mycobacterium tuberculosis producidas en Streptomyces lividans* (Tesis de maestría), Universidad Nacional Autónoma de México. México.
- Berks, B. (2000). *The bacterial Tat protein export pathway*. *Biochim. Soc. Trans*, 28 (3), 58.2- 58.3.
- Brennan, P.J. (1989). *Structure of mycobacteria: recent developments in defining cell wall carbohydrates and proteins*. *Rev. Infect. Dis*, 11(2), 420–430.
- Chater, K., Biró, S., Lee, K.J., Palmer, T., Schrempf, H. (2010). *The complex extracellular biology of Streptomyces*. *FEMS Microbiol Rev*, 34 (2), 171-198.
- Corbière, V., van Els, C., Smits, K., van Gaans-van den Brink, J., Poelen, M., Mascart, F., Meiring, H., Locht, C. (2014). *Toward understanding the essence of post-translational modifications for the Mycobacterium tuberculosis immunoproteome*. *Frontiers in immunology*, 1-2.
- Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E. (2004). *WebLogo: A sequence logo generator*. *Genome Research*, 14, 1188-1190.
- Dobos, K. M., Swiderek, K., Khoo, K. H., Brennan, P. J., & Belisle, J. T. (1995). *Evidence for glycosylation sites on the 45-kilodalton glycoprotein of Mycobacterium tuberculosis*. *Infection and Immunity*, 63 (8), 2846–2853.
- Flärdh, K. & Buttner, M.J. (2009). *Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium*. *Nature reviews*, 7, 36-49.
- Forrellad, M.A., Klepp, L.I., Gioffré, A., Sabio y García, J., Morbidoni, H.R., Santangelo, M.P., Bigi, F. (2013). *Virulence factors of the Mycobacterium tuberculosis complex*. *Virulence*, 4 (1), 3–66.

- González-Cerón, G., Miranda-Olivares, O., Servín-González, L. (2009). *Characterization of the methyl-specific restriction system of Streptomyces coelicolor A3 (2) and of the role played by laterally acquired nucleases*. FEMS Microbiology Letters, 301 (1), 35-43.
- González-Zamorano, M., Mendoza-Hernández, G., Xolalpa, W., Parada, C., Vallecillo, A.J., Bigi, F., Espitia, C. (2009). *Mycobacterium tuberculosis Glycoproteomics Based on ConA-Lectin Affinity Capture of Mannosylated Proteins*. J. Proteome Res., 8 (2), 721–733.
- Goosens V.J., Monteferrante C.G., van Dijk J.M. (2014). *The Tat system of Gram- positive bacteria*. Biochim Biophys Acta. 1843(8):1698-706.
- Grant, S.G., Jesse, J., Bloom, F.R., Hannan, D. (1990). *Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants*. Proc. Natl. Acad. Sci, 42, 46-51.
- Henao-Tamayo, M., Junqueira-Kipnis, A.P., Ordway, D., Gonzales-Juarrero, M., Stewart, G.R., Young, D.B. (2007) *A mutant of Mycobacterium tuberculosis lacking the 19-kDa lipoprotein Rv3763 is highly attenuated in vivo but retains potent vaccinogenic properties*. Vaccine, 25, 7153–7159.
- Hermann, J.L., Gaora, R.O., Gallagher, A., Thole, J.E.R., Young, D.B. (1996). *Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from Mycobacterium tuberculosis*. The EMBO Journal, 15 (14), 3547-3554.
- Jiansong, T., Wing-Cheong, Y., Zhiwei, C. (2016). *Mycobacterium tuberculosis infection and vaccine development*. Tuberculosis. 98, 30-41.
- Kamalakkannan, S., Murugan, V., Jagannadham, M., Nagaraj, R., Sankaran, K. (2004). *Bacterial lipid modification of proteins for novel protein engineering applications*. Protein Engineering Design And Selection, 17, 10.
- López Portillo, M. (2016). *Predicción y expresión de lipoproteínas glicosiladas de Streptomyces coelicolor, incluyendo una peptidil-prolil isomerasa de la familia FKBP*. (Tesis de licenciatura). Universidad Nacional Autónoma de México.
- Mahenthalingam, E., Marklund, B.I., Brooks, L.A., Smith, D.A., Bancroft, G.J., Stokes, R.W. (1998). *Site-Directed Mutagenesis of the 19-Kilodalton*

Lipoprotein Antigen Reveals No Essential Role for the Protein in the Growth and Virulence of Mycobacterium intracellulare. Infection and Immunity, 66 (8), 3626–3634.

- McMurray, D.N. (1996) Cap. 33. *Mycobacteria and Nocardia*. En: Baron S, editor. (4^a ed.) Medical Microbiology. Galveston (TX): University of Texas Medical Branch at Galveston, 33.
- Mead, D.A., Szczesna-Skorupa, E., Kemper, B. (1986). *Single-stranded DNA blue T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering*. Protein Engineering, 1, 67-74.
- Ocampo, M., Curtidor, H., Vanegas, M., Patarroyo, M. A., Patarroyo, M. E. (2014). *Specific interaction between Mycobacterium tuberculosis lipoprotein-derived peptides and target cells inhibits mycobacterial entry in vitro*. Chemical Biology & Drug Design, 84 (6), 626–641.
- Parra, J., Marcoux, J., Poncin, I., Canaan, S., Herrmann, J.L., Nigou, J., Burlet-Schiltz, O., Rivière, M. (2017). *Scrutiny of Mycobacterium tuberculosis 19 kDa antigen proteoforms provides new insights in the lipoglycoprotein biogenesis paradigm*. Sci. Rep, 7, 43682.
- Peter-Katalinić, J. (2005) *Methods in Enzymology: O-Glycosylation of Proteins*. Methods Enzymol, 405, 139-71.
- Rebrova E.M. (2017). *Cell wall remodeling proteins in Mycobacterium tuberculosis: structure, function and inhibition*. Karolinska Institutet, Solna, 14.
- Schneider, T., Stephens, R.M. (1990). *Sequence Logos: A New Way to Display Consensus Sequences*. Nucleic Acids Res, 18, 6097-6100.
- Smith, G.T., Sweredoski, M.J., Hess, S. (2014). *O-linked glycosylation sites profiling in Mycobacterium tuberculosis culture filtrate proteins*. J. Proteomics, 97, 296–306.
- Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Bennett, E.P., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., Clausen, H. (2013). *Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology*. The EMBO Journal, 32, 1478-1488.

- Sutcliffe, I.C., Harrington, D.J., (2004). *Lipoproteins of Mycobacterium tuberculosis: an abundant and functionally diverse class of cell envelope components*. FEMS: Microbiology reviews, 28 (5), 645-659.
- Takano, E., White, J., Thompson, C.J., Bibb, M.J. (1995). *Construction of thiostrepton-inducible, high-copy-number expression vectors for use in Streptomyces spp.* Gene, 166 (1), 133-7.
- van Dissel, D., van Wezel, G. (2018) *Morphology-driven downscaling of Streptomyces lividans to micro-cultivation*. Antonie van Leeuwenhoek, 111, 457-469.
- World Health Organization. (2014). *World Health Organization Global Tuberculosis Report, 2014*.
- Yanisch-Perron, C., Vieira, J., Messing, J. (1985). *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors*. Gene, 33 (1),103-19.
- Young, D.B., Garbe, T.R. (1991). *Lipoprotein antigens of Mycobacterium tuberculosis*. Res Microbiol, 142 (1): 55-65.