

Caracterización *in silico* de la proteína Rv3494c y evaluación *in vitro* del efecto inhibitorio de sus péptidos sobre la invasión de *Mycobacterium tuberculosis* H37Rv a células humanas U937.

MIKE YEFERSON VASQUEZ RODRIGUEZ

TRABAJO DE GRADO

Presentado como requisito para optar al título de

BIÓLOGO

PONTIFICIA UNIVERSIDAD JAVERIANA

FACULTAD DE CIENCIAS

BIOLOGÍA

Bogotá D. C.

Junio de 2012

Caracterización *in silico* de la proteína Rv3494c y evaluación *in vitro* del efecto inhibitorio de sus péptidos sobre la invasión de *Mycobacterium tuberculosis* H37Rv a células humanas U937.

MIKE YEFERSON VASQUEZ RODRIGUEZ

INGRID SCHULER Ph. D.

Decana Académica

Facultad de Ciencias

ANDREA FORERO

Directora Carrera de Biología

Facultad de Ciencias

Caracterización *in silico* de la proteína Rv3494c y evaluación *in vitro* del efecto inhibitorio de sus péptidos sobre la invasión de *Mycobacterium tuberculosis* H37Rv a células humanas U937.

MIKE YEFERSON VASQUEZ RODRIGUEZ

MARISOL OCAMPO CIFUENTES

Directora

RICARDO VERA BRAVO

Jurado

NOTA DE ADVERTENCIA

Artículo 23 de la resolución No 13 de Julio de 1946

“La universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Sólo velara porque no se publique nada contrario al dogma y a la moral católica, y por que las tesis no contengan ataques personales contra persona alguna, antes bien se vea en ellas el anhelo de buscar la verdad y la justicia”

Agradecimientos

Agradezco a Dios y a la vida por mostrarme el mejor camino para lograr todos mis objetivos. Por darme la oportunidad de desarrollarme como profesional y ser humano día a día.

Gracias a mi mamá, a mi hermana, a mi sobrino y a mi papá por acompañarme durante estos cinco años de proceso de formación profesional y personal. A mi mamá y a mi hermana por darme sabios consejos para tomar decisiones importantes en mi vida. A mi sobrino que se ha convertido en un hijo para mí. A mi papá por sus enseñanzas.

A la Pontificia Universidad Javeriana, por ser el sitio donde me forme como profesional y ser humano. A todos los profesores, directivos y estudiantes que durante cinco años me enseñaron cosas valiosas para formarme como persona y profesional.

A Marisol Ocampo quien me permitió realizar este proyecto en la Fundación Instituto de Inmunología de Colombia y me enseñó cosas valiosas para desarrollarlo de una manera muy disciplinada y profesional. A mis compañeros de laboratorio que me aportaron sus conocimientos para realizar experimentos. Al grupo funcional de Estructura tridimensional de la Fundación Instituto de Inmunología de Colombia por sus consejos e inmensa colaboración en el desarrollo del proyecto. Al profesor Ricardo Vera quien me colaboro durante toda la elaboración de mi trabajo de grado.

TABLA DE CONTENIDO

Lista de figuras.....	VIII
Lista de tablas.....	IX
Lista de abreviaturas.....	X
1. Resumen.....	1
2. Introducción.....	2
3. Formulación del problema.....	3
4. Estado del arte.....	5
4.1 Generalidades de la tuberculosis.....	5
4.1.1. Definición y características.....	5
4.1.2. Epidemiología.....	5
4.2 Micobacteria.....	6
4.2.1 <i>Mycobacterium tuberculosis</i>	6
4.2.2 Complejo <i>Mycobacterium tuberculosis</i> (MTC).....	7
4.2.3 Estructura de la envoltura celular.....	7
4.2.4 Receptores para la infección por <i>Mycobacterium tuberculosis</i>	8
4.2.5 Mecanismos de secreción de proteínas.....	10
4.2.6 Complejo Mayor de Histocompatibilidad (MHC) y epítopes.....	11
4.3 Herramientas bioinformáticas usadas en el análisis de proteínas secretadas.....	11
4.3.1 Análisis <i>in silico</i> de la Localización subcelular.....	12
4.3.2 Análisis <i>in silico</i> de secuencias señal y tipos de secreción.....	13
4.3.3 Análisis <i>in silico</i> de hélices transmembranales.....	13
4.3.4 Inmunoinformática.....	13
4.3.5 Modelamiento de proteínas y Docking Molecular.....	14
5. Objetivos.....	15
5.1 Objetivo general.....	15
5.2 Objetivos específicos.....	15
6. Metodología.....	16
6.1 Caracterización <i>in silico</i> de la proteína.....	16
6.1.1 Localización subcelular de la proteína.....	16
6.1.2 Predicción de péptido señal y tipo de secreción.....	16
6.1.3 Predicción de hélices transmembranales.....	17
6.1.4 Predicción de Epítopes B.....	17

6.1.5 Predicción de Epítopes T.....	17
6.1.6 Predicción de estructura tridimensional de la proteína.....	18
6.1.7 Docking molecular y Refinamiento de los epítopes de la proteína.....	18
6.2 Evaluación <i>in vitro</i> del efecto inhibitorio de los péptidos de la proteína.....	19
6.2.1 Macrófagos derivados de monocitos U937.....	19
6.2.2 Normas de bioseguridad para el manejo de <i>M. tuberculosis</i>	20
6.2.3 Cultivo de <i>M. tuberculosis</i> H37Rv.....	20
6.2.4 Cosecha y marcaje de <i>M. tuberculosis</i> H37Rv.....	20
6.2.5 Síntesis de péptidos de la proteína Rv3494c.....	21
6.2.6 Ensayos de Invasión.....	21
6.2.7 Ensayo de citotoxicidad.....	23
6.3 Análisis estadísticos.....	23
7. Resultados.....	24
7.1 Caracterización bioinformática de la proteína.....	24
7.1.1 Localización subcelular de la proteína.....	24
7.1.2 Predicción de péptido señal y tipo de secreción.....	24
7.1.3 Predicción de hélices transmembranales.....	26
7.1.4 Predicción de Epítopes B.....	27
7.1.5 Predicción de Epítopes T.....	28
7.1.6 Predicción de estructura 3D de la proteína.....	31
7.1.7 Docking molecular y refinamiento de epítopes.....	37
7.2 Inhibición de la invasión a macrófagos derivados de monocitos.....	40
7.3 Citotoxicidad de los péptidos en células U937.....	43
8. Discusión y conclusiones.....	43
9. Recomendaciones.....	50
10. Bibliografía.....	51
11. Anexos.....	59
11.1 Matriz de todos los epítopes T predichos.....	59
11.2 Docking molecular de todos los epítopes predichos y el CMHII.....	65
11.3 Fluorescencias emitidas por las células con y sin péptido.....	68

Lista de figuras

Figura 1. Esquema de la envoltura de la micobacteria.....	8
Figura 2. Predicciones de péptido señal.....	25
Figura 3. Predicción de hélices y regiones transmembranales.....	27
Figura 4. Regiones de la proteína con mayor predicción de epítopes T.....	31
Figura 5. Predicción de Estructura secundaria de la proteína.....	32
Figura 6. Coeficientes QMEAN, Anolea y Gromos.....	34
Figura 7. Coeficientes Z (QMEAN6).....	35
Figura 8. Plot de Ramachandran.....	36
Figura 9. Modelo final de la proteína.....	36
Figura 10. Docking molecular entre epítopes T y el CMHII.....	37
Figura 11. Solapamiento de los epítopes refinados.....	38
Figura 12. Docking molecular entre epítopes T refinados y el CMHII.....	39
Figura 13. Puentes de hidrogeno entre epítope y CMHII.....	39
Figura 14. Inhibición de la invasión.....	40
Figura 15. Porcentajes de inhibición de la invasión.....	43
Figura 16. Visualización de péptidos de mayor inhibición.....	42
Figura 17. Visualización de epítopes T promiscuos.....	45
Figura 18. Visualización de epítopes B.....	46

Lista de tablas

Tabla 1. Epítopes B predichos.....	28
Tabla 2. Numero total de epítopes predichos.....	29
Tabla 3. Epítopes predichos con más frecuencia.....	30
Tabla 4. Péptidos utilizados en lo ensayos de invasión.....	42

Lista de abreviaturas

MCE: Mammalian Cell entry

MCE4F: Mammalian Cell entry 4F

HLA: Human Leucocyte Antigen

CMHI: Complejo Mayor de Histocompatibilidad I

CMHII: Complejo Mayor de Histocompatibilidad II

HABP's: High Activity Binding Peptides

CDR's: Complementary-Determining Regions

1. Resumen

En este trabajo se hizo una caracterización *in silico* de la proteína Rv3494c (MCE4F) y una evaluación *in vitro* de la capacidad que tienen sus péptidos de inhibir la invasión de *M. tuberculosis* H37Rv en macrófagos derivados de monocitos humanos U937. El componente bioinformático fue generado a partir de un conjunto de las herramientas de más uso en la literatura. Debido a la importancia que tienen las características topológicas en la interacción de proteínas de *M. tuberculosis* con la célula blanco de infección, se realizó la predicción de localización subcelular, presencia de secuencia señal y de hélices transmembranales, encontrándose una posible localización en la envoltura celular y una posible secuencia señal ligada a una secreción de vía clásica. De igual forma, se hizo una predicción de epítopes T para diferentes HLA (Human leukocyte antigen, en inglés) y de epítopes B ya que se espera que empleando la aproximación planteada por la Fundación Instituto de Inmunología se encuentren secuencias peptídicas de la proteína Rv3494c que puedan incluirse en el diseño de métodos diagnósticos y profilácticos contra tuberculosis. Se determinaron quince epítopes T promiscuos y cuatro epítopes B fueron determinados por análisis *in silico*. También fue predicho un modelo de estructura tridimensional para la proteína, observando validaciones positivas y negativas a través de diferentes métodos. Un docking o acoplamiento molecular se realizó entre el moléculas del complejo mayor de histocompatibilidad clase II (CMH II) HLADRB1 y los epítopes T, mostrando que 14 de los quince péptidos se ubicaron en la zona comprendida por los bolsillos o “pockets” de interacción con el HLA, pero solo uno de ellos se acopló de manera adecuada a los bolsillos relevantes (1, 4, 6 y 9) de la molécula esta molécula. Así, una minimización de energía se realizó en dos de los péptidos mas relevantes (el mas promiscuo y el de mejor acople al HLA) y un nuevo docking molecular fue generado, evidenciando un acople adecuado al CMHII por parte de un solo epítope. A pesar de no tener estructuras secundarias similares ni un patrón de topología, los epítopes B y T, y los péptidos que mostraron inhibiciones significativas se ubicaron en la periferia de la estructura tridimensional de la proteína. Por otro lado, el componente *in vitro* mostró que seis de los veintiséis péptidos evaluados mostraron inhibiciones significativas mayores al 30% en todas las concentraciones a las que fueron probados. De igual forma ninguno de los péptidos mostró efectos tóxicos sobre las células U937. Además se determinaron zonas comunes de la proteína con capacidad de ser epítopes T, epítopes B y de inhibir la invasión de la micobacteria.

2. Introducción

La tuberculosis es una enfermedad infecciosa que afecta principalmente los pulmones y es causada por el bacilo ácido alcohol resistente (BAAR) *Mycobacterium tuberculosis*. Se transmite de una persona a otra a través de expectoraciones que contienen la bacteria y que son generadas en el aparato respiratorio de pacientes con enfermedad pulmonar activa (1). La infección por *M. tuberculosis* suele ser asintomática, dado que el sistema inmune del hospedero sano actúa formando una barrera alrededor de la bacteria. La tuberculosis se puede tratar mediante la administración de un conjunto de antibióticos administrados durante períodos de seis meses (1), sin embargo la aparición de cepas resistentes y extremadamente resistentes a diferentes antibióticos, la co-infección de pacientes con SIDA y la variable eficacia de la vacuna BCG (Bacilo *Mycobacterium bovis Calmette-Guerin*) (0-80%) han llevado a la búsqueda de nuevas terapias y medidas preventivas, tales como el desarrollo de una nueva vacuna que pueda controlar efectivamente la enfermedad (2).

Se ha postulado que las proteínas que se encuentran en la envoltura de *M. tuberculosis* H37Rv, basados en la descripción del genoma (3), tienen funciones importantes en la patogénesis y han sido objeto de estudio por parte del grupo Funcional Tuberculosis de la Fundación Instituto de Inmunología de Colombia, FIDIC. Adicionalmente, estudios recientes han reportado proteínas externas de membrana (Outer Membrane Proteins) en *M. tuberculosis* implicadas en procesos de secreción, transporte de nutrientes e interacciones entre hospedero - patógeno (4).

La metodología propuesta en los estudios de FIDIC se basa en la identificación de péptidos de alta actividad de unión (High Activity Binding Peptides, en inglés) a la célula hospedera, dentro de la secuencia de la proteína de interés. Estos HABPs serán usados para diseñar péptidos inmunogénicos (5 - 15) que, de acuerdo con los resultados obtenidos en el desarrollo de la vacuna contra malaria se basan en la modificación de los residuos críticos de unión a las células blanco de cada uno de los HABPs. Estas secuencias harán parte de una vacuna sintética, multiepitópica definitiva contra la tuberculosis.

La proteína Rv3494c (Mce4F) de *M. tuberculosis*, perteneciente a la familia de proteínas MCE (mammalian cell entry), ha sido anotada e identificada en el genoma de este microorganismo como una posible proteína de la membrana externa que es secretada al

exterior de la bacteria y que cumple funciones importantes dentro de la patogénesis de la micobacteria (3). Sin embargo, no se ha hecho la caracterización que mediante el empleo de herramientas bioinformáticas pueda ser usada en la predicción de propiedades topológicas y antigénicas así como tampoco la determinación funcional *in vitro* de los péptidos de esta proteína como posibles inhibidores de la invasión del patógeno en células humanas.

A partir de lo anterior, se propone realizar una caracterización bioinformática de la proteína Rv3494c (MCE4F) con la cual se determinarán características topológicas y antigénicas de la proteína y se definirá si presenta alguna importancia en la búsqueda de posibles candidatos a una vacuna contra tuberculosis. Por otro lado, es importante realizar una parte funcional con el fin de evaluar el efecto de los péptidos de la proteína en la inhibición de la invasión del patógeno en macrófagos pulmonares U937 y determinar cual de ellos es relevante para estudios futuros.

3. Formulación del problema

La tuberculosis es una enfermedad infecciosa causada por especies del género *Mycobacterium* pertenecientes al Complejo de *Mycobacterium tuberculosis* (MTC) (16). Para el año 2009, la población afectada por esta patología aumentó considerablemente comparada con años anteriores. Año tras año, la detección de nuevos casos sigue en aumento y hay una baja reducción en la tasa de incidencia *per capita* en la población (1; 17). Teniendo en cuenta la alta incidencia de la tuberculosis a nivel global y nacional y partiendo del hecho que en 1998 se publicó el genoma de *Mycobacterium tuberculosis* (3) muchas expectativas se han generado para desarrollar diversos procesos que permitan el desarrollo de vacunas en contra de la tuberculosis. De igual forma, la aparición de cepas multi-drogas resistentes, la infección de pacientes con SIDA y la variable eficacia de la vacuna BCG han llevado a la búsqueda de nuevas terapias y medidas preventivas, tales como el desarrollo de una nueva vacuna que pueda controlar efectivamente la enfermedad (18; 19).

Conocer la morfología de la micobacteria permite tener ventajas para el propósito mencionado anteriormente. Así pues, la pared celular de la micobacteria es una compleja estructura compuesta por una gran variedad y cantidad de lípidos. En la capa interna de la pared celular de la bacteria se encuentran largas cadenas de ácidos grasos (ácidos micólicos) que están covalentemente unidos al polímero de arabinogalactano y peptidoglicano, mientras

que la capa externa está compuesta por otros tipos de lípidos (4;5). En la envoltura posiblemente están expuestas las proteínas que cumplen roles importantes en la patogénesis y en el contacto de la bacteria con la célula hospedera, como se ha encontrado en diferentes estudios de proteómica (4; 5; 20; 21).

Dentro de las proteínas de *Mycobacterium tuberculosis* H37Rv para las cuales experimentalmente se ha determinado su presencia en la envoltura de la micobacteriana y que han sido estudiadas para determinar la presencia de estos HABPs se encuentran: Rv3166c (5), Rv1980 (6), Rv0679c (7), Rv0180c (8), Rv1490 (9), Rv2707 (10), Rv2560 (11), Rv2969c (12), Rv1510 (13), Rv2004c (14) y Rv2536 (15). Se han identificado para cada una de ellas los péptidos de alta actividad de unión (HABPs) a la célula hospedera, los cuales serán usados para diseñar péptidos antigénicos por modificación de los residuos de unión a las células blanco. A partir de ello podría ser posible el diseño de una vacuna sintética, subunitaria definitiva contra la tuberculosis.

La familia inmunogénica de proteínas MCE (mammalian cell entry) se ha considerado dentro de los factores de virulencia que están implicados en la entrada y supervivencia del patógeno en los macrófagos pulmonares (22). La secuenciación del genoma de *M. tuberculosis* reveló que la proteína MCE1A (Rv0169) era parte de un operón que codifica ocho proteínas de membrana: YrbEA-B y MceA-F; este operón tiene cuatro repeticiones (mce1-4) en el genoma *M. tuberculosis* (23). La proteína Rv3494c (MCE4F) de *M. tuberculosis*, perteneciente a la familia de proteínas MCE ha sido considerada como una proteína de la membrana externa que es secretada al exterior de la bacteria y que cumple funciones importantes dentro de la patogénesis (4). Sin embargo, su caracterización mediante el empleo de herramientas bioinformáticas y evaluación funcional no ha sido determinada. Esta caracterización puede ser iniciada con el uso de herramientas bioinformáticas que permitan la identificación y caracterización de posibles fragmentos peptídicos que puedan ser incluidos en el diseño de subunidades de una vacuna contra la tuberculosis y el estudio de sus péptidos como posibles inhibidores de la entrada de la micobacteria al espacio celular (24; 25).

4. Estado del arte

4.1 Generalidades de la tuberculosis

4.1.1. Definición y características

La tuberculosis es una enfermedad infecciosa que afecta principalmente los pulmones y es causada por una bacteria (*Mycobacterium tuberculosis*). Se transmite de una persona a otra a través de partículas generadas en el aparato respiratorio de pacientes con enfermedad pulmonar activa. Esta enfermedad infecto-contagiosa es de las más letales y antiguas que afecta al ser humano, posee una amplia distribución en el mundo, produciendo cada año la muerte de alrededor de 2 millones de personas (1).

La infección por *M. tuberculosis* suele ser asintomática en personas sanas, dado que su sistema inmunitario actúa formando una barrera alrededor de la bacteria. Los síntomas de la tuberculosis pulmonar activa son tos, a veces con esputo que puede ser sanguinolento, dolor torácico, debilidad, pérdida de peso, fiebre y sudoración nocturna. La tuberculosis se puede tratar mediante la administración de antibióticos durante seis meses (26).

La patogénesis de la tuberculosis puede ser considerada como una constante batalla entre la célula blanco y el bacilo infectante. A partir de esto se pueden definir algunas estrategias y debilidades que tienen las células blanco y los bacilos para enfrentar esta relación de infección (27). Los bacilos tienen la habilidad de multiplicarse logarítmicamente dentro de los macrófagos inactivos, como en los monocitos que son transportados a través del torrente sanguíneo al sitio de infección. También la habilidad de multiplicarse extracelularmente le permite al bacilo colonizar cavidades externas a los pulmones.

4.1.2. Epidemiología

Según el más reciente informe entregado por la Organización Mundial de la salud (OMS), en 2009 se diagnosticaron 9.4 millones de casos de tuberculosis en todo el mundo (137 casos por cada 100 000 habitantes) (1). Año tras año, la detección de nuevos casos sigue en aumento y hay una baja reducción en la tasa de incidencia *per capita* en la población (1). De los 9.4 millones de casos en 2009, 1.2 millones de casos (12%) pertenecían a personas con VIH positivo (1). Mostrando estas cifras una gran problemática en salud pública.

En Colombia durante el año 2010 se notificaron un total de 11362 casos de tuberculosis al SIVIGILA (Sistema de vigilancia en salud pública) con una tasa de incidencia general de 24,9 personas por cada 100 000 habitantes, aumentando en relación a los casos notificados para el año 2009 (10 070 con una incidencia general de 22,4 personas por cada 100 000 habitantes), la forma más común para ambos años fue la tuberculosis de tipo pulmonar, con una incidencia de 20,3 por 100 000 habitantes para el año 2010. En cuanto a la coinfección tuberculosis/VIH, se presentaron 1 231 casos de los cuáles el 32% presentó tuberculosis de tipo extrapulmonar, porcentaje muy superior a los casos de tuberculosis extrapulmonar en pacientes sin coinfección que para el año 2010 fue del 17% y el 2009 fue del 18,5%; sin embargo un alto porcentaje de casos de tuberculosis desconocen que tengan coinfección con VIH. En cuanto a la vigilancia de tuberculosis fármacorresistente, a partir del año 2010 se inició el ingreso de casos al SIVIGILA; según datos de la Red Nacional de Laboratorios el número de casos para el 2010 llegó a los 131 (17).

4.2 Micobacteria

4.2.1 *Mycobacterium tuberculosis*

Mycobacterium tuberculosis es un patógeno intracelular facultativo, el cual puede sobrevivir y replicarse dentro de los macrófagos de un hospedero. Esta bacteria es transmitida vía aérea (inhalación), llevando el bacilo hasta el alveolo de los pulmones (28). Las células epiteliales pulmonares son la primera barrera que el organismo tiene para defenderse ante el patógeno; además son las células que más predominan en el pulmón. Así, *M. tuberculosis* se adhiere, invade y replica dentro de este tipo de células (29).

Cuando se inicia la replicación intracelular de la micobacteria dentro del macrófago, se genera una reacción inflamatoria con el reclutamiento de leucocitos mononucleares (monocitos, macrófagos y linfocitos). La acumulación de ese tipo de células es conocido como granuloma. El crecimiento de la micobacteria en esta etapa de la infección es intracelular, lo que permite evidenciar al bacilo dentro de los macrófagos del granuloma. Como la respuesta inmune continúa y el tamaño del granuloma se incrementa, los macrófagos del centro del granuloma empiezan a morir, probablemente por la combinación de factores de virulencia de la micobacteria y falta de oxígeno y nutrientes, formando una zona central de necrosis. La

micobacteria puede sobrevivir y replicarse extracelularmente, dentro de los centros necróticos del granuloma. Sin embargo, esa zona se empieza a solidificar haciendo que la micobacteria entre en un estado de latencia donde disminuye su metabolismo (28).

4.2.2 Complejo *Mycobacterium tuberculosis* (MTC)

Al complejo *M. tuberculosis* pertenecen las especies *M. tuberculosis*, *M. africanum*, *M. canettii*, *M. bovis* y *M. microti* los cuales están muy relacionados con un 99% de identidad a nivel de nucleótidos para algunos loci (16). Se han realizado muchos estudios de genómica, proteómica y transcriptómica para determinar aspectos específicos de las relaciones de estas especies del MTC. Sin embargo, difieren significativamente en morfología, bioquímica, rangos de células blanco y patrones de infección en modelos animales. *M. tuberculosis* mata a tres millones de personas anualmente; en muchas partes de África, *M. africanum* causa más casos de tuberculosis que *M. tuberculosis*. Por otro lado, la infección por *M. canettii* es muy rara (16).

A pesar que el humano es el único hospedero del patógeno, existen modelos animales que muestran los patrones de comportamiento de las especies. *M. bovis* produce tuberculosis en un gran espectro en mamíferos, incluyendo humanos y ganado; esta especie también fue la que más casos de tuberculosis causó antes de que se pudiera pasteurizar la leche. A partir de las especies *M. microti* y *M. bovis* se han elaborado vacunas que han permitido manejar la tuberculosis hasta el presente (16).

4.2.3 Estructura de la envoltura celular

La micobacteria posee una envoltura celular muy compleja que le permite tener una eficiente barrera de permeabilidad que juega un papel importante en la resistencia a drogas y en la supervivencia bajo una gran variedad de condiciones (30). Básicamente, la envoltura celular de *M. tuberculosis* esta constituida por tres elementos. La cápsula, que es una envoltura externa que protege a la bacteria del medio externo y tiene una interacción directa con los elementos de la respuesta inmune del hospedero; esta compuesta por ácidos micólicos (que dan propiedades hidrofóbicas a la envoltura celular) y glicolípidos, que dan identidad antigénica a la bacteria (31). La pared celular, que esta separada de la cápsula por un espacio periplásmico y posee un alto contenido de lípidos (50-60%) que le dan un carácter

hidrofóbico, esta compuesta por un complejo macromolecular de largas cadenas de ácidos grasos (ácidos micólicos) covalentemente unidos al polímero de arabinogalactano y peptidoglicano. Recientemente se han descrito proteínas y lipoproteínas que hacen parte de esta estructura celular (32). Por último, la membrana celular es una bicapa lipídica donde los fosfolípidos están altamente glicosilados y forman moléculas como la lipoarabinomana (LAM) que está implicada en la patogénesis (33); también han sido encontradas proteínas inmunogénicamente importantes dentro de esta región (20;21) (Figura 1).

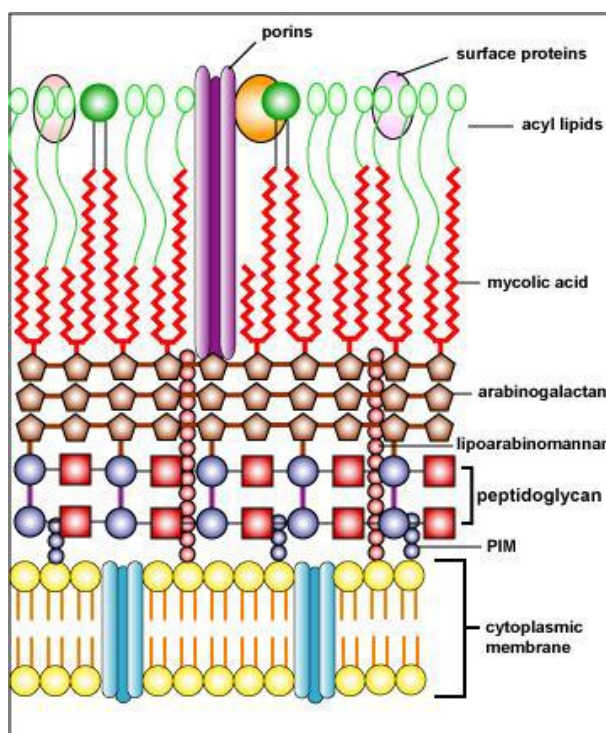


Figura 1. Esquema de la envoltura celular de la micobacteria. Tomado de (85).

4.2.4 Receptores para la infección por *Mycobacterium tuberculosis*

Los microorganismos patogénicos han desarrollado muchas estrategias para superar la respuesta antimicrobiana de las células hospedadas. Así, por ejemplo, la fagocitosis de los macrófagos se puede dar a través de receptores de los macrófagos, que crean un “bypass” o mecanismo por el que el patógeno puede manipular el entorno de la célula hospedada para su propio beneficio (31).

Cuando los macrófagos humanos reconocen e internalizan a *M. tuberculosis*, la bacteria se ubica en un compartimento fagosomal que resiste su fusión con los lisosomas. Cuando la bacteria es opsonizada, el compartimento fagosomal adquiere de nuevo sus propiedades de

fusión a lisosomas (31). Así, el proceso de internalización (no opsonica) se da a través de diferentes tipos de receptores que reconocen ligandos que son expresados en la superficie del bacilo (34).

Muchos de los receptores de los macrófagos tienen el potencial de reconocer, unirse y fagocitar a *M. tuberculosis* para iniciar respuesta inmune: dentro de estos, se encuentran el receptor de manosa (MR), los receptores del complemento (CRs), el receptor CD14, los receptores de surfactantes de Proteína A y los receptores “Scavenger” o basurero (35-38).

El receptor de manosa es una proteína transmembranal monomérica que tiene ocho dominios (tipo lectina dependientes de calcio) de reconocimiento de carbohidratos. Este receptor, que solo es expresado en macrófagos maduros y no en monocitos, solo se une a la cepa virulenta de *M. tuberculosis* H37Rv y no a la cepa no virulenta H37Ra, haciendo fagocitosis no opsonica de ésta. Así, se da una interacción con los residuos de manosa que se encuentran en el extremo del lipoarabinomano (LAM) de la envoltura celular de la bacteria (36; 37). Por otra parte, los receptores del complemento tienen dos formas estructurales en los macrófagos: El CR1 es una proteína transmembranal monovalente y los CR3 y CR4 son proteínas heterodímeras, pertenecientes a la superfamilia de las integrinas, que poseen subunidades alfa idénticas (CD18) y subunidades beta diferentes (CD11b y CD11c). *M. tuberculosis* usa los polisacáridos de su pared celular para unirse a los sitios de unión de los receptores y adherirse a los productos de degradación C3b y C3bi del complemento C3, lo que permite su opsonización por parte del receptor CR1 y CR3/CR4, respectivamente (34; 35). La proteína surfactante A (la más predominante del surfactante pulmonar), Sp-A, es un miembro de la familia de proteínas colectinas, que incluyen la proteína de unión a manosa (MBP) y la C1q (componente del complemento). La glicosilación del receptor hace que este se una a la micobacteria, en un proceso dependiente de calcio, que posiblemente altere procesos de unión a otros tipos de receptores (36).

Existen otros tipos de receptores como el CD14, que es una proteína de membrana unida por glicofosfatidilinositol que puede interactuar con lipopolisacáridos de bacterias Gram negativas y con lipoarabinomano (LAM) de *M. tuberculosis*, permitiendo la secreción de interleucina-8 por parte de los macrófagos. Los receptores “Scavenger” (SR) o basurero que se unen a los lípidos de baja densidad, a fosfatidilserina y a macromoléculas polianiónicas como sulfolípidos de trehalosa (ácidos micólicos unidos al disacárido de trehalosa que se

encuentran en la periferia de la pared celular), son una familia de receptores que presentan diversas estructuras, que en el caso de las micobacterias también están involucrados en la invasión de la célula hospedera, al unirse a los SR tipo A. (33; 36).

Igualmente están los receptores denominados de tipo Toll o Toll-Like receptors (TLR), los cuales son importantes, tanto en la respuesta inmune innata como en la específica, y son responsables de las primeras señales producidas ante estímulos antigénicos. Estos receptores reconocen numerosos ligandos provenientes de patógenos, tales como los lipopolisacáridos, peptidoglucano, LAM, ácido lipoteicoico, ácidos nucleicos y sulfolípidos. La activación de los macrófagos por el receptor tipo Toll-2 (TLR2) activa el factor nuclear- κ B (NF- κ B), responsable de la producción de citoquinas para iniciar una respuesta inmune eficiente contra la micobacteria y la activación del macrófago vía TLR6 regula la activación de células estimuladas por moléculas como la lipoproteína conocida como factor soluble de la tuberculosis (39; 40).

4.2.5. Mecanismos de secreción de proteínas

La secreción de proteínas micobacterianas a la envoltura y fuera de ésta es regulada por varios mecanismos o rutas de secreción. En las bacterias uno de los más conocidos sistemas de secreción es el Sec-dependiente o clásico, el cual transporta precursores de proteínas a la membrana. La proteína SecA-ATPasa reconoce una secuencia señal que contiene un dominio cargado positivamente en la región N terminal, seguido de un dominio hidrofóbico y de una región polar. Hay otro mecanismo de secreción alternativo al clásico y es el de translocación arginina-doble (Tat) que básicamente posee una secuencia señal como la del sistema de secreción clásica con un motivo de doble arginina en la región N terminal. La secreción de proteínas se da con una secuencia señal caracterizada por la presencia de una alanina o una glicina en la posición -1 y una leucina en la posición -3 de la región C terminal en el sitio de clivaje y una cisteína en la posición 1 después del sitio de clivaje (41). Además existe un tipo de secreción no clásica en el que no existe ningún tipo de secuencia señal, sino que simplemente se secreta al exterior la proteína madura (42). Muchas moléculas localizadas y exportadas en la superficie juegan un papel importante en la virulencia dentro de las cuales se destacan las familias de proteínas Erp, HbhA, PE-PGRS y PPE en *Mycobacterium*. Estas dos últimas conforman el 10% del genoma de *M. tuberculosis* H37Rv. La familia PE se encuentra en todos los genomas del género de *Mycobacterium*, aunque su número es ampliamente

variable. *M. tuberculosis*, *M. bovis* BCG y *M. marinum* tienen de 80-140 genes PE. En *M. leprae* la familia PE se convierte en pseudogenes. Esta familia es la encargada de la expresión de los granulomas en etapas iniciales de la infección y puede variar según la especie (16).

4.2.6 Complejo Mayor de Histocompatibilidad (MHC) y Epítopes

De los muchos genes relacionados con la respuesta inmune, los genes del CMH, HLA en humanos, están dentro de los más cruciales y más variables. Las proteínas CMH presentan antígenos (péptidos) extraños de patógenos intracelulares (CMH I) y extracelulares (CMH II) a las células CD8+ citotóxicas (CMH I) y a las células ayudadoras CD4+ (CMH II) para iniciar la respuesta inmune. Al haber cientos de alelos de HLA, la variación en estos puede variar y significativamente cambiar la respuesta en los individuos (43).

Para patógenos extracelulares como *P. falcifarum*, los péptidos que interactúan con el CMH II se ubican en un loop formado por aminoácidos de la cadena alfa y beta altamente conservados en humano y ratones, generando que el péptido se acople a 9 bolsillos o “pockets” que determinan la presentación de este antígeno a los receptores de linfocitos T (TCR). De estos nueve bolsillos, se ha visto que cuatro de ellos (bolsillos 1, 4, 6 y 9) son relevantes en el acople de péptidos al CMH II y su posterior presentación a TCR (44 - 46).

Las moléculas que interactúan directamente con el CMH para son secuencias cortas denominadas epítopes, siendo entonces definidos como péptidos de las proteínas bacterianas que interactúan de manera específica con las células hospederas, generando la respuesta celular o humoral. Los epítopes T, que tienen una longitud de 9 a 15 aminoácidos, generan una respuesta celular al activar los linfocitos T. Esta respuesta es dada por el complejo formado entre el epítope, el CMH y el TCR (47). Los epítopes B, que generalmente tiene una longitud de 20 aminoácidos, son los encargados de generar la respuesta humoral inmune, es decir la producción de anticuerpos. La identificación de estos epítopes es importante en el desarrollo de vacunas y pruebas de inmunodiagnósticas (48).

4.3. Herramientas bioinformáticas usadas en el análisis de proteínas secretadas

Una gran cantidad de proteínas de superficie de *M. tuberculosis* han sido identificadas gracias a la publicación del genoma de este patógeno y a su posible rol en la patogénesis y virulencia

dentro de las células hospederas, sin embargo ha sido difícil determinar su estructura, función y localización por dificultades relacionadas con los análisis de proteómica y transcriptómica. Con el uso de herramientas bioinformáticas es posible determinar la localización subcelular de la proteína, la presencia de una secuencia señal, el tipo de secreción hacia la membrana externa y su estructura (24; 25) para microorganismos como *M. tuberculosis*. Además se pueden detectar proteínas y secuencias dentro de esta que posiblemente tengan importancia en el contacto inicial entre el patógeno y la célula hospedera, aportando soluciones e información en la búsqueda de una vacuna.

4.3.1 Análisis *in silico* de la Localización subcelular

La predicción *in silico* de la localización subcelular de proteínas representa una alternativa a los estudios experimentales. El uso de herramientas bioinformáticas para este fin es llamativo, ya que asigna una posible localización subcelular de forma rápida y económica, eliminando horas de trabajo en actividades experimentales (49). El primer servidor creado para este fin fue Psort (50) y una de sus variantes más populares es Psortb (51). Gpos-Ploc (52; 53) es un servidor perteneciente al paquete de servidores de predicción de localización subcelular en organismos Cell-PLoc (52). Predicciones de la localización subcelular de proteínas de *M. tuberculosis* ya se han realizado utilizando Psortb y Gpos-Ploc, mostrando una gran eficacia de estas herramientas para hacer las predicciones (24; 25). TBPred es el único servidor que predice la localización subcelular de proteínas exclusivas del género *Mycobacterium*; las localizaciones que da el servidor son: citoplasmáticas, de membrana, unidas a la membrana (lipoproteínas) y secretadas (54).

La predicción de la localización subcelular puede ser deducida a partir de características como la presencia de secuencias señal, composición de aminoácidos, análisis comparativos con bases de datos de proteínas, información evolutiva o estructural y una combinación de varias de estas características. De igual forma, los algoritmos de análisis incluyen redes artificiales neurales (ANN), “Support Vector Machines” (SVM), Fuzzy-logic o métodos basados en evidencia y teoría. A partir de ello, los servidores arrojan puntuaciones mostrando la probabilidad de que la proteína se ubique en un sitio específico dentro de la célula (49).

4.3.2 Análisis *in silico* de secuencias señal y tipos de secreción

Los métodos computacionales para la predicción de péptidos señal N-terminal iniciaron desde hace veinte años, usando enfoques basados en matrices de tamaño. A mediados de los años noventa se desarrollaron métodos de predicción basados en algoritmos, incrementando y mejorando la predicción de estas secuencias señal. El servidor SignalP es uno de los más usados actualmente y muestra la presencia de péptidos señal tipo I (sec-depend). La versión 3.0 (55) de este servidor hace predicciones basadas en Modelos ocultos de Markov (HMM) y redes artificiales neurales (ANN) y muestra el sitio de clivaje de la secuencia señal, mientras que la última versión 4.0 permite discernir entre lo que puede ser un péptido señal o una hélice transmembranal (56). Como complemento a la predicción de péptidos señal tipo I, se desarrollaron los servidores TatP 1.0 (57) y LipoP (58), que predicen secuencias señal de tipo II (doble Arginina) usando ANN y secuencias señal de lipoproteínas usando HMM. Por último el servidor Secretome 2.0 (42) hace predicciones (basadas en ANN) de proteínas secretadas sin poseer alguna secuencia señal. De esta manera, al usar el consenso de estas herramientas se obtiene una aproximación a lo que puede ser el péptido señal de la proteína y su tipo de secreción al medio extracelular.

4.3.3 Análisis *in silico* de hélices transmembranales

Los servidores de identificación de regiones transmembranales predicen en ocasiones péptidos señal como hélices transmembranales y viceversa. Por ejemplo, el servidor TMHMM (59), basado en modelos ocultos de Markov, hace la predicción de regiones transmembranales sin tener en cuenta la presencia de péptidos señal; a pesar de esto la eficacia en las predicciones es muy alta. Por otro lado, el servidor Phobius (60), también basado en HMM, da solución a este problema y predice hélices transmembranales y al mismo tiempo discierne si estas son péptidos señal. El uso de estas dos herramientas (que son de gran popularidad), permite dar una aproximación a la presencia o ausencia de regiones transmembranales.

4.3.4 Inmunoinformática

Recientemente, muchos algoritmos y herramientas se han desarrollado y validado con el fin de predecir la afinidad de secuencias de péptidos a diferentes alelos de HLA tipo I y II, con el

fin de determinar epítopes T. De esta manera, un consenso de estas herramientas (conformada por tres o más de estas) permite abordar con más efectividad la predicción de epítopes para el CMH II (43). De los servidores disponibles actualmente en línea para la predicción de epítopes T, los más eficaces y más usados son ARB (61), basado en la matriz de predicción TEPITOPE (62) que tiene en cuenta los “pockets” de unión del CMH para hacer las predicciones, MHCPRED 2.0 (63), NETMHCII 2.2 (64), PROPPRED (65) y NetMHCIIpan 2.1 (66). El factor común de estos servidores en línea es que arrojan una puntuación a partir del valor IC50 que sea asignado a cada péptido, mostrando la afinidad que estos tengan para unirse a cada uno de los HLA analizados; así pues, los epítopes con menor valor de IC50 serán los que tengan una mayor probabilidad de unión al CMH. Esto permite hacer una comparación homogénea entre todos los servidores. A pesar que PROPPRED no arroja valores de IC50, la matriz (TEPITOPE) con la que está sustentado este servidor es de las más verídicas y confiables (62).

Por otro lado, la identificación y caracterización de los epítopes B sirve de manera fundamental en el diseño de vacunas basadas en epítopes, en el desarrollo de pruebas inmuno diagnósticas y en la producción de anticuerpos (48). Muchas técnicas experimentales se han desarrollado para determinar epítopes B. Sin embargo, el alto costo y los esfuerzos que conllevan realizarlas impiden el avance en la detección de estos péptidos. A partir de ello, las técnicas computacionales ofrecen una predicción rápida, poco costosa y confiable de epítopes B (67). La identificación de estos epítopes a través de métodos de escala de propensión (del inglés: *Propensity Scale Methods; PSM*), asigna valores a cada aminoácido de la secuencia teniendo en cuenta propiedades físico químicas como antigenicidad, hidrofiliidad y accesibilidad al solvente (48), lo que permite una confiable identificación de epítopes B. Así, el programa Antheptot 3D (68) y el servidor en línea Bepipred 1.0 (67), que han sido usados en investigaciones de validaciones de predicción de epítopes B (25), representan dos aproximaciones de predicción de epítopes B que utilizan el método PSM.

4.3.5 Modelamiento de proteínas y Docking Molecular

El conocimiento de la estructura tridimensional de las proteínas permite conocer sus características evolutivas y funcionales. Además la información estructural de esta puede ser usada en iniciativas de diseños de medicamentos. Sin embargo los procedimientos experimentales para determinar estas estructuras son muy costosos y conllevan grandes

esfuerzos. Es así como las predicciones computacionales de estructuras proteicas juegan un papel importante en el desarrollo de la genómica y la proteómica (69). Dentro de los métodos existentes para modelar proteínas se encuentran: Modelación por homología de secuencias, en donde se destaca el servidor SWISS MODEL (70) y EsyPred (71), modelación por reconocimiento de moldes, en donde se destacan el servidor PHYRE (72) y I-TASSER (73) y modelación *ab initio*, en donde se destaca el servidor Robbeta (74). En todos los servidores la estructura tridimensional es predicha a partir de una estructura secundaria determinada por el mismo servidor o por servidores especializados en este campo.

El Docking molecular entre la proteína modelada o los péptidos de esta y moléculas de interés se puede realizar a través de servidores que simulan interacciones de tipo enzima-inhibidor, antígeno-anticuerpo y proteína-pequeño ligando. Uno de los servidores en línea más usados y con más fiabilidad es PATCHDOCK (75; 76), el cual hace una simulación de interacción en tres pasos: 1) Computa la superficie molecular de la molécula dada; 2) detecta “parches” geométricos usando algoritmos; 3) filtra cada parche geoméricamente y le da una puntuación con la cual se obtiene la mejor (es) interacción (es) (75; 76). A partir de ello, por ejemplo, se pueden determinar las interacciones entre epítopes T y moléculas del CMH I y II, que a su vez permitirían hacer otra interacción con los receptores de los linfocitos T (TCR).

5. Objetivos

5.1 Objetivo general

Determinar *in silico* las características de la proteína Rv3494c y la capacidad que presentan sus péptidos de inhibir la invasión de *Mycobacterium tuberculosis* H37Rv en células U937.

5.2 Objetivos específicos

- Determinar la posible localización subcelular y el potencial antigénico de la proteína Rv3494c por medio de herramientas bioinformáticas.
- Evaluar la capacidad inhibitoria de péptidos de la proteína Rv3494c sobre la invasión de *M. tuberculosis* en células humanas U937.
- Establecer relaciones entre las características estructurales de la proteína y la capacidad de inhibición de la invasión de algunas regiones de ésta.

6. Metodología

La metodología se siguió de acuerdo a criterios planteados anteriormente por el grupo de tuberculosis de la FIDIC (24; 25) como se menciona a continuación:

6.1 Caracterización *in silico* de la proteína

La secuencia FASTA de la proteína Rv3494c se descargó de la base de datos del NCBI (<http://www.ncbi.nlm.nih.gov/protein/2924431?report=fasta>), con el fin de realizar todos los procedimientos bioinformáticos de análisis.

6.1.1 Localización subcelular de la proteína

La localización de la proteína dentro de un compartimento subcelular específico se realizó con las herramientas bioinformáticas de “aprendizaje automatizado” (learning machines) Gpos-Ploc (52; 53), Psortb v 2.0.4 (51) y TBpred (54) basado en el protocolo bioinformático de predicción de localización subcelular de proteínas y péptidos señal. Solo se tomó en cuenta si la localización de la proteína es extracelular, de membrana citoplasmática o de pared celular. Todas las herramientas fueron usadas con las opciones predeterminadas por los mismos servidores.

6.1.2 Predicción de péptido señal y tipo de secreción

La predicción de un péptido señal y el tipo de péptido señal que la proteína pueda tener (Vías clásicas como Sec-depent, Tat o Lipoproteína, y vías no clásicas) se realizó con las herramientas bioinformáticas de “aprendizaje automatizado” (learning machines) SignalP V3.0 (55), SignalP V4.0 (56), TatP 1.0 (57) y LipoP 1.0 (58) y Secretome 2.0 (42) basado en el protocolo bioinformático de predicción de localización subcelular de proteínas y péptidos señal. El umbral para cada una de estas herramientas fue dado acorde a valores de cada una de las herramientas, teniendo en cuenta que la proteína solo debe ser secretada por uno de estos mecanismos. Para SignalP, TatP y Secretome es ≥ 0.5 ; para LipoP la proteína tiene que estar entre los primeros 70 aminoácidos y tener como mejor resultado (best type) una señal SPII y una mayor probabilidad de sitios de clivaje en SPII.

6.1.3 Predicción de hélices transmembranales

La predicción de hélices transmembranales dentro de la proteína se realizó con las herramientas bioinformáticas de “aprendizaje automatizado” (learning machines) TMHMM 2.0 (59) y Phobius (60) basado en el protocolo bioinformático de predicción de localización subcelular y péptidos señal de proteínas. Se determinó el número de hélices que contiene la proteína y su significancia se tuvo en cuenta si el número esperado de aminoácidos para esta (s) hélices (s) era mayor a 18, mencionado en el servidor TMHMM 2.0 como expAA (expectedAminoAcids).

6.1.4 Predicción de Epítopes B

La predicción de posibles péptidos de la proteína que puedan ser epítopes B se realizó utilizando herramientas bioinformáticas de “aprendizaje automatizado” (learning machines) Antheprot 3D (68) y Bepipred 1.0 (67) basado en el protocolo bioinformático de predicción de localización subcelular y péptidos señal de proteínas. Para seleccionar los péptidos con más probabilidad de ser epítopes B se hizo un consenso entre ambas herramientas bioinformáticas. Teniendo los posibles epítopes B de ambas herramientas, se hizo una comparación para determinar cual (es) péptido (s) es el que tuvo mayor probabilidad de ser epítope B según ambas herramientas. Además se corroboró en la estructura tridimensional de la proteína (usando el programa Pymol) si estos epítopes se encontraban expuestos en la superficie de la proteína.

6.1.5 Predicción de Epítopes T

La predicción de posibles péptidos de la proteína que puedan ser epítopes T se realizó utilizando herramientas bioinformáticas de “aprendizaje automatizado” (learning machines) ARB (61), MHCPRED 2.0 (63), NETMHCII 2.2 (64), PROPRED (65) y NetMHCIIpan 2.1 (66) basado en el protocolo de uso de predicción de epítopes para evaluar la eficacia y cubrimiento de la vacuna Mtb72f contra tuberculosis (43). La elección de los péptidos con mayor probabilidad de ser epítope T se hizo con base al valor del IC50 que arroja cada una de las herramientas para cada posible péptido. A partir de una comparación entre las herramientas se seleccionaron los péptidos con el valor de IC50 menor para las herramientas y

se generó una tabla de epítopes promiscuos. Además se corroboró en la estructura tridimensional de la proteína (usando el programa Pymol) si estos epítopes se encontraban expuestos en la superficie de la proteína.

6.1.6 Predicción de estructura tridimensional de la proteína

La predicción de la estructura terciaria de la proteína se realizó utilizando el servidor de predicción de estructura de proteína Robetta (<http://rosetta.bakerlab.org/>) (74). La validación del modelo (s) se hizo utilizando el servidor de modelamiento por homología de proteínas SWISS MODEL (<http://swissmodel.expasy.org/>) (70).

6.1.7 Docking molecular y refinamiento de los epítopes de la proteína

Los epítopes T promiscuos determinados con las herramientas bioinformáticas fueron seleccionados del modelo 3D de la proteína y guardados como archivos PDB independientes. Una vez esto, se hizo un docking molecular entre los epítopes y moléculas MHCII HLADRB1010 (PDB: 1DLH) utilizando el servidor PatchDock (<http://bioinfo3d.cs.tau.ac.il/PatchDock/>) (75; 76) utilizando un RMSD de 4.0 (predeterminado) y un tipo de interacción antígeno-anticuerpo, con el fin de verificar si los epítopes se acoplaban de manera adecuada a los pockets de las cadena alfa y beta del complejo. En total se hicieron 15 docking, correspondientes a los quince epítopes T predichos.

Por último se hizo un refinamiento de los campos de energía del epítope con mayor promiscuidad y del epítope con mejor ubicación sobre los pockets del MHCII utilizando el software Insight II (Accelrys) en el módulo Discover (77) con el algoritmo Steepest Descent para mejorar la estabilidad de los péptidos y realizar de nuevo el docking molecular para verificar los resultados con los obtenidos antes del refinamiento. El máximo de minimizaciones de energía para cada epítope fueron diez refinamientos. Las mediciones de los puentes de hidrógeno se hicieron con base a Stern *et. al.*, 1994 (45) para comparar los resultados con estructuras cristalográficas de interacciones entre péptidos y el CMH II. La visualización de los archivos en PDB se hizo a través del programa PyMol.

6.2 Evaluación *in vitro* del efecto inhibitorio de los péptidos de la proteína

Para la realización de los estudios *in vitro* se manejó como la variable dependiente el porcentaje de invasión de la bacteria y como la variable independiente la concentración de cada uno de los péptidos a tres diferentes concentraciones (Concentración A: 2 μ M, Concentración B: 20 μ M, Concentración C: 200 μ M).

6.2.1 Macrófagos derivados de monocitos U937

Las células U937 se cultivaron en monocapa en cajas de cultivo (25 cm²) con medio RPMI completo (medio RPMI 1640 + suero fetal bovino al 10% + antibiótico al 10%); manteniéndolas en incubación a temperatura constante 37°C, con 5% de CO₂, durante todo el proceso del experimento. Una vez determinada la viabilidad de las células en las cajas de cultivo, se procedió a ampliar la cantidad de células en seis cajas de cultivo de 75 cm², las cuales se mantuvieron bajo las mismas condiciones de las cajas de 25 cm². Con ello, se tenían de 6-7 millones de células/caja de cultivo para realizar los ensayos de invasión. Luego de la formación de la monocapa celular, las U937 se desprendieron y con parte de ellas se hizo el repoblamiento celular cada 4 días para mantener su viabilidad. Rutinariamente se revisó el estado de las células para determinar la presencia o no de posible contaminación. El medio contenido en cada caja fue eliminado para dejar solo las células adheridas. Luego, se le adicionaron 2 ml de solución de tripsina (tripsina 3% y EDTA 1%) a cada caja durante 3-5 minutos para desprender las células de la monocapa. Una vez desprendidas las células, se adicionaron 3 ml de medio RPMI completo a cada caja, haciendo 2 lavados en la caja con el mismo medio y se colectó el contenido de cada caja en un tubo Falcon de 50 ml (un tubo para todas las cajas), el cual se centrifugó a 2500 rpm durante 5 minutos. El sobrenadante se descartó y se agregaron 6 ml de medio RPMI completo homogenizando al pellet de células. Se hizo conteo de células en cámara de Neubauer utilizando tinción de azul de Tripán, para determinar el porcentaje de viabilidad de las mismas. Por último a las cajas de cultivo de se adicionaron 10 ml de medio RPMI completo y el volumen de suspensión celular determinado para tener 700.000 células/caja. Por último fueron incubadas a temperatura constante de 37°C.

6.2.2 Normas de bioseguridad para el manejo de *M. tuberculosis*

Para las manipulaciones de especímenes clínicos donde no hay aerosoles, como la preparación de especímenes acidorresistentes, se recomiendan las prácticas, el equipo de contención y las instalaciones del Nivel de Bioseguridad 2 y 3. Todas las actividades que generen aerosoles deben realizarse en un Cabina de de Seguridad Biológica Clase I o II (78).

Al manipular agentes del Nivel de Bioseguridad 3 se dá mayor énfasis en las barreras primarias y secundarias para proteger al personal en áreas contiguas, a la comunidad y al medio ambiente de la exposición a aerosoles potencialmente infecciosos. Por ejemplo, todas las manipulaciones de laboratorio se deben llevar a cabo en cabinas de extracción de aerosoles. Las barreras secundarias para este nivel incluyen el acceso controlado al laboratorio y requisitos de ventilación que minimizan la liberación de aerosoles infecciosos desde el laboratorio (78). De igual forma el uso de bata de laboratorio, batas quirúrgicas, guantes, tapabocas y polainas es indispensable para trabajar de manera correcta con este tipo de patógeno.

6.2.3 Cultivo de *M. tuberculosis* H37Rv

El cultivo de las cepas de *M. tuberculosis* H37Rv se hizo en medio de cultivo líquido 7H9 (Sulfato de amonio, Acido L-glutámico, Citrato de sodio, Piridoxina, Biotina, Fosfato disódico, sulfato de magnesio, sulfato de zinc) con suplemento OADC (Albumina, Dextrosa, Catalasa y acido Oleico) al 10%. En un frasco de cultivo de 2000 mL se prepararon 200 ml de medio 7H9 al cual se le agregó suplemento OADC al 10% (20 ml) y 200 µl de *M. tuberculosis* H37Rv teniendo precaución de que todo el material estuviera libre de contaminación. El frasco de cultivo se incubó a 37°C, con agitación constante.

6.2.4 Cosecha y marcaje de *M. tuberculosis* H37Rv

La cosecha de *M. tuberculosis* H37Rv (en fase estacionaria) se hizo a partir de los cultivos de la bacteria realizados (15 días después de la siembra, de acuerdo con la curva de crecimiento), colocando 30 ml de la suspensión del cultivo en tubos Falcon, los cuales fueron centrifugados a 12,500 rpm durante 20 minutos, para su posterior eliminación de sobrenadante y

resuspensión en 1ml de PBS1X. El pellet resultante se almacenó en tubos eppendorf de 1.5 ml a -20°C hasta su uso.

La concentración de bacteria que se utilizó por ensayo fue calculado por escala de Mcfarland a partir de la densidad óptica medida el día en que se realizó la cosecha de la bacteria. Así, la micobacteria se marcó con SYBR Safe 20X (3-4ul) y fue incubada durante 20 minutos a 37°C para su uso en el ensayo de invasión. Luego, la bacteria marcada se centrifugó a 6000 RPM durante 5 minutos y se eliminó el sobrenadante para resuspender el pelett en medio RPMI de ensayos (suplementado con SFB al 10% sin antibióticos). Por otro lado, el uso de fenol 5% (mata la bacteria) permitió la manipulación segura de materiales que estuviesen en contacto con la bacteria o que la contuvieran.

6.2.5 Síntesis de péptidos de la proteína Rv3494c

Los péptidos de la proteína Rv3494c fueron sintetizados por el Grupo Funcional de síntesis química de la Fundación Instituto de Inmunología de Colombia (FIDIC) empleando la técnica de síntesis en fase sólida (79). Para cada péptido se realizaron análisis de HPLC y espectrometría de masa para determinar su pureza. Los péptidos liofilizados fueron resuspendidos en 20 μl de DMSO y 980 μl de buffer (HBSS) y se almacenaron a -20°C hasta su uso.

6.2.6 Ensayos de Invasión

En cajas de 24 pozos, se adicionaron 3 ml de medio RPMI completo y el volumen de suspensión celular determinado para obtener 250.000 células/pozo. Se incubaron a 37°C durante 24 horas, con el fin de obtener una monocapa de células adherida en el fondo de cada uno de los pozos. Los ensayos de invasión que se realizaron siguiendo el protocolo estandarizado en FIDIC (80), con algunas modificaciones detalladas a continuación. En cada ensayo de invasión se evaluaron 3 péptidos a 3 concentraciones diferentes y por triplicado, además se contó con los controles positivo de invasión y de inhibición. Cada uno de los péptidos a ensayar se llevó a 1mL (980 μl de HBSS y 20 μl de DMSO). El volumen utilizado de cada uno de los péptidos, para obtener las concentraciones de 2 μM , 20 μM y 200 μM , se determinó con base al peso molecular de cada péptido, utilizando HBSS.

Luego de 24 horas, cuando las células U937 han formado la monocapa en cada pozo se eliminó el medio sobrenadante y se adicionaron 200 µl de cada uno de los péptidos a las diferentes concentraciones. A los pozos del control positivo se le agregó 200 µl de HBSS y a los pozos del control negativo se les adicionó 200 µl de Citocalasina D (que inhibe la polimerización de los filamentos de actina, evitando el movimiento del citoesqueleto y por ende la fagocitosis) a 10 µM (concentración final). Luego, las células fueron incubadas a 4°C por 2 horas.

Posteriormente, se preparó una solución de medio RPMI y bacteria marcada, calculando el volumen final teniendo en cuenta que en cada pozo se adicionarían 300 µl de esta suspensión. Esta solución se sonicó durante 15 segundos con el fin de disgregar los grupos de bacterias. Después de incubar por 2 horas los péptidos con las células, el medio de cada pozo fue eliminado y remplazado con 300 µl de suspensión bacteriana, a un MOI (del inglés: *Multiplicity of infection*) de 1:100 (100 bacterias por cada célula), siendo 25.000.000 el número aproximado de bacterias por pozo. Las cajas se incubaron a 37°C durante 24 horas.

Pasadas las 24 horas de incubación de las células U937 (tratadas con los péptidos) con micobacteria marcada, se eliminó el medio y se realizaron dos lavados con 700 µl de HBSS (cada lavado), lo que permite eliminar la bacteria que no invadió las células. A cada pozo se le agregaron 150 µl de tripsina (Tripsina 3% y EDTA 1%) por 1 minuto para desprender las células. Luego, se adicionaron 300 µl de medio RPMI a cada pozo, homogenizando el contenido. Este contenido se pasó a un tubo eppendorf de 0.5 ml (una bala por pozo) y se centrifugó a 2500 rpm durante 5 minutos. Posteriormente, se eliminó el sobrenadante y se resuspendió el contenido en 100 µl de paraformaldehído al 4% para fijar las células e inactivar la bacteria extracelular y la bacteria restante de los lavados. Esta suspensión se pasó a tubos de citometría (un tubo por pozo), que se incubaron a -4°C durante 2 horas para su lectura en el citómetro de flujo. 5 minutos antes de la lectura en el citómetro se adicionaron 150 µl de azul de metileno filtrado a cada una de las muestras. La lectura se llevó a cabo en un citómetro de flujo FACScan y un total de 10000 eventos/muestra fueron adquiridos para su análisis. El análisis de estos datos se realizó usando el software FlowJo 7.2.

El cálculo del porcentaje de inhibición (I) para cada replica de cada tratamiento se hizo a partir de los datos obtenidos de los porcentajes de invasión, empleando la siguiente ecuación:

$$I=100-((\text{porcentaje promedio de invasión en presencia de inhibidor} \times 100)/\text{promedio de invasión en ausencia de inhibidor})$$

Una vez obtenido el porcentaje de inhibición para cada replica se pudo calcular el promedio de inhibición para el tratamiento.

6.2.7 Ensayo de citotoxicidad

Para determinar el efecto tóxico que los péptidos pudieran generar en las células U937, se realizó un ensayo de reducción de Resazurina, utilizando el kit toxicológico *in vitro* (Base Resazurina) de SIGMA (84). Cada péptido fue probado a una concentración de 200 μ M (la más alta usada en ensayos de invasión). Como control positivo de bio-reducción del colorante se utilizaron células sin adicionarles péptido. De igual forma se tuvo un control de solo medio RPMI completo y de solo buffer HBSS con el fin de determinar la posible influencia que tengan estas soluciones sobre los péptidos.

En una placa de 96 pozos se sembró 250.000 células/pozo por duplicado para cada uno de los péptidos y el control positivo de reducción del colorante. Las células se incubaron a 37°C durante 24 horas. Pasado este tiempo se retiró el sobrenadante y se hizo un lavado con 200 μ l de HBSS para retirar las células que no estuvieran adheridas a la monocapa. Luego se adicionó 100 μ l de cada péptido a cada uno de los pozos, teniendo en cuenta que los péptidos se prepararon de la misma manera que en los ensayos de invasión, es decir, con base a su peso molecular y llevando el volumen final utilizando HBSS. Así pues las células fueron incubadas a 4°C durante 1 hora. Luego se sacó el sobrenadante y se adicionó 10 μ l de colorante Resazurina (10% del volumen de péptido añadido a las células), llevando a un volumen final de 100 μ l con medio RPMI completo e incubando a 37°C durante 3 horas.

La lectura de resultados se llevo a cabo en un lector de ELISA utilizando una longitud de onda de emisión de 580 nm y de excitación de 535 nm.

6.3 Análisis estadísticos

El análisis estadístico de los datos se realizó utilizando el software GraphPad PRISM 5.01 mediante una prueba de Dunnet, que compara cada uno de los tratamientos con el control negativo de inhibición. Se maneja un intervalo de confianza del 95%, donde $p \leq 0.01$ es altamente significativo y $p \leq 0.05$ es significativo.

7. Resultados

7.1 Caracterización bioinformática de la proteína Rv3494c

7.1.1 Localización subcelular de la proteína

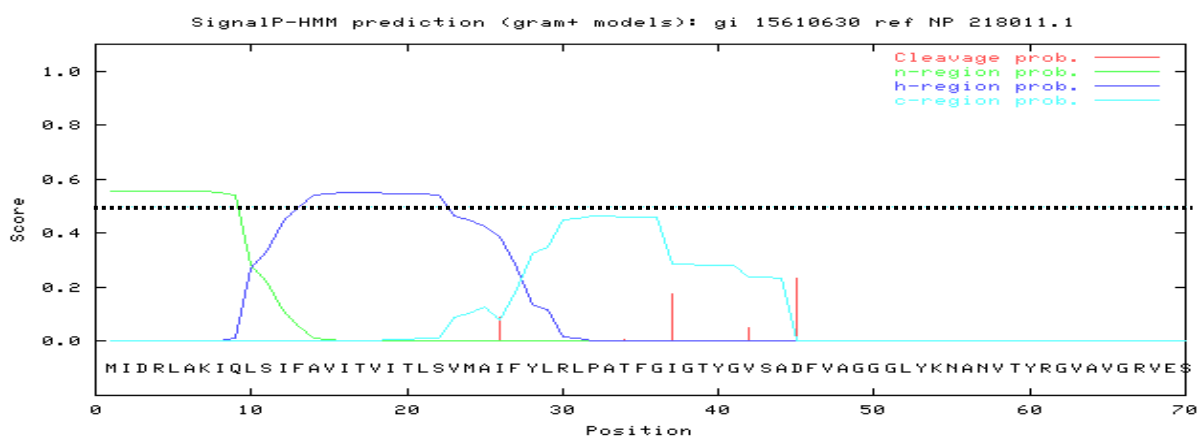
La predicción de una posible localización subcelular de la proteína en *Mycobacterium tuberculosis* H37Rv utilizando los servidores Gpos-PLoc, Psortb y TBpred, un servidor especializado en detectar la localización de proteína de micobacteria, demostró que posiblemente la proteína este localizada en la zona extracelular de la bacteria (Gpos-Ploc) o en la membrana citoplasmática (Psortb y TBpred). El servidor Gpos-PLoc determino una localización de la proteína en la zona extracelular basado en Ontología Génica (GO). El servidor Psortb dio puntuaciones menores a 1 para ubicar la proteína en el citoplasma, en la pared celular y en la zona extracelular, mientras que la ubicación en la membrana citoplasmática tuvo una puntuación de 9.26, lo que muestra la probable ubicación de la proteína. El servidor TBpred también asigna puntuaciones a las diferentes localizaciones subcelulares dependiendo de las características de los aminoácidos usando algoritmos SVM (support vector machine). Así, la zona de mayor puntaje fue la de membrana (proteína de membrana) con un valor de 2.43, mientras que las localizaciones en el citoplasma, unida a la membrana o proteína secretada tuvieron puntuaciones negativas (menores a cero), por lo que la predicción fue de proteína de membrana.

7.1.2. Predicción de péptido señal y tipo de secreción

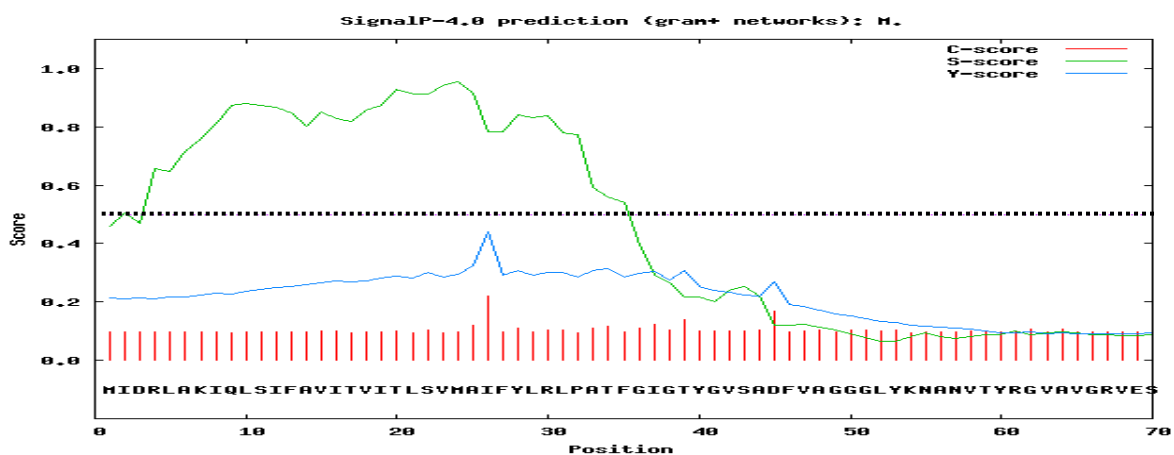
La figura 2 muestra el resultado para la predicción de secuencia señal en la proteína Rv3494c con la herramienta SignalP 3.0, basada en modelos ocultos de Markov (HMM). Se predice que la proteína presenta una secuencia señal de vía clásica, con una puntuación de 0.554 (figura 2), con dos posibles sitios de clivaje: uno entre los aminoácidos 25 y 26 (con una puntuación de 0.220), y otro, entre los aminoácidos 44 y 45 (con una puntuación de 0.234). Ambos sitios de clivaje presentan una alanina en la posición -1 pero no tienen otra alanina en la posición -3 de esos sitios, como es de esperarse en péptidos señal tipo I de vía clásica. Sin embargo esto no se presenta en algunos casos de secuencia señal tipo I (42;55;56). La herramienta SignalP 4.0 se utilizó para corroborar la predicción de secuencia señal por el servidor SignalP 3.0 y para discernir si esa secuencia señal hace referencia a una región

transmembranal. La opción que ofrece SignalP 4.0 de hacer predicciones teniendo y no teniendo en cuenta regiones transmembranales (TM) permitió determinar la presencia de una secuencia señal y no de una hélice TM (figura 2). La probabilidad de presencia de la secuencia señal sin tener en cuenta regiones TM fue de 0.570, con sitios de clivaje entre los aminoácidos 25 y 26 (como en SignalP 3.0); la probabilidad de presencia de péptido señal teniendo en cuenta regiones TM fue de 0.45, lo que muestra que el péptido señal aparece cuando no hay regiones TM. Por su parte, el programa SecretomeP 2.0 predijo que esta proteína posiblemente es secretada por vía no clásica, arrojando una puntuación de 0.68. Por otro lado el servidor TatP 1.0 mostro que la proteína no presenta secuencia señal de arginina doble, ya que las probabilidad de que ocurra ese evento es menor a 0.5 (figura 2). El servidor LipoP 1.0 hizo la predicción para una hélice transmembranal, lo cual descarta una secuencia señal de lipoproteína; sin embargo el servidor detecto un posible sitio de clivaje entre los aminoácidos 25-26. El umbral de aceptación para todas herramientas fue de 0.5; por encima de este valor la predicción se hizo valida.

A)



B)



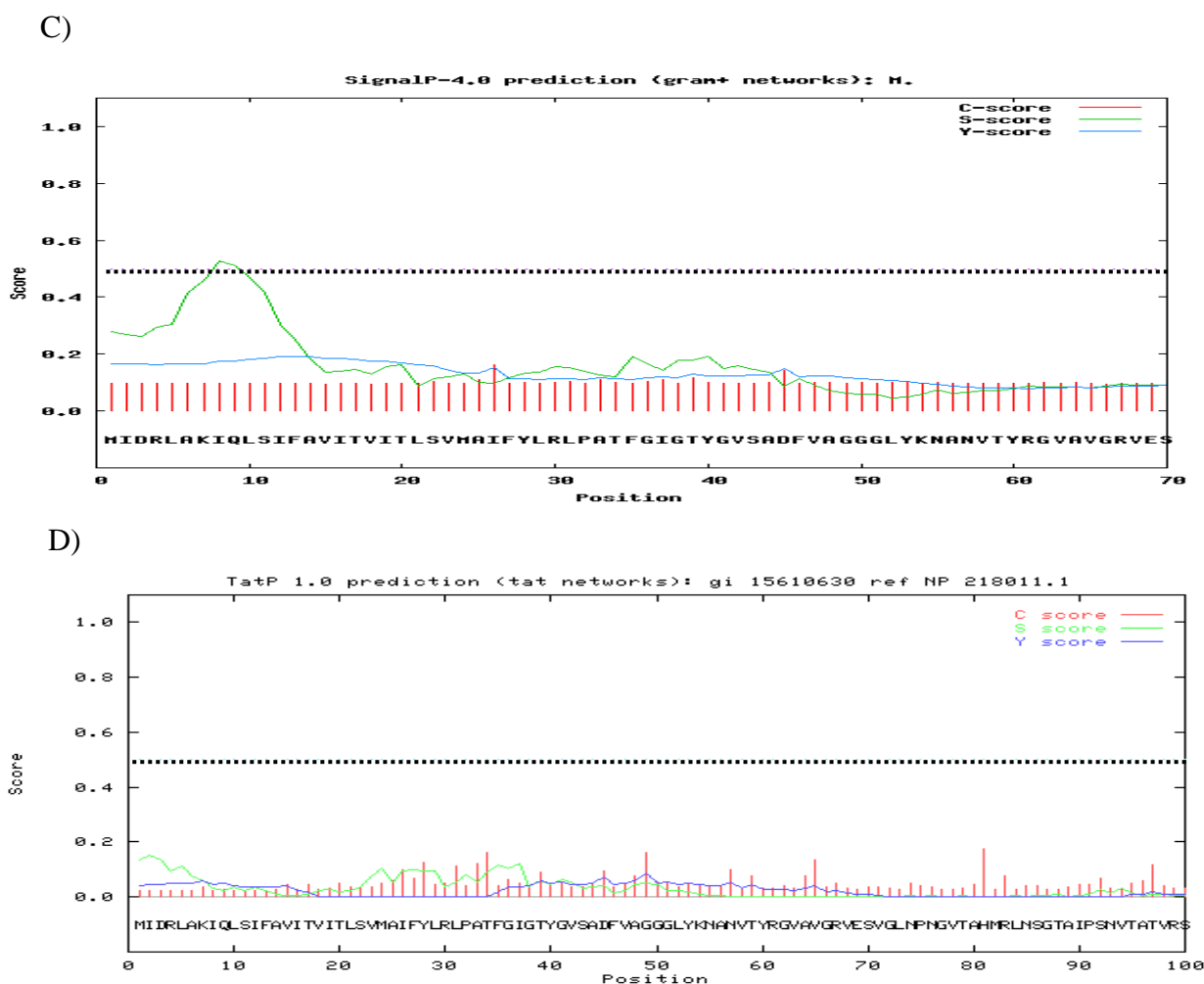


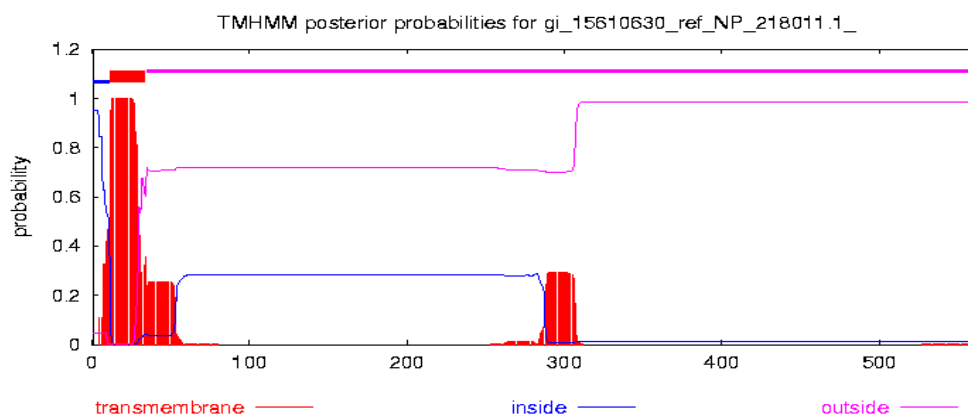
Figura 2. Predicciones de péptido señal. A) Predicción con SignalP 3.0 (HMM). B) Predicción con SignalP 4.0 (redes neurales artificiales sin tener en cuenta regiones transmembranales). C) Predicción con SignalP 4.0 (redes neurales artificiales teniendo en cuenta regiones transmembranales). D) Predicción con TatP 1.0. La curva verde significa probabilidad de presencia de péptido señal y las rectas rojas probabilidad y posición de sitio de clivaje.

7.1.3 Predicción de hélices transmembranales

En cuanto a la predicción de regiones transmembranales utilizando el servidor TMHMM 2.0, se evidenció que probablemente la proteína tenga organización de la siguiente manera (figura 3): del aminoácido 1 al 11 la proteína es citoplasmática, del aminoácido 12 al 34 se encuentra una hélice transmembranal y del aminoácido 35 al 564 esta la parte de la proteína que esta expuesta al exterior celular y probablemente interactúe con células hospederas. Además, la predicción de regiones y hélices transmembranales con el servidor Phobius mostró la misma organización de la proteína y también que puede haber una hélice transmembranal que inicia

en el aminoácido 12 (en la parte citoplasmática de la bacteria) y termina en el aminoácido 31 en la parte no citoplasmática.

A)



B)

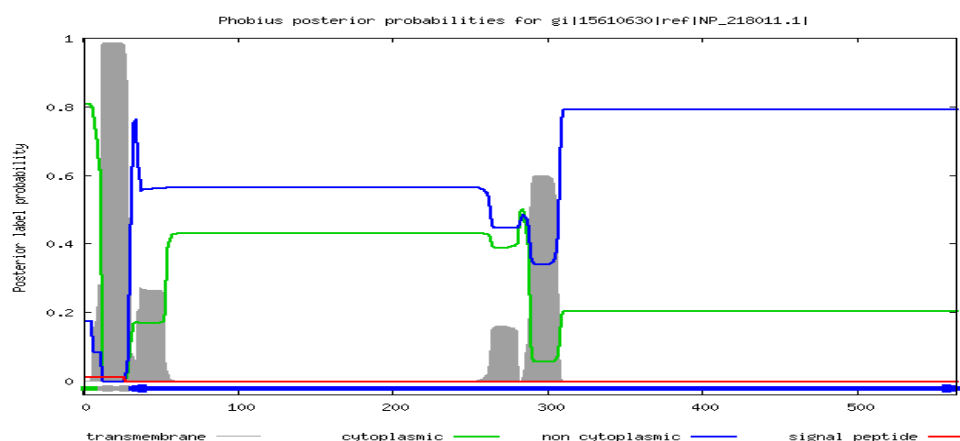


Figura 3. Predicción de hélices y regiones transmembranales. A) Predicción con TMHMM, donde la curva y la línea azul, roja y fucsia muestran la probabilidad y la distribución de aminoácidos en las regiones citoplasmáticas, transmembranales y extracelulares, respectivamente. B) Predicción con Phobius, donde la curva y la línea verde, gris y azul muestran la probabilidad y la distribución de aminoácidos en las regiones citoplasmáticas, transmembranales y extracelulares, respectivamente; la línea roja representa la probabilidad y ubicación de péptido señal.

7.1.4 Predicción de Epítopes B

La predicción de Epítopes B que se realizó con el servidor Bepipred 1.0 y con el software Antheprot mostró que hay tres secuencias de aminoácidos (de 20 aminoácidos cada uno) que podrían ser epítopes B. La predicción con el programa Antheprot se hizo con base en los perfiles fisicoquímicos: antigenicidad de Parker, hidrofiliadad y accesibilidad a solvente, los

cuales en conjunto permiten definir las áreas de la proteína que se encuentran en la superficie (antigenicidad de Parker) y cuyas regiones son hidrofílicas, flexibles y accesibles al solvente (81). Los umbrales de validez son >50 , >-0.1 y >0.62 , respectivamente. El servidor Bepipred además de hacer este mismo análisis fisicoquímico, utiliza algoritmos como Modelos Ocultos de Markov (HMM) para realizar predicciones más robustas (48). En los dos programas, para que una secuencia fuera seleccionada como posible epítipo B debería contener por lo menos 8 aminoácidos contiguos favorables. A partir de lo anterior se determinaron solo 4 epítopos B cuya secuencia se muestra en la tabla 1. Las secuencias de estos epítopos presentan altos valores de antigenicidad, hidrofiliidad y accesibilidad al solvente, lo que indica que los 4 epítopos deberían ubicarse en la superficie de la proteína.

Tabla 1. Epítopos B predichos utilizando el consenso del servidor Bepipred 1.0 y el programa Antheprot.

POSICION	SECUENCIA	ANTIGENICIDAD	HIDROFILIDAD	ACCESIBILIDAD A SOLVENTE
355-375	TAQNDPSTVRGARNYPCQEF	92	0.39	1.00
494-514	PPPPPEGTGPPPGPAPGPQP	90	-1.26	1.33
115-135	ENPSSTKLRNGFRIQRQNTR	92	0.22	3.16
533-553	DPAGGTGIFAPGMTGASSAE	81	0.61	0.52

7.1.5 Predicción de Epítopos T

La predicción de epítopos T (nonámeros) se hizo con los programas ARB, MHCPRED 2.0 (que permite seleccionar los aminoácidos para de las posiciones 1,4, 6 y 9 del péptido, con el fin de determinar el mejor P-anchor o aminoácidos de esas posiciones que den estabilidad de unión a los “pockets” del MHC II), NETMHCII 2.2, NETMHCIIPAN 2.1 y PROPRED (que tiene un P-anchor obligatorio, es decir que la posición 1 del epítipo predicho solo tendrá permitido tener los aminoácidos M(Met), F(Phe), I(Ile), L(Leu), V(Val), W(Trp) y Y(Tyr) (62). Los valores de IC50 tomados para las herramientas programas ARB, MHCPRED 2.0 y NETMHCII 2.2 estuvieron entre 0-50, mientras que para el servidor NETMHCIIPAN 2.1 los valores IC50 analizados estuvieron entre 0-1000, ya que arrojó valores mucho más altos que los otros tres servidores. Para el servidor PROPRED, los valores de análisis fueron los porcentajes de mayor puntuación para los péptidos que arrojó el mismo. Todos los programas mostraron un margen de error de un aminoácido al realizar la predicción (excepto ARB). TEPITOPE es un software de alta predicción de epítopos T, el cual no se pudo adquirir; sin embargo el servidor PROPRED maneja las matrices de este software.

Los alelos para los cuales se hizo la predicción de epítopes se encuentran dentro de los más comunes en zonas de alta incidencia de tuberculosis (82) y son los que mas se presentan dentro de los mismos servidores. Estos son HLADRB1: 0101, 0301, 0401, 0404, 0405, 0701, 0802, 0901, 1101, 1301, 1302, y 1501.

En los servidores ARB y NETMHCIIPAN 2.1 aparecen los 12 alelos seleccionados, en el servidor NETMHCII 2.2 aparecen 11 alelos (excepto el 1301), en el servidor PROPRED también aparecen 11 alelos (excepto el 0901) y en el servidor MHCPRED 2.0 aparecen solo los alelos DRB10101, DRB10401 y DRB10701. Los resultados obtenidos se muestran en la tabla 2.

Tabla 2. Numero total de epítopes predichos para cada uno de los alelos y para cada uno de los servidores utilizados.

ALELO DRB1	SERVIDOR				
	ARB	MHCPRED 2.0	NETMHCII	NETMHCIIPAN	PROPREL
0101	38	13	3	3	12
0301	7	0	0	1	12
0401	12	13	0	0	12
0404	4	0	0	0	14
0405	7	0	0	0	15
0701	8	13	1	0	14
0802	6	0	0	1	8
0901	9	0	0	1	0
1101	4	0	0	4	12
1301	0	0	0	0	11
1302	8	0	1	1	14
1501	6	0	0	0	14

Para cada uno de los 12 alelos se obtuvo la lista de las secuencias predichas como epítopes T específicos de ese alelo (Anexo 11.1). Se diseñó una matriz que contenía todos los resultados, con el fin de determinar cuales epítopes están presentes en al menos 4 de los 12 alelos y fueron determinados por los menos en 2 de las herramientas bioinformáticas descritas (se escogió un umbral de 2 herramientas ya que la herramienta NETMHCII 2.2 solo predijo epítopes para 3 alelos) (Tabla 3).

Tabla 3. Epítopes predichos con más frecuencia en los diferentes alelos y por lo menos en 2 de las herramientas de predicción utilizadas. Con X se representa el número de aminoácidos del epítope específicos a los “pockets” 1, 4, 6 y 9 del MHC II.

POSICION	SECUENCIA	ALELO
10XX	LSIFAVITV	0401, 0404, 0701, 0802, 01101, 01301, 01501
12XX	FAVITVITL	0101, 0405, 0701, 01101
15XX	VITVITLSV	0101, 0401, 0404, 0405, 0701, 01101, 01501
18XX	VITLSVMAI	0101, 0301, 0401, 0404, 0405, 0701, 0802, 01101, 01301, 01302, 01501
21XX	LSVMAIFYL	0404, 0405, 0802, 0901
26XX	IFYLRLPAT	0101, 0401, 0404, 0802, 01101
27X	FYLRLPATF	0101, 0401, 0802, 0901, 01101, 01302
28X	LRLPATFGI	0301, 0404, 0405, 0701, 01301, 01302, 01501
60XX	YRGVAVGRV	0101, 0301, 0405, 0701, 0802, 0901, 01101
78X	VTAHMRLNS	01101, 01301, 01302, 01501
82XXX	MRLNSGTAI	0101, 0301, 0401, 0701, 01302, 01501
126XX	FRIQRQNTR	0101, 0401, 01101, 01301, 01301
128XXX	IQRQNTRIG	0404, 0405, 01301, 01302
142XXX	LRQAETLLG	0301, 0401, 0404, 0405, 01101, 01301, 01302, 01501
179X	LIESARLLV	0301, 0701, 0802, 01501
260X	IRPSFPALA	0101, 0301, 0401, 0404, 01101
278X	VIYHKSIEQ	0401, 0404, 0405, 01101, 01501
288XXXX	LLVVFPALF	0101, 0401, 0405, 0701, 01501
289XXXX	LVVFPALFA	0101, 0401, 0404, 0405, 0802, 01101, 01301, 01302, 01501
296XXXX	FAAITSAG	0101, 0401, 0405, 0701
362XX	VRGARNYPT	0101, 0401, 0404, 0405, 01501
390XXX	YVPVGTNPW	0401, 0404, 0405, 0701, 0802, 01101
523XXXX	IYDQLSGAF	0101, 0401, 0701, 0901, 01501

De igual forma se estableció las secuencias de aminoácidos que tuvieron una mayor predicción de epítopes T para la mayor cantidad de alelos (figura 4), esto con el fin de establecer cuáles son las zonas de la secuencia de la proteína de mayor relevancia cuando se genera la respuesta celular ante la entrada de la bacteria al organismo.

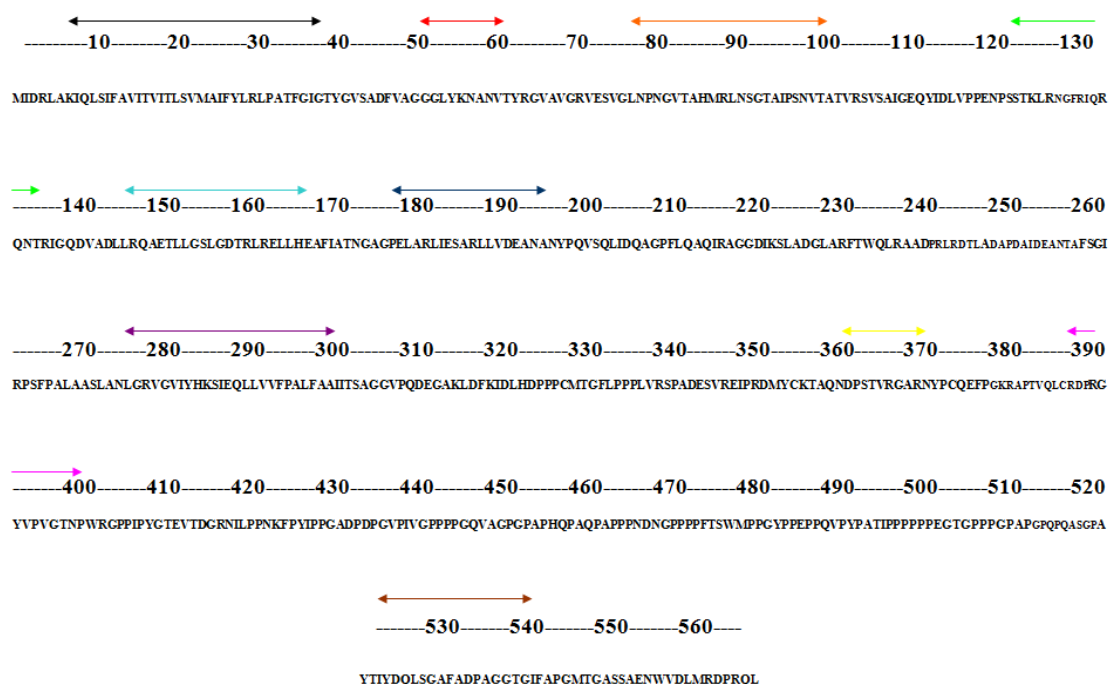


Figura 4. Regiones de la proteína con mayor predicción de epitopes T para los 12 epitopes analizados. Línea negra: presente en todos los alelos (7-37); Línea roja: todos los alelos menos 0404 y 01302 (50-60); Línea naranja: todos los alelos menos 0401, 0404, 0802 y 0901 (77-100); Línea verde: todos los alelos menos 0404, 0802, 0701, 0901 (121-137); Línea aguamarina: todos los alelos menos 0101, 0802, 0701 y 0901 (145-168); Línea verde azulada: alelos 0301, 0701, 0401 y 01302 (179-195); Línea morada: todos los alelos (278-300); Línea amarilla: alelos 0101, 0401, 0405, 01501 (360-370); Línea fucsia: alelos 0401, 0405, 0701, 0802, 01101 (390-400); Línea café: alelos 0101, 0401, 0801, 0901, 01302, 01501 (520-540).

7.1.6 Predicción de estructura 3D de la proteína

En cuanto al modelamiento de la proteína, se hizo la predicción empleando el servidor Robetta, el cual es uno de los más robustos. La validación de los modelos predichos con Robetta se hizo a través la opción de validación de estructura del servidor SWISS MODEL, el cual permite verificar el modelo proteico a través de comparaciones con estructuras cristalográficas de la base de datos PDB. Además revisa las fuerzas atómicas y de torsión que rigen el modelo y crea un análisis en el Plot de Ramachandran para verificar la distribución optima de aminoácidos a lo largo del modelo. La predicción del modelo de la proteína a través de métodos *ab initio*, donde se parte de la ausencia de moldes para construir una estructura tridimensional, que enmarcan también metodologías de modelamiento por homología fue la más eficaz, ya que también se hizo el modelamiento con otros servidores como Swiss Model, I Tasser y EsyPred 3D, los cuales a través de metodologías como modelamiento por homología de secuencias y reconocimiento de moldes no iniciaron la construcción de una estructura proteica. De igual forma el servidor Phyre2 modeló solo el

29% de la proteína usando metodologías de reconocimiento de moldes y métodos *ab initio*. A través de la metodología GINZU, que permite buscar dominios remotamente relacionados filogenéticamente pero estructuralmente similares (usando BLAST, PSIBLAST, FFAS03, 3D-Jury), los dominios utilizados por Robetta como molde para el modelamiento fueron: familia de proteínas MCE (del aminoácido 1-132) (codigo Pfam PF02470), proteína de señalización de la alga roja *Galdieria sulphuraria* (del aminoácido 133-212) (PDB: 3CAZ A) y proteína estructural/célula de adhesión vinculina de humano (del aminoácido 213-337) (PDB: 1YDI A). Para los últimos 227 aminoácidos no se obtuvo un molde apropiado, por lo que no fueron modelados por el servidor. La estructura secundaria de la proteína predicha (figura 5), que también es predicha a través del servidor Robetta (utiliza los predictores de estructura secundaria Psipred, Sam y Jufo), en su mayoría esta conformada por hélices alfa (H) en los aminoácidos 1-30 y 140-300, y por una pequeña parte de hojas plegadas B (E) en los aminoácidos 40-130.

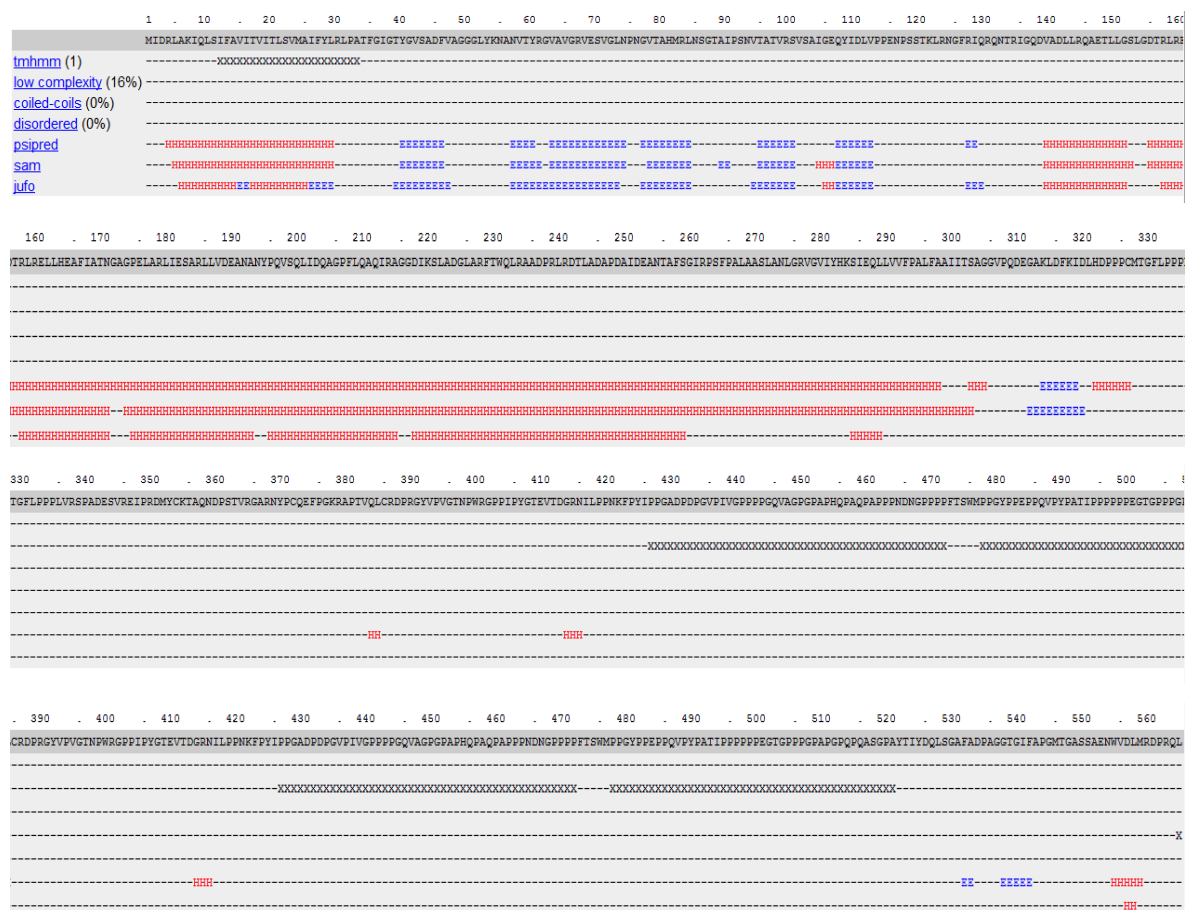


Figura 5. Estructura secundaria de la proteína predicha por Robetta a través de la búsqueda en tres servidores (Psipred, sam, jufo). En rojo se observan las predicciones de hélice alfa y en azul las de hojas plegadas.

En total se obtuvieron 5 estructuras proteicas con el servidor Robetta, las cuales fueron validados con el servidor SWISS MODEL, con el fin de escoger el mejor modelo. Para ello, la validación de la estructura tridimensional se basa en valores dados al modelo acorde a los siguientes coeficientes (a continuación se muestran los resultados para el mejor modelo de los cinco validados):

ANOLEA (Atomic Non-Local Environment Assessment): Este indicador muestra la energía de interacción favorable para cada uno de los átomos de la molécula. La mayoría de aminoácidos del modelo presentan poca energía favorable para que cada átomo interactúe adecuadamente con otros átomos en la proteína (grafica con curvas rojas) (Figura 6).

QMEAN4: Evalúa potenciales de ángulo de torsión por cada tres aminoácidos. Se evalúa la interacción atómica de la proteína y los ángulos de torsión que presenta la estructura secundaria. El color rojo de la curva indica bajos potenciales de ángulo de torsión. La proteína presenta buenos potenciales de ángulo de torsión a lo largo de la secuencia (Figura 6).

GROMOS: Valida el resultado de ANOLEA (figura 6).

QMEAN6 : Incorpora 6 patrones de análisis geométrico (Qmean4 incorpora solo 4). Muestra el grado de fidelidad del modelo comparado con estructuras cristalográficas disponibles en PDB. El valor para este modelo fue 0.521 (para ser un buen modelo el valor de este coeficiente debe estar entre 0.7 y 1).

QMEAN6 (Z SCORE): Muestra una comparación entre el modelo predicho y estructuras cristalográficas de rayos X de proteínas de PDB. El modelo predicho tiene un valor de -2.68, lo que indica que el modelo no muestra una estructura idónea para la proteína (figura 7).

DFIRE: mide interacciones Van der Waals, fuerza de interacción o repulsión por enlaces covalentes o interacciones electrostáticas, e interacciones electrostáticas. El valor dado para el modelo fue de -569.8 (el valor de un buen modelo debe ser menor a -1000) indicando que las interacciones dadas son buenas.

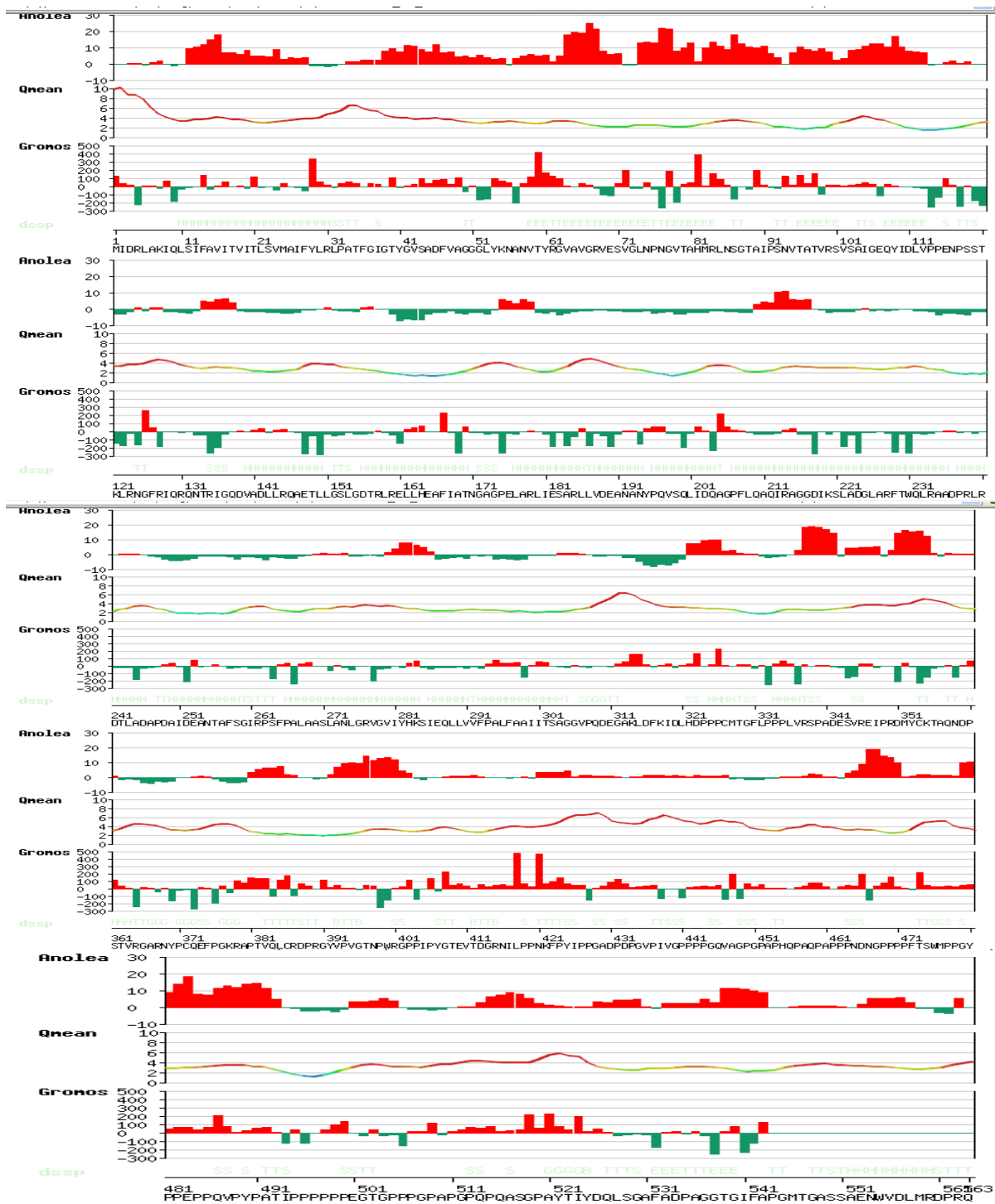


Figura 6. Coeficientes QMEAN, Anolea y Gromos predichos para el mejor modelo de la proteína con SWISS MODEL

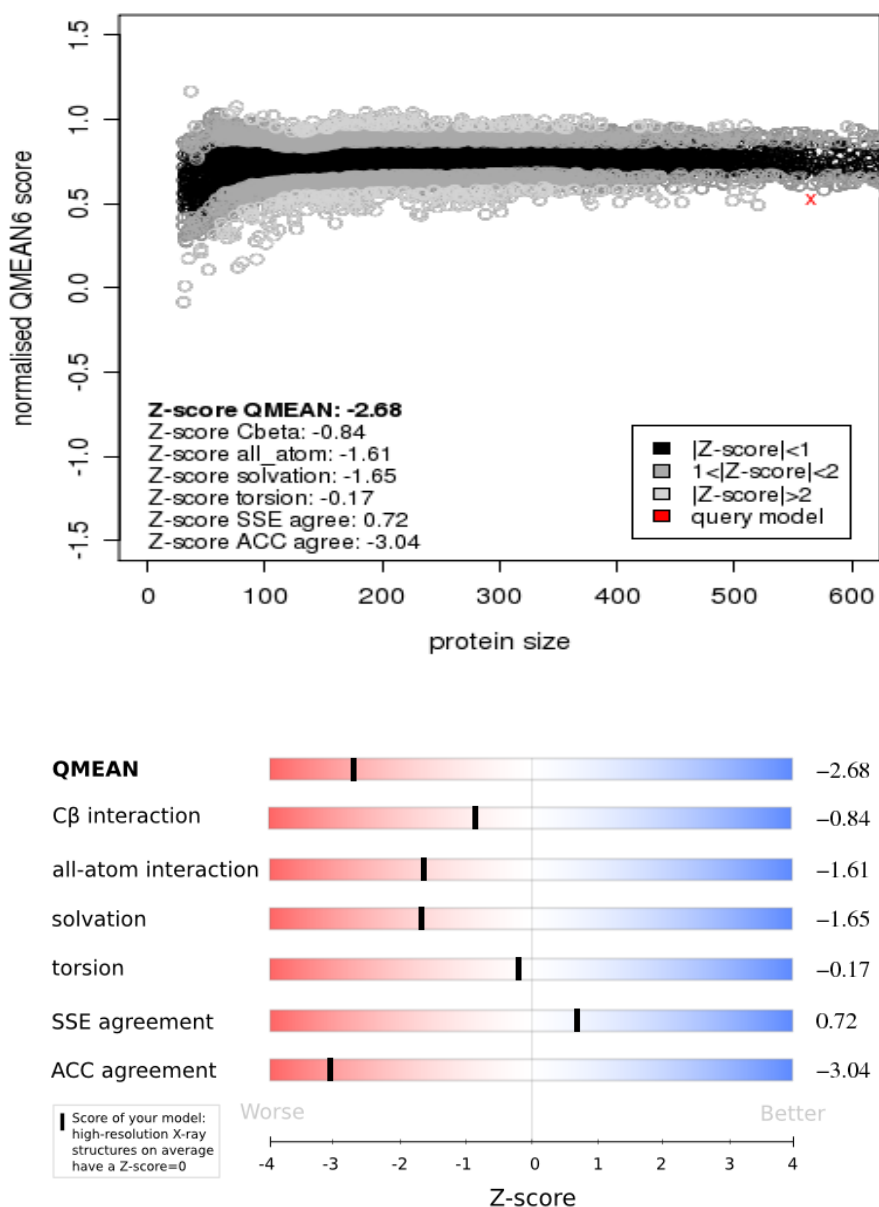


Figura 7. Coeficientes Z (QMEAN6) predicho para el mejor modelo de la proteína con SWISS MODEL. A) Comparación del modelo con estructuras cristalográficas almacenadas en PDB. B) Coeficientes de potenciales geométricos del modelo evaluado a partir de estructuras cristalográficas depositadas en el PDB; un buen modelo tendría sus valores en 0.

RAMACHANDRAN: Tiene una distribución en zonas favorables para los ángulos diedricos de los enlaces peptídicos de la proteína (es mayor al 90 %). En los plots rojos se ven la mayoría de aminoácidos (puntos), lo que indica que los aminoácidos de la proteína se unen de manera adecuada (Figura 8) y refleja una validación positiva del modelo proteico.

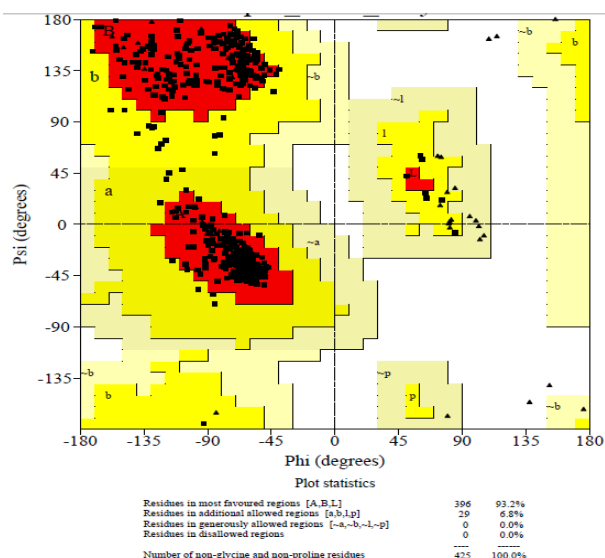


Figura 8. Plot de Ramachandran predicho para el mejor modelo de la proteína con SWISS MODEL. El plot rojo de la parte superior indica distribución de hojas plegadas β , el plot de la parte inferior indica distribución de aminoácidos en hélice alfa hacia la derecha y el Plot del segundo cuadrante (superior derecho) indica distribución de hélices alfa hacia la izquierda.

Por último, con el programa Pymol se visualiza el modelo final de la proteína (figura 9), donde se muestra el péptido señal predicho entre los aminoácidos 12-34 y las diferentes conformaciones estructurales de la proteína.

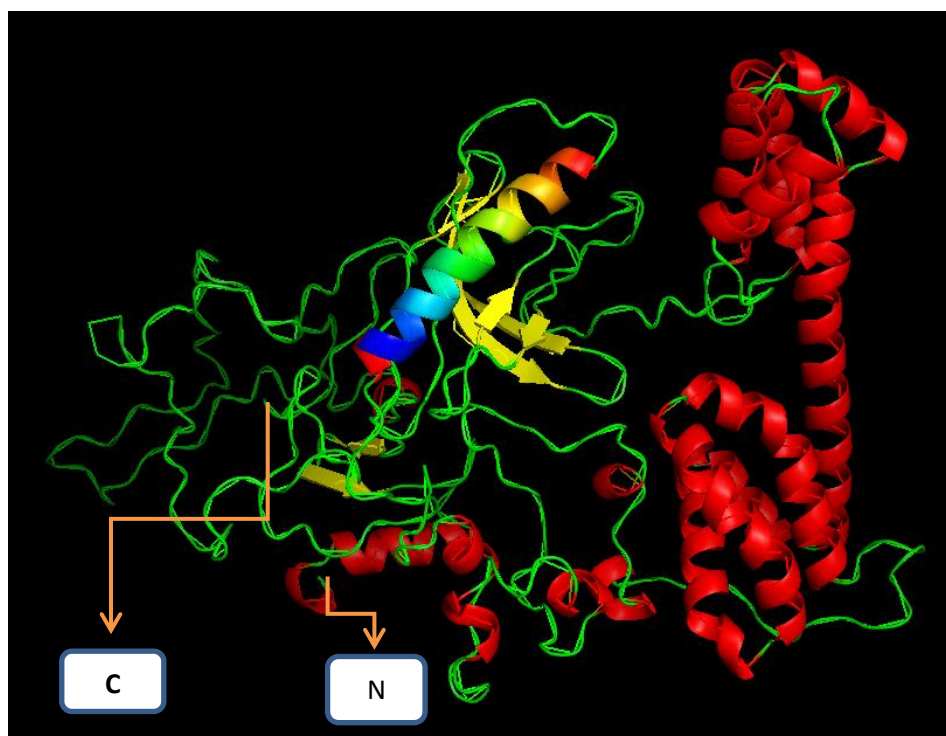


Figura 9. Modelo final de la proteína. La hélice alfa de colores representa el péptido señal predicho anteriormente (aminoácidos 12-34). En rojo se muestran las hélices alfa y en amarillo las hojas plegadas β . Las líneas verdes indican las zonas modeladas con estructura secundaria aleatoria.

7.1.7 Docking molecular y refinamiento de epitopes

Los epítopes T promiscuos determinados con las herramientas bioinformáticas fueron seleccionados y extraídos del modelo 3D de la proteína y guardados como archivos PDB independientes. Con estos segmentos independientes se hizo un docking molecular entre los epítopes y el MHCII HLADRB1010 (PDB: 1DLH) utilizando el servidor PatchDock utilizando un RMSD de 4.0 (predeterminado) y un tipo de interacción antígeno-anticuerpo, con el fin de verificar si los epítopes se acoplaban de manera adecuada a los pockets de las cadena alfa y beta del CMH II (Anexo 11.2). Al cambiar el tipo de interacción antígeno-anticuerpo por una interacción proteína-pequeño ligando las probabilidades de reconocer el CDR del complejo se disminuirían, por lo que la interacción se podría dar en otro sitio del CMH II. El docking molecular que se hizo con cada uno de los 15 epítopes T predichos y la molécula del complejo mayor de histocompatibilidad II HLA DRB1 mostró 14 regiones CDR (Complementary-Determining Regions, en inglés), que corresponden al loop donde se presenta la interacción entre antígenos y anticuerpos (75); formado por los aminoácidos 51 y 76 de la cadena alfa y los aminoácidos 59 y 82 de la cadena beta. Sin embargo, no se presenta un acople adecuado de los aminoácidos que conforman cada uno de los epitopes y los respectivos bolsillos de la molécula CMHII. En la figura 10A se muestran los resultados del docking correspondientes al epítipo más promiscuo (posición ¹⁸VITLSVMAI²⁶) y en la figura 10B los resultados del epítipo que mejor acople tuvo a los bolsillos del CMHII (posición ⁷⁸VTAHMLNS⁸⁶). El epítipo YRGVAVGRV, correspondiente a los aminoácidos 60 a 68 de la proteína fue el único epítipo en no mostrar especificidad por la región CDR del HLA DRB1.

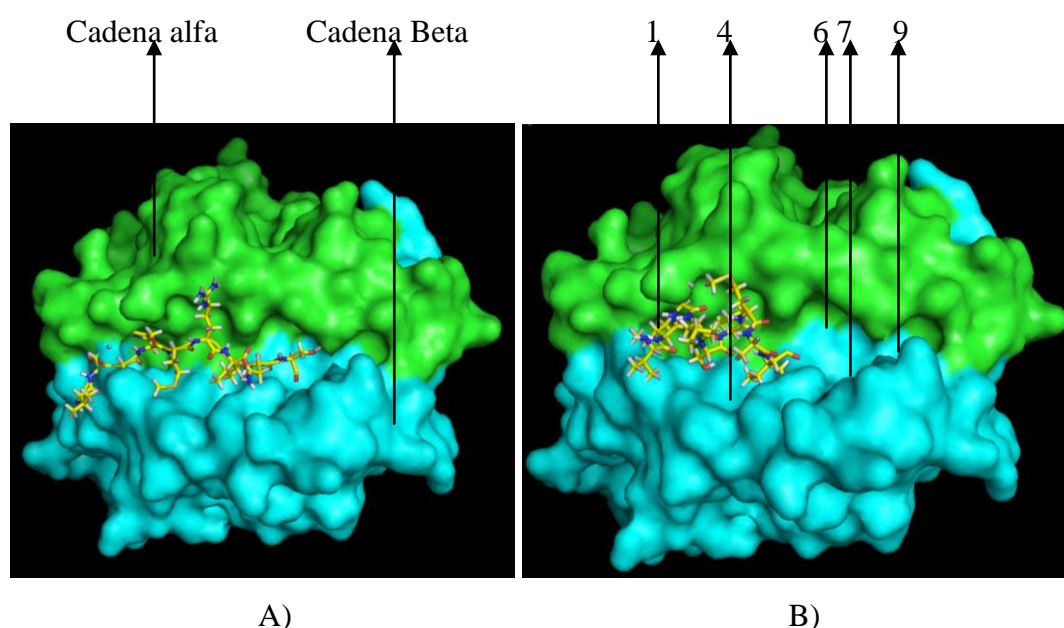
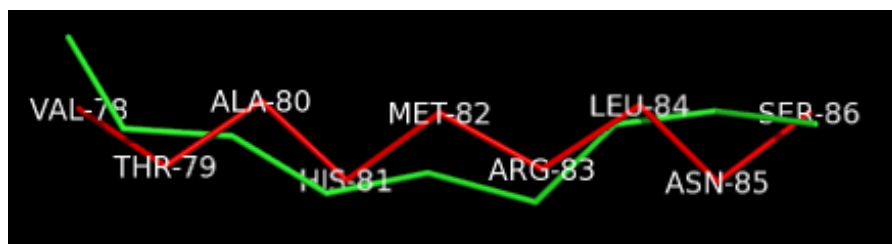


Figura 10. Docking molecular entre epítopes T y el CMHII. A) Interacción entre el epítipo $^{78}\text{VTAHMRLNS}^{86}$ y el CMHII. Se observa que el epítipo reconoce el CDR pero no se ubica dentro de los bolsillos. Cadena alfa vista en verde; cadena beta vista en azul. B) Interacción entre el epítipo $^{18}\text{VITLSVMAI}^{26}$ más promiscuo y el CMHII. El epítipo reconoce el CDR pero la estructura secundaria del péptido es enrollada. Con flechas se representan los bolsillos relevantes del CMH II.

Los refinamientos energéticos de los epítopes mostraron cambios diversos en los péptidos analizados. Al realizar 10 minimizaciones de energía al epítipo que mejor acople tuvo a los bolsillos del CMHII (VTAHMRLNS) la energía total del péptido paso de 38.14 Kcal/mol a -37.32 Kcal/mol, mostrando una molécula de mayor estabilidad. Por otro lado, para el epítipo más promiscuo (VITLSVMAI) solo fue posible realizar 4 minimizaciones de energía, pasando de 141 Kcal/mol a 127.78 Kcal/mol de energía total, lo que muestra una reducción mínima en la estabilidad del péptido (Figura 11).

A)



B)

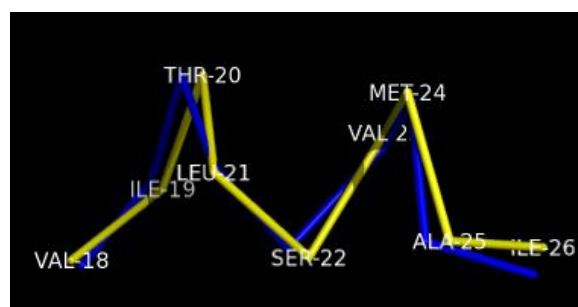


Figura 11. Solapamiento de los epítopes antes y después de la minimización de energía. A) Epítipo $^{78}\text{VTAHMRLNS}^{86}$. Estructura verde: antes del refinamiento; estructura roja: después del refinamiento. B) Epítipo $^{18}\text{VITLSVMAI}^{26}$. Estructura azul: antes del refinamiento; estructura amarilla: después del refinamiento. De izquierda a derecha extremo N-terminal a C-terminal.

El docking molecular entre el epítipo $^{78}\text{VTAHMRLNS}^{86}$ refinado y la molécula de CMH II mostró un mejor acople del péptido en los bolsillos del complejo, mientras que la interacción entre el epítipo $^{18}\text{VITLSVMAI}^{26}$ siguió desacoplada a los bolsillos del CMHII (Figura 12).

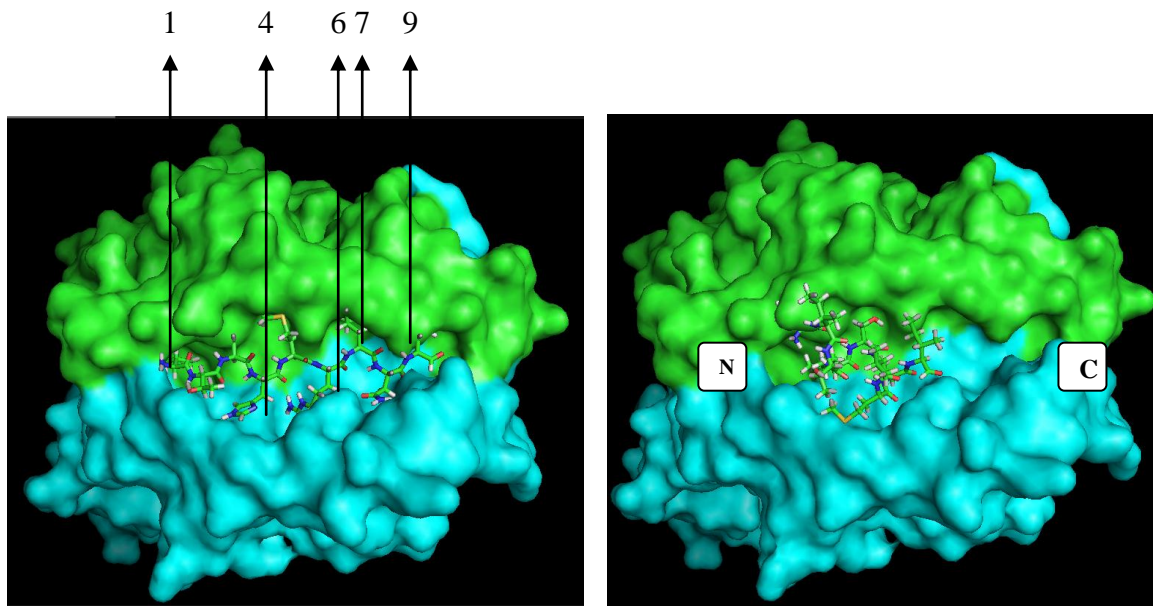


Figura 12. Docking molecular entre epítopes T y el CMHII después de la minimización energética. A) Interacción entre el epítipo $^{78}\text{VTAHMRLNS}^{86}$ y el CMHII. Se observa que el epítipo reconoce el CDR y se ubica dentro de los pockets respectivos del MHCII. Las flechas indican los aminoácidos relevantes ubicados en los bolsillos. B) Interacción entre el epítipo $^{18}\text{VITLSVMAI}^{26}$ más promiscuo y el CMHII.

La medición y cálculo de los puentes de hidrogeno, que se realizó con el programa Pymol (opción medición de interacciones polares), solo se hizo entre la cadena principal del epítipo $^{78}\text{VTAHMRLNS}^{86}$ y el CMH II HLA DRB1, ya que el otro epítipo no mostró buen acople al HLA. Como se observa en la figura 13, solo se presentan tres enlaces de hidrógeno entre la cadena principal del epítipo (en los dos primeros residuos) y dos de los cinco residuos conservados en las moléculas de HLA (Serina 53 de la cadena alfa y Asparagina 82 de la cadena beta).

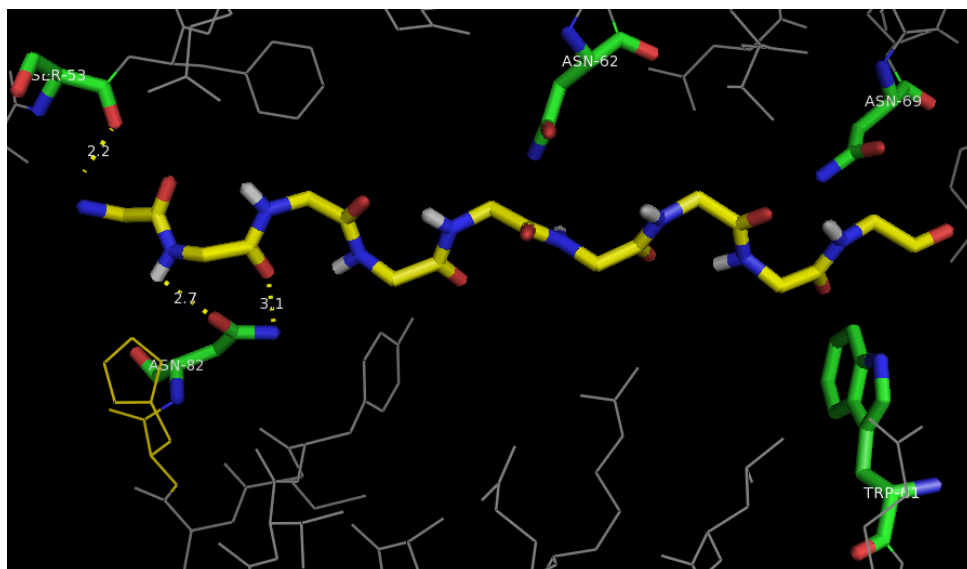
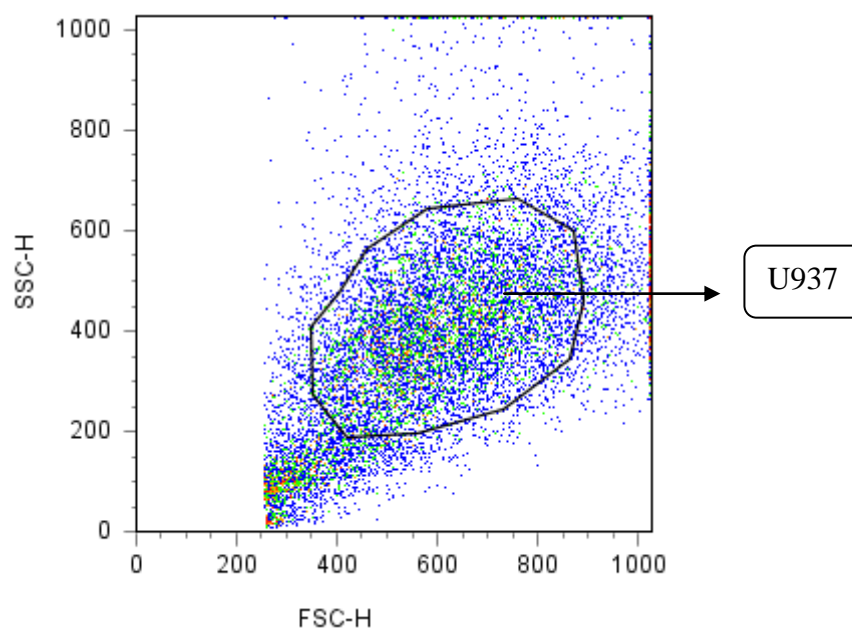


Figura 13. Puentes de hidrogeno entre la cadena principal del péptido $^{78}\text{VTAHMLNS}^{86}$ y el HLADRB1. Los puentes de hidrogeno son mostrados como líneas discontinuas. De los ocho enlaces de hidrogeno presentes entre péptidos y HLA DRB1 consultados en Stern *et al*, 1994 (para 9 aminoácidos del epítipo), solo se dieron tres en la estructura analizada. Los aminoácidos resaltados del HLA representan residuos conservados en la estructura que son relevantes en los acoples de péptidos a moléculas de HLA. En la parte superior se encuentran 3 residuos conservados de la cadena alfa; en la parte inferior 2 residuos conservados de la cadena beta.

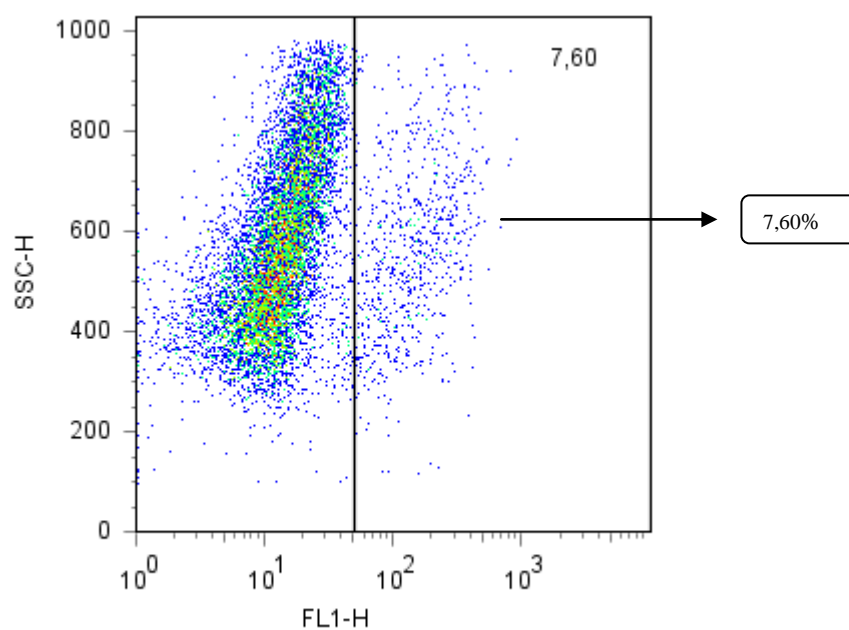
7.2 Inhibición de la invasión a macrófagos derivados de monocitos

La evaluación *in vitro* del efecto inhibitorio de los 26 péptidos que comprenden la proteína Rv3494c (tabla 4) sobre la invasión de *M. tuberculosis* en la línea celular de macrófagos pulmonares U937 tuvo como control positivo de invasión células con bacteria únicamente, este sería el 100% de invasión base para calcular los porcentajes de inhibición debidos a la presencia de péptidos en los tratamientos posteriores. Como control negativo de invasión se utilizó citocalasina D 10 μM , la cual mostró un efecto inhibitorio entre el 60% y el 70% (Figura 14).

A)



B)



C)

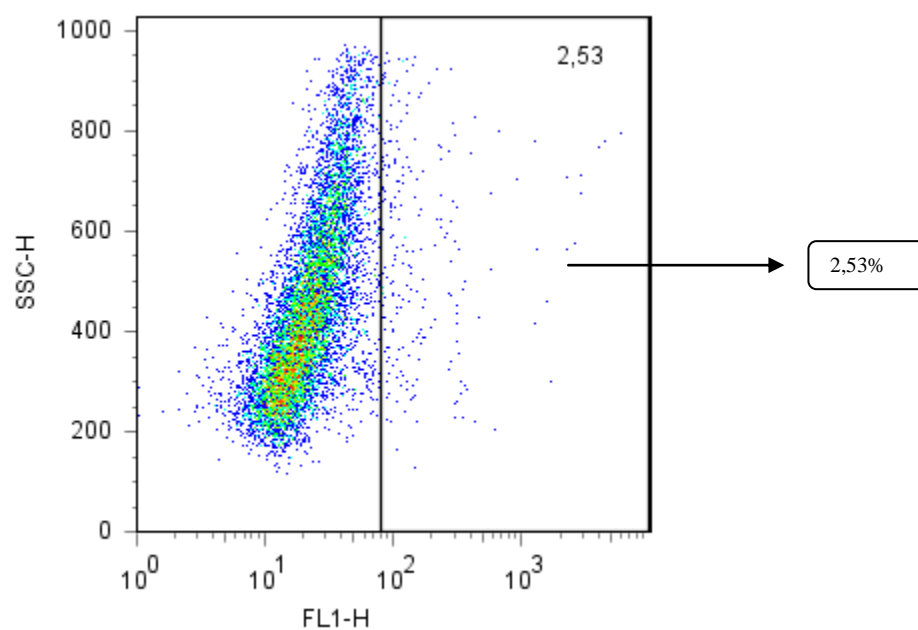


Figura 14. Inhibición de la invasión. A) Población de células U937 seleccionadas como plot de análisis en ensayos de invasión. B) Control positivo de invasión. C) Control negativo de invasión con Citocalasina D 10 μ M.

La inhibición de la entrada de *M. tuberculosis* H37Rv a células U937 fue dependiente de la concentración para los péptidos 38371, 38378 y 38384 los cuales mostraron inhibiciones de 5%, 32% y 42% para la concentración de 2 μ M respectivamente; 17%, 48% y 50% para la concentración de 20 μ M la inhibición fue de y para la concentración de 200 μ M la inhibición

fue de 20%, 58% y 55% respectivamente. De igual forma, los péptidos 38391, 38394 y 38397 mostraron inhibición de la invasión dependiente de concentración, pero sin cambios significativos entre algunas de sus concentraciones. Así, el péptido 38391 inhibió 49%, 59,4% y 59,3% a las concentraciones de 2 μM , 20 μM y 200 μM respectivamente mostrando similar efecto entre las dos concentraciones más altas. El péptido 38394 tuvo un comportamiento de concentración muy similar al péptido 38391 mostrando inhibiciones de 57%, 65% y 64% para las concentraciones de 2 μM , 20 μM y 200 μM con una diferencia de 1% de inhibición entre las dos últimas concentraciones. Los porcentajes de inhibición del péptido 38397 fueron 48%, 47% y 58% para las tres concentraciones mostrando una diferencia de 1% entre las concentraciones de 2 μM y 200 μM . El péptido 38385 mostró alta inhibición y dependencia en las concentraciones de 2 μM y 200 μM (la concentración 2 μM no presentó inhibición) con porcentajes de 35% y 72% respectivamente. El péptido 38381 mostró inhibición solo en las concentraciones de 20 μM y 200 μM con porcentajes de 9% y 28%. Por último los péptidos 38376, 38379, 38390, 38392 y 38393 mostraron efectos inversamente proporcionales a la concentración de péptido (Figura 15).

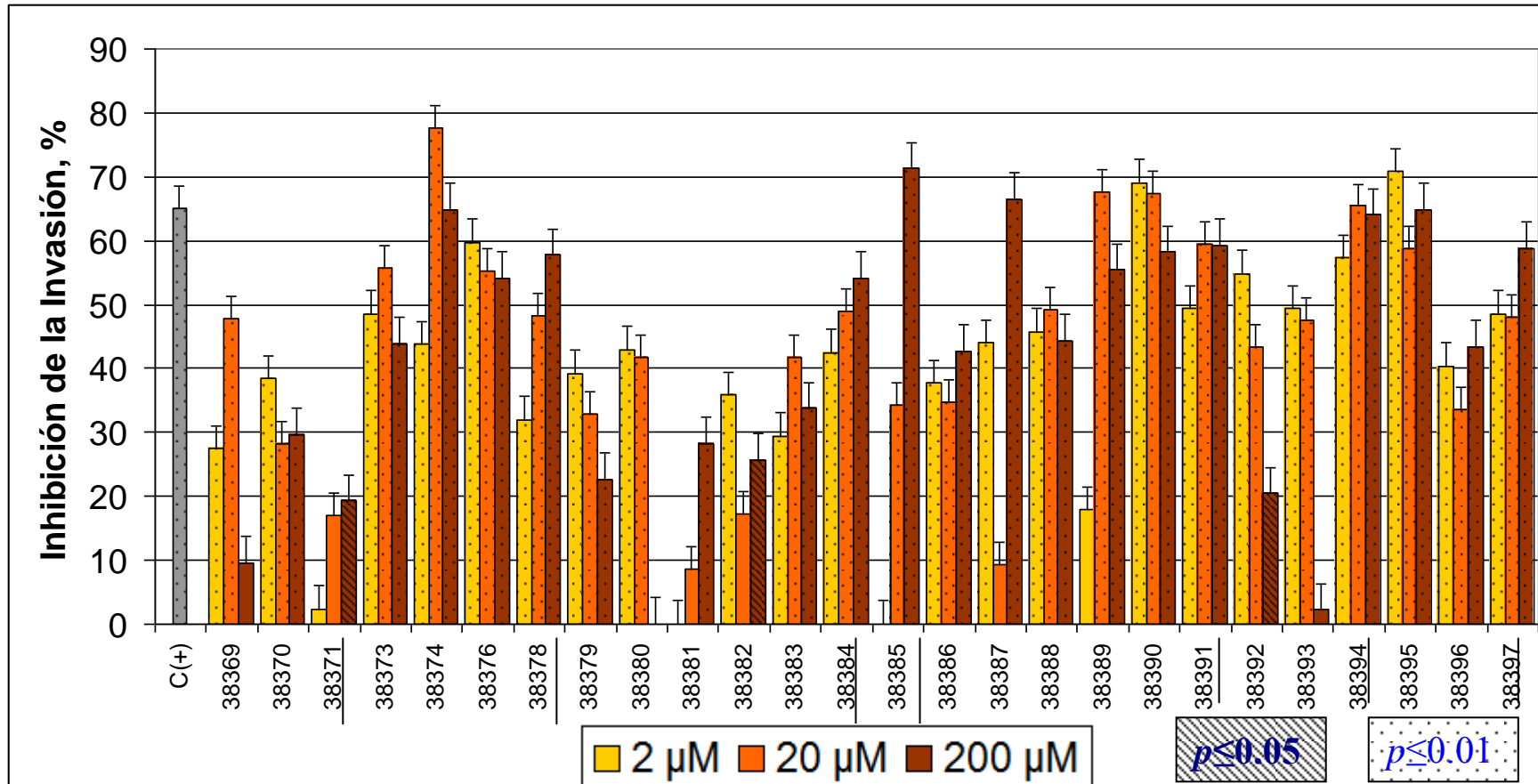


Figura 15. Porcentajes de inhibición de la invasión de *M. tuberculosis* H37Rv a células U937. Cada péptido de la proteína se probó a tres concentraciones diferentes (2 μM , 20 μM , 200 μM) teniendo como control positivo de inhibición Citocalasina D 10 μM (barra gris). El error estándar de cada tratamiento es indicado por las barras negras. Las barras sobre el número del péptido indican dependencia de concentración. Columnas punteadas $p \leq 0.01$. Columnas con rayas diagonales $p \leq 0.05$.

Tabla 4. Listado y secuencia de los 26 péptidos utilizados en los ensayos de invasión.

PEPTIDO	SECUENCIA
38369	MIDRLAKIQLSIFA VITVITY
38370	LSVMAIFYLRLPATFGIGTY
38371	GVSAD FVAGGGLYKNANVTY
38373	HMRLNSGTAIPSNVTATVRSY
38374	VSAIGE QYIDL VPPENPSST
38376	DLLRQAETLLGSLGDTRLREY
38378	ESARLLVDEANANYPQVSQL
38379	IDQAGPFLQAQIRAGGDIKSY
38380	LADGLARFTWQLRAADPRLRY
38381	DTLADAPDAIDEANTAFSGIY
38382	RPSFPALAASLANLGRVGVIIY
38383	YHKSIEQLLVVFPALFAAII
38384	TSAGGVPQDEGAKLDFKIDLY
38385	HDPPPTMTGFLPPPLVRSPAY
38386	DESVREIPRDMYTKTAQNDP
38387	STVRGARNYPTQEFPGKRAP
38388	TVQLTRDPRGYVPVGTNPWR
38389	GPPIPYGTEVTDGRNLPN
38390	KFPYIPPGADPDGPVPIVGP
38391	PPPGQVAGPGPAPHQPAQPAY
38392	PPPNDNGPPPFTSWMPPGY
38393	PPEPPQVPYPATIPPPPPPE
38394	GTGPPPGPAPGPQPQASGPAY
38395	YTIYDQLSGAFADPAGGTGI
38396	FAPGMTGASSAENWVDLMRDY
38397	MTGASSAENWVDLMRDPRQLY

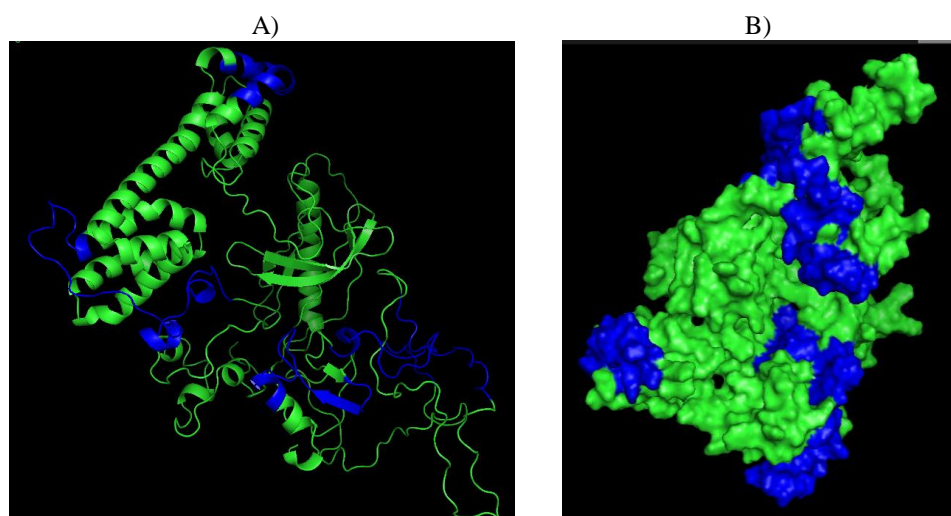


Figura 16. Visualización de péptidos de mayor inhibición. A) Estructura tridimensional (Cartoon) de péptidos de mayor inhibición; péptidos (azul); resto de la proteína (verde). B) Los péptidos se ven expuestos en la superficie de la proteína. Péptidos (azul); resto de la proteína (verde).

La figura 16A muestra la estructura secundaria de los 6 péptidos con mayor actividad inhibitoria, los cuales no tienen una estructura topológica definida, dado a que están ubicados después del aminoácido 337 (hasta donde se modeló la proteína). A pesar de ello los péptidos se ubican en la periferia de la proteína como lo indica la figura 16B.

7.3 Citotoxicidad de los péptidos en células U937

Para cada uno de los pozos (tratamiento) se obtuvo la fluorescencia emitida por las células (anexo 11.3). Así pues, las fluorescencias emitidas por las células tratadas con cada uno de los péptidos no tuvieron diferencias significativas con respecto al control positivo de bio-reducción de Resazurina. De igual forma la fluorescencia emitida por el buffer HBSS y por el medio RPMI completo fue similar entre sí y muy inferior a la emitida por las células con y sin péptido.

8. Discusión y conclusiones

De acuerdo con la metodología planteada por la Fundación Instituto de Inmunología de Colombia (FIDIC), a partir del estudio de proteínas localizadas en la envoltura celular de la micobacteria que presentan roles importantes en procesos invasivos y de virulencia se podría llegar a generar candidatos para el diseño de vacunas sintéticas contra la tuberculosis. Así, el estudio *in silico* e *in vitro* de la proteína Rv3494c (MCE4F) se encuentra en el marco de la búsqueda de péptidos candidatos a subunidades de vacuna con capacidad de inhibición de la entrada de la micobacteria a los macrófagos pulmonares.

La determinación de la posible localización de la proteína en la envoltura celular de la micobacteria y el potencial antigénico de estas a través de herramientas bioinformáticas permite guiar las validaciones experimentales de dichas características. La proteína Rv3494c (MCE4F) mostró que tiene una posible localización subcelular en la envoltura de la bacteria. Sin embargo, la zona específica no fue clarificada, ya que el servidor Gpos-Ploc determinó una localización extracelular, mientras que el servidor Psortb y TB pred la localizaron en la membrana citoplasmática. La base de datos Tuberculist (<http://tuberculist.epfl.ch/quicksearch.php?gene+name=Rv3494c>) muestra que una posible función de esta proteína sea la de procesos de invasión en células hospederas y estudios

bioinformáticos recientes han mostrado que todas las proteínas pertenecientes a la familia MCE, incluida MCE4F, pueden ser consideradas como proteínas de membrana externa (4) y están involucradas en procesos de invasión celular (83).

La presencia de un péptido señal tipo I, con dos posibles sitios de clivaje, y un tipo de secreción por vía clásica fue evidenciado en este estudio a través de los servidores SignalP 3.0 y SignalP 4.0. Estudios anteriores demostraron bioinformáticamente esta misma característica (4). A pesar que el servidor SecretomeP 2.0 haya mostrado una secreción por vía no clásica, hay que tener en cuenta que las predicciones que se hacen a través de esta herramienta no se basan en la búsqueda de secuencias con características y motivos específicos, sino simplemente en algún tipo de secreción dada por mecanismos no específicos, los cuales podrían contener en ocasiones tipos de secreción ya conocidos (56).

La presencia de una hélice transmembranal fue demostrada con los servidores TMHMM 2.0 y Phobius entre los aminoácidos 12 al 34, sobrelapándose con la secuencia predicha como péptido señal. Sin embargo, los servidores Phobius y SignalP 4.0 permiten descartar que la secuencia ¹² IFAVITVITLSVMAIFYLRLPAT ³⁴ fuese una hélice transmembranal confirmándose como una secuencia señal. Las características predichas de estructura secundaria coinciden con el modelo tridimensional de la proteína, ya que la proteína tiene estructura secundaria hasta el aminoácido 337, el cual fue el último residuo en ser modelado por Robetta. Esto muestra el acople existente entre estructura secundaria y terciaria.

Dado a que el complejo mayor de histocompatibilidad II presenta muchos polimorfismos, la predicción de epítopes T se hizo con base a los alelos de HLA más comunes para tuberculosis en las zonas de mayor incidencia de esta enfermedad, correspondientes a 12 alelos de la molécula DRB1 (1; 43; 82). A partir de la predicción de epítopes T usando el grupo de cinco servidores, se obtuvo un total de 38 epítopes promiscuos para los diferentes alelos de HLA analizados. De esta manera se obtuvieron cinco zonas dentro de la secuencia con probabilidad de ser antigénicas. A nivel de secuencia, los epítopes promiscuos presentaron por lo menos uno de los cuatro aminoácidos específicos (en las posiciones 1, 4, 6 y/o 9) para acoplarse de manera adecuada a los bolsillos o “pockets” 1, 4, 6 y 9 de CMH DRB1 (44; 45).

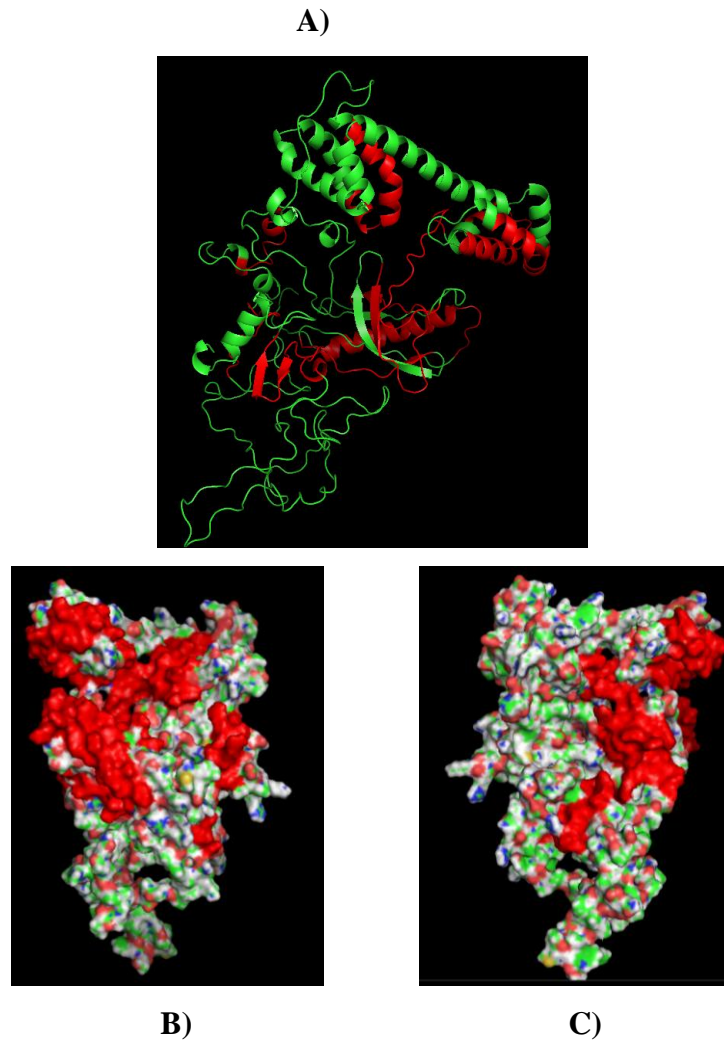


Figura 17. Visualización de epítopes T promiscuos predichos. A) Estructura secundaria de epítopes T; epítopes (rojo); resto de la proteína (verde). B) Los epítopes se ven expuestos en la superficie de la proteína (dos vistas diferentes de la proteína). Epítopes T (rojo); resto de la proteína (multicolor).

La figura 17A muestra la estructura secundaria de los epítopes T, los cuales muestran topología de 3 hélices alfa y 2 hojas plegadas. Además la figura 17B muestra que los epítopes están expuestos sobre la superficie de la proteína.

La predicción de epítopes B, se hizo como complemento para evidenciar regiones compartidas con epítopes T, que probablemente sirvan para el desarrollo de inmunoterapias y para la producción de anticuerpos, fue realizada con 2 herramientas de predicción: El servidor Bepipred 1.0 esta basado en métodos de escala de propensión (del ingles: *Propensity Scale Methods*), el cual asigna valores a cada aminoácido de la secuencia teniendo en cuenta propiedades físico químicas como antigenicidad, hidrofiliidad y accesibilidad al solvente.

Además incorpora modelos ocultos de Markov (HMM) para generar de esta manera métodos de escala propensión mejorados (del inglés: *Improved Propensity Scale Methods*) (67).

El programa Antheprot también se basa en métodos de escala de propensión, generando valores de antigenicidad de Parker, hidrofiliidad, accesibilidad al solvente e hidrofobicidad para cada aminoácido. El uso de estos métodos se aplicó a esta predicción, ya que proveen bases teóricas para el estudio posterior de epítopes conformacionales más complejos (47). En total se detectaron cuatro posibles epítopes B, de los cuales tres poseen secuencias que se solapan con algunas de las secuencias de las zonas predichas a comportarse como epítopes T. Dichas zonas se encuentran entre el aminoácido 121 al 137, entre el 360 al 370 y entre el 530 al 540.

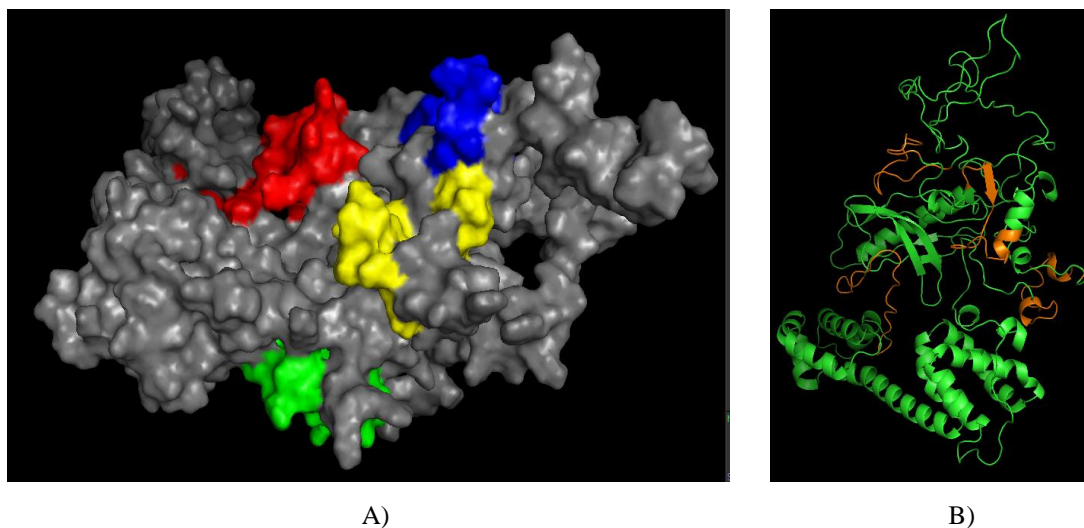


Figura 18. Visualización de epítopes B predichos con Bepipred 1.0 y Antheprot 3D. A) Los epítopes se ven expuestos en la superficie de la proteína. Epítope 115 ENPSSTKLRNGFRIQRQNTR 135 (rojo); epítope 355 TAQNDPSTVRGARNYPCQEF 375 (verde); epítope 494 PPPPPEGTGPPPGPAPGPQP 514 (azul); epítope 533 DPAGGTGIFAPGMTGASSAE 553 (amarillo); resto de la proteína (gris). B) Estructura secundaria de epítopes B; epítopes (naranja); resto de la proteína (verde).

Al corroborar la ubicación de los epítopes B mediante la visualización de la estructura tridimensional de la proteína, se puede ver que estos epítopes evidentemente se ubican en zonas superficiales de la proteína (figura 18).

La predicción del modelo de la proteína con otros servidores como ITasser, Esypred3D y Swiss model no se pudo realizar debido a la carencia de moldes o templates para generar el modelo, mientras que el servidor Phyre2 solo modeló el 29% de la proteína, lo que incidió en

resultados de validación menos significativos que el modelo determinado con Robetta. El servidor de modelamiento por homología y *ab initio* Robbeta permitió modelar la proteína hasta el aminoácido 337, correspondiente a más de la mitad de la secuencia total. El método utilizado por este servidor reconoce secuencias homologas filogenéticamente no emparentadas para construir el modelo; de igual forma el método Rosseta *de novo* permite modelar zonas de la secuencia que no tengan homología con estructuras cristalográficas de la base de datos PDB, llevando a la generación de un estructura con mayor número de aminoácidos modelados (Robetta). La mayoría de los aminoácidos presentan una estructura secundaria alfa helical, las cuales se encuentran muy enrolladas. Las zonas con hojas plegadas presentan una distribución de aminoácidos mas abierta que las hélices. Una característica evidenciada en los epítopes T y B predichos es que se ubican en la periferia del modelo tridimensional, validando las predicciones realizadas mediante los servidores de predicción de epítopes T y B. A pesar que la validación del modelo usando el servidor SWISS MODEL mostró que la conformación de la estructura proteica no es la más idónea con respecto a otras estructuras cristalográficas que se encuentran en PDB (Protein Data Bank), el Plot de Ramachandran indica que los aminoácidos se encuentran en zonas favorables para su conformación estructural idónea; la mayoría de aminoácidos tienen una conformación de hélice alfa derecha y las distribuciones para la glicina y prolina son favorables. Esto indica que la estructura terciaria y secundaria predichas para la proteína son favorables.

El docking molecular entre los quince epítopes promiscuos predichos y la molécula HLA DRB1 (0101) mostró un reconocimiento del sitio de unión de péptidos (llamado CDR en el servidor PATCHDOCK) al CMHII por parte de catorce de los quince epítopes. Sin embargo, el acople de dichos péptidos a los bolsillos de este sitio de unión no fue bien establecido, lo que condujo a un refinamiento y minimización de la energía de dos de los epítopes predichos. La baja especificidad de uno de los epítopes promiscuos predichos al HLA muestra la fiabilidad de las otras predicciones. El servidor Patchdock permite hacer un Docking a través de la generación de “parches” a lo largo de la molécula receptora para detectar posibles sitios de unión en la molécula y reducir las posibles soluciones erróneas de interacción; además reconoce tipos de residuos (polares y no polares) y genera las transformaciones de interacción de manera rápida (75), conduciendo a una mejor interacción molecular entre dos estructuras. A pesar que el servidor PATCHDOCK sugiere hacer la interacción con un RMSD de 1.5, en el presente estudio se hizo con la opción predeterminada (RMSD 4.0); esto no interfiere en el acople de los péptidos con el CMHII, ya que el péptido no cambiara su conformación

tridimensional original (la que viene del modelo completo de la proteína predicho con el servidor Robetta) el unirse al HLA. El servidor PATHCDOCK no cambia conformaciones tridimensionales de las estructuras analizadas; simplemente busca los mejores sitios de acople entre estas utilizando la geometría del área de las moléculas.

El refinamiento y minimización de energía se hizo utilizando un algoritmo muy robusto en el que a través de interacciones generadas en un gradiente descendiente XY se puede conseguir el mínimo de energía con alta confiabilidad (77). A partir de lo anterior, se evidenció que solo el epítipo refinado VTAHMRLNS (posición 78) mostró significancia en cuanto a su disminución de energía y acople a los bolsillos del HLA. El otro epítipo VITLSVMAI (posición 18) no tuvo cambios significativos en su interacción con el HLA. El péptido de mejor acople al HLA presenta estructura secundaria de hoja plegada, mientras que el de menor acople es una hélice alfa derecha con tendencia a enrollarse sobre si misma. Esta característica explicaría la baja especificidad del péptido VITLSVMAI (posición 18) a los bolsillos del HLA.

La estructura cristalográfica del complejo HLA DRB1-Péptido de virus de influenza (45) mostró que en los enlaces de puentes de hidrogeno entre el péptido y el HLA intervenían péptidos conservados en las moléculas de CMH tipo II de humanos y ratones. Estos aminoácidos corresponden a PHE-51, SER-53, ASN-62, ASN-69 y ARG-76 de la cadena alfa, y a TRP-61, HIS-81 y ASN-82 de la cadena beta. En total se forman 12 puentes de hidrogeno. Para un péptido de nueve aminoácidos (como el de este estudio), los enlaces de puente de hidrogeno se dan solo en 5 de estos aminoácidos (alfa SER-53, ASN-62, ASN-69 y beta TRP-61 y ASN-82) formando ocho puentes de hidrogeno. De estos, solo tres enlaces se dieron entre el epítipo VTAHMRLNS (posición 78), entre los dos primeros residuos del péptido y dos de los cinco residuos conservados en las moléculas de HLA (Serina 53 de la cadena alfa y Asparagina 82 de la cadena beta). El remplazo de los aminoácidos de las posiciones críticas 1, 4, 6 y 9 del péptido podría mejorar la interacción del complejo (43), dependiendo de la molécula de HLA y los requerimientos de residuos que tenga cada bolsillo o "pocket" del CMHII (44, 45).

En cuanto a la capacidad que tienen los péptidos de la proteína para inhibir la invasión de la micobacteria a los macrófagos pulmonares, los porcentajes de inhibición son comparables al porcentaje del control de citocalasina D 10 uM. Los péptidos que presentaron mayor

inhibición dependiente de concentración fueron el 38378 y 38384 con porcentajes entre el 30% y el 60%. También es importante resaltar que los péptidos 38391, 38394 y 38397, que tuvieron efectos muy similares entre algunas de sus concentraciones, mostraron altas inhibiciones por encima del 50 % en sus tres concentraciones. Un caso especial fue el del péptido 38385 que mostró alta inhibición solo en las concentraciones de 20 uM (35%) y 200 uM (72 %). De esta manera se evidencio que 6 de los 26 péptidos presentaron inhibiciones altamente significativas, sugiriendo que probablemente haya una unión específica a receptores en la superficie de las células blanco, lo que evita la entrada de la micobacteria a los macrófagos. Debido a que el modelo tridimensional solo fue predicho hasta el aminoácido 337 y que los 6 péptidos de mayor inhibición de la invasión se ubican después de este residuo, no es posible identificar una estructura secundaria similar entre estos péptidos. Sin embargo, todos estos péptidos se ubican en la periferia de la proteína.

El ensayo de reducción de Resazurina mostró que ninguno de los péptidos utilizados en los ensayos de invasión en su concentración máxima es citotóxico para las células U937, debido a que la fluorescencia emitida por las células tratadas con péptido es igual o similar a las de las células sin péptido (control positivo de bio-reducción del colorante). Con esto se puede afirmar que el efecto inhibitorio que los péptidos mostraron en los ensayos de invasión es confiable y verdadero. De igual forma se evidencio que la fluorescencia emitida por el medio RPMI completo (Control de medio RPMI completo) y por el buffer HBSS (control de HBSS) es muy similar entre si y muy inferior a la fluorescencia emitida por los tratamientos y el control positivo de reducción de colorante, por lo cual, la alteración que puedan generar estas soluciones sobre los péptidos unidos a las células U937 es mínimo o nulo.

A partir de los resultados se pueden evidenciar zonas en la secuencia que probablemente tengan probabilidad de ser epítopes T, epítopes B y/o inhibir la entrada de la micobacteria a las células. El péptido 38378, cuya secuencia empieza en el aminoácido 181, se solapa con la zona de epítopes T 180-190 (aminoácidos). Los péptidos 38384 y 38385, con inicio de secuencia en los aminoácidos 301 y 321 respectivamente, se solapan con el epítope B ubicado en los aminoácidos 355-375 y se ubica entre dos zonas de epítope T. El péptido 38391 no se solapa con nada al igual que el 38394 que se haya junto a la zona de epítopes T en el aminoácido 520 y por último, el péptidos 38397 se solapa con el epítope B ubicado en el aminoácido 533. Esto muestra que los péptidos predichos como epítopes T, epítopes B y buenos inhidores de la invasión de la micobacteria en los macrófagos se solapan o se hayan en

zonas continuas de la secuencia de la proteína que podrían ser utilizadas como base para futuros estudios.

9. Recomendaciones

1. Probar los epítopes T promiscuos en diferentes alelos de HLA DRB1, con el fin de comparar la interacción entre estos con los resultados obtenidos en el presente estudio.
2. Realizar cambios de residuos en los diferentes epítopes T a miras de mejorar el acople que estos presenten en los bolsillos del HLA.
3. Generar minimizaciones de energía a cada uno de los epítopes T predichos, con el fin de mejorar la interacción de estos con el HLA.
4. Evaluar la interacción que tengan los mejores complejos de epítopes-HLA con el receptor de linfocitos T (TCR).
5. Evaluar el efecto inhibitorio que puedan tener los péptidos de la proteína sobre la invasión de la micobacteria en células A549 epiteliales.
6. Realizar validaciones experimentales de los resultados obtenidos a través de herramientas bioinformáticas.

10. Bibliografía

1. World Health Organization (WHO). *Global Tuberculosis Control*. World Health Organization. Switzerland. 2010, 205 p.
2. Kaufmann S, Baumann S, Nasser E. Exploiting immunology and molecular genetics for rational vaccine design against tuberculosis. *Int J Tuberc Lung Dis* 2006; 10: 1068-1079.
3. Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon S, Eiglmeier K *et al*. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393: 537-544
4. Song H, Sandie R, Wang Y, Andrade-Navarro M, Niederweis M. Identification of outer membrane proteins of *Mycobacterium tuberculosis*. *Tuberculosis* 2008; 88: 526-544.
5. Ocampo M, Aristizábal-Ramírez D, Rodríguez DM, Muñoz M, Curtidor H, Vanegas M, Patarroyo MA, Patarroyo ME. The role of *Mycobacterium tuberculosis* Rv3166c protein-derived high-activity binding peptides in inhibiting invasion of human cell lines. *Protein Eng Des Sel*. 2012; 25(5):235-42.
6. Rodriguez D, Vizcaino C, Ocampo M, Curtidor H, Pinto M, Patarroyo M, Patarroyo M. Peptides from the *Mycobacterium tuberculosis* Rv1980c protein involved in human cell infection: insights into new synthetic subunit vaccine candidates. *Biol. Chem*. 2010; 391: 207-217.
7. Cifuentes D, Ocampo M, Curtidor H, Vanegas M, Forero M, Patarroyo M, Patarroyo M. *Mycobacterium tuberculosis* Rv0679c protein sequences involved in host-cell infection: Potential TB vaccine candidate antigen. *BMC Microbiology* 2010; 10:109.
8. Caceres S, Ocampo M, Arevalo-Pinzon G, Jimenez R, Patarroyo M, Patarroyo M. The *Mycobacterium tuberculosis* membrane protein Rv0180c: Evaluation of peptide sequences implicated in mycobacterial invasion of two human cell lines. *Peptides* 2010.

9. Patarroyo M A, Plaza D, Ocampo M, Curtidor H, Forero M, Rodríguez L, Patarroyo M E. Functional characterization of *Mycobacterium tuberculosis* Rv2969c membrane protein. *Biochemical and Biophysical Research Communications* 2008; 372: 935–940.
10. Chapeton-Montes J, Plaza D, Curtidor H, Forero M, Vanegas M, Patarroyo M E, Patarroyo M A. Characterizing the *Mycobacterium tuberculosis* Rv2707 protein and determining its sequences which specifically bind to two human cell lines. *Protein Science* 2008; 17: 342-351.
11. Plaza D, Curtidor H, Patarroyo M A, Chapetón-Montes, Reyes C, Barreto J, Patarroyo M E. The *Mycobacterium tuberculosis* membrane protein Rv2560 - biochemical and functional studies. *Febs Journal* 2007.
12. Patarroyo M A, Curtidor H, Plaza D, Ocampo M, Reyes C, Saboya O, Barrera G, Patarroyo M E. Peptides derived from the *Mycobacterium tuberculosis* Rv1490 surface protein implicated in inhibition of epithelial cell entry: Potential vaccine candidates?. *Vaccine* 2008; 26: 4387-4385.
13. Vera-Bravo R, Torres E, Valbuena J, Ocampo M, Rodríguez L, Puentes A, Garcia J *et al.* Characterising *Mycobacterium tuberculosis* Rv1510c protein and determining its sequences that specifically bind to two target cell lines. *Biochemical and Biophysical Research Communications* 2005; 332: 771-781.
14. Forero M, Puentes A, Cortés J, Castillo F, Vera R, Rodríguez L, Valbuena J *et al.* Identifying putative *Mycobacterium tuberculosis* Rv2004c protein sequences that bind specifically to U937 macrophages and A549 epithelial cells. *Protein Science* 2005; 14: 2767-2780.
15. García J, Puentes A, Rodríguez L, Ocampo M, Curtidor H, Vera R, Lopez R *et al.* *Mycobacterium tuberculosis* Rv2536 protein implicated in specific binding to human cell lines. *Protein Science* 2005; 14: 2236-2245.
16. Cosma C, Sherman D, Ramakrishnan L. The secret lives of the pathogenic mycobacteria. *Annu. Rev. Microb.* 2003; 57: 641-676.
17. Ministerio De Protección Social (MPS). *Boletín Epidemiológico semanal*. Semana epidemiológica numero 10 de 2011 (6 al 12 de marzo de 2011). 16 p.

18. Snider D, Castro JR and K. The global threat of drug-resistant tuberculosis. *N Engl J Med* 1998; 338: 1689-1690.
19. Brandt L, Feinho J, Olsen A, Chilima B, Hirsch P, Appleberg R, Andersen P. Failure of the *Mycobacterium bovis* BCG Vaccine: Some Species of Environmental Mycobacteria Block Multiplication of BCG and Induction of Protective Immunity to Tuberculosis. *Infection and Immunity* 2002; 70(2): 672-678.
20. Sinha S, Kosalai K, Arora S, Namane A, Sharma P, Gaikwad AN, Brodin P, Cole ST. Immunogenic membrane-associated proteins of *Mycobacterium tuberculosis* revealed by proteomics. *Microbiology* 2005; 151: 2411-2419.
21. Gu S, Chen J, Dobos KM, Bradbury EM, Belisle JT, Chen X. Comprehensive Proteomic Profiling of the Membrane Constituents of a *Mycobacterium tuberculosis* Strain. *Molecular & Cellular Proteomics* 2003; 2: 1284-1296.
22. Checkley A, Willie D, Scriba T, Golubchik T, Hill A, Hanekom W, Mcshane H. Identification of Antigens Specific to Non-Tuberculous Mycobacteria: The Mce Family of Proteins as a Target of T Cell Immune Responses. *PLoS ONE* 2011; 6(10).
23. Casali N, Riley L. A phylogenomic analysis of the *Actinomycetales* *mce* operons. *Bmc Genomics* 2007; 8:60.
24. Restrepo-Montoya D, Vizcaino C, Niño L, Ocampo M, Patarroyo M, Patarroyo M. Validating subcellular localization prediction tools with mycobacterial proteins. *BMC Bioinformatics* 2009; 10:134.
25. Vizcaíno C, Restrepo-Montoya D, Rodríguez D, Niño LF, Ocampo M, *et al.* Computational Prediction and Experimental Assessment of Secreted/Surface Proteins from *Mycobacterium tuberculosis* H37Rv. *PLoS Comput Biol* 2010; 6:6.
26. Ruiz-Manzano J, Blanquer R, Calpe JL, Caminero JA, Caylá J, Domínguez JA, García JM, Vidal R. Diagnostico y tratamiento de la tuberculosis. *Arch Bronconeumol* 2008; 44 (10): 551-566.
27. Scholssberg D. *Tuberculosis and nontuberculous mycobacterial infections*. Fifth edition illustrated. McGraw-Hill Prof Med/Tech. Philadelphia. UE. 2006, 523 p.

28. Parish T, Stoker N. *Mycobacteria protocols*. Illustrated edition. Humana Press. London, UK. 1998, 472 p.
29. Olubisi T, Manormoney P, Willem S. Adhesion to and invasion of pulmonary epithelial cells by the F15/LAM4/KZN and Beijing strains of *Mycobacterium tuberculosis*. *Journal of medical Microbiology* 2010; 59: 528-533.
30. Niederweis M, Danilchanka O, Huff J, Hoffman C, Engelhardt H. Mycobacterial outer membranes: in search of proteins. *Trends in microbiology* 2009; 18(3): 109-116.
31. Astarie-Dequeker C, N'diaye EN, Le Cabec V, Rittig M, Prandi J, Maridonneau-Parini I. The Mannose Receptor Mediates Uptake of Pathogenic and Nonpathogenic Mycobacteria and Bypasses Bactericidal Responses in Human Macrophages. *Infection and Immunity* 1999; 67(2): 469-477.
32. Wolfe LM, Mahaffey SB, Kruh NA, Dobos KM. Proteomic Definition of the Cell Wall of *Mycobacterium tuberculosis*. *Journal of Proteome Research* 2010; 9: 5816-5826.
33. Gorocica P, Jimenez-Martínez M, Garfias Y, Sada I, Lascurain R. Componentes glicosilados de la envoltura de *M. tuberculosis* que intervienen en la patogénesis de la tuberculosis. *Rev Inst Nal Enf Resp Mex* 2005; 18(2): 142-153.
34. Villeneuve C, Gilleron M, Maridonneau-Parini I, Daffé M, Astarie-Dequeker C, Etienne G. Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *Journal of Lipid Research* 2005; 46: 475-483.
35. Hu C, Mayadas-Norton T, Tanaka K, Chan J, Salgame P. *Mycobacterium tuberculosis* Infection in Complement Receptor 3-Deficient Mice. *The Journal of Immunology* 2000; 165: 2596-2602.
36. Ernst J. Macrophage Receptors for *Mycobacterium tuberculosis*. *Infection and Immunity* 1998; 66(4): 1277-1281.
37. kang PB, Azad AK, Torelles JB, Kaufman TM, Beharka A, Tibesar E, DesJardin L, Schlesinger LS. The human macrophage mannose receptor directs *Mycobacterium tuberculosis* lipoarabinomannan-mediated phagosome biogenesis. *JEM* 2005; 202 (7): 987-999.
38. Areschoug T, Gordon Siamon. Scavenger receptors: role in innate immunity and microbial patogénesis. *Cellular Microbiology* 2009; 11 (8): 1160-1169.

39. Means TK, Wang S, Lien E, Yoshimura A, Golenbock DT, Fenton MJ. Human Toll-Like receptors Mediate Cellular Activation by *Mycobacterium tuberculosis*. *J Immunol* 1999; 163:3920-3927.
40. Radolf C, Raldolf B, Belisle J, Golenbock DT, Henry W, Noss EH, Pai RK, Sellati TJ, Justin D. Toll-Like Receptor 2-Dependent Inhibition of Macrophage Class II MHC Expression and Antigen Processing by 19-kDa Lipoprotein of *Mycobacterium tuberculosis*. *J Immunol* 2001; 167:910-918.
41. DiGiuseppe PA, Cox JS. Protein secretion systems in Mycobacteria. *Cellular Microbiology* 2007; **9** (6): 1376-1384.
42. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S. Non-classical protein secretion in bacteria. *BMC Microbiology* 2005; **5** (58): 1-13.
43. Mcnamara L, He Y, Yang Z. Using epitope predictions to evaluate efficacy and population coverage of the Mtb72f vaccine for tuberculosis. *BMC Immunology* 2008; 11:18.
44. Patarroyo ME, Bermudez A, Patarroyo MA. Structural and Immunological Principles Leading to Chemically Synthesized, Multiantigenic, Multistage, Minimal Subunit-Based Vaccine Development. *Chemical Reviews* 2011; A-AW.
45. Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, Wiley DC. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994; 368: 215-221.
46. De Oliveira DB, Harfouch-Hammoud E, Otto H, Papandrea NA, Stern LJ, Cohen H, Boehm BO, Bach J-M, Caillat-Zucman S, Walk T, Jung G, Eliopoulos E, Papadopoulos GK, Van Ender PM. Structural analysis of two HLA-DR-presented autoantigenic epitopes: crucial role of peripheral but not central peptide residues for T-cell receptor recognition. *Molecular Immunology* 2000; 37: 813-825.
47. Nemeč M: Web alert. Lymphocyte activation and effector functions. *Current opinion in immunology* 2001;13(3):265.
48. Manzalawy EL, Hinavar V. Recent advances in B-cell epitope prediction methods. *Immunome Research* 2010; 6: S2.
49. Sadowski PG, Groen AJ, Dupree P, Lilley KS. Sub-cellular localization of membrane proteins. *Proteomics* 2008; 8: 3991-4011.
50. Horton P, Park KJ, Obayashi T, Fujita N. *et al.*, WoLF PSORT, protein localization predictor. *Nucleic Acids Res* 2007; 35: 585-587.

51. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, Brinkman FSL. PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* 2005; 21 (5): 617-623.
52. Kuo-Chen C, Hong-Bin S. "Cell-PLoc: A package of web-servers for predicting subcellular localization of proteins in various organisms". *Nature Protocols* 2008; 3, 153-162.
53. Hong-Bin S, Kuo-Chen C. "Gpos-PLoc: an ensemble classifier for predicting subcellular localization of Gram-positive bacterial proteins". *Protein Engineering, Design and Selection* 2007; 20: 39-46.
54. Rashid M, Saha S, Raghava GPS. Support Vector Machine-based method for predicting subcellular localization of mycobacterial proteins using evolutionary information and motifs. *BMC Bioinformatics* 2007; 8: 337.
55. Bendtsen JD, Nielsen H, Von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol* 2004; 340: 783-795.
56. Petersen TN, Brunak S, Von Heijne G, Nielsen H. SignalP 4.0: Discriminating peptides from transmembrane regions. *Nature Methods* 2011; 8: 785-786.
57. Bendtsen JD, Nielsen H, Widdick D, Palmer T, Brunak S. Prediction of twin-arginine signal peptides. *BMC Bioinformatics* 2005; 6: 167.
58. Juncker AS, Willenbrock H, Von Heijne G, Nielsen H, Brunak S, Krogh A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* 2003; 12 (8): 1625-1662.
59. THMMH Server Version 2.0. Prediction of Transmembrane Helices in Proteins. <http://www.cbs.dtu.dk/services/TMHMM/>. Consultado en 2011.
60. Käll L, Kragh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction: the Phobius web server. *Nucleic Acid Res* 2007; 35: W429-W432.
61. Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B. Peptide Binding prediction for HLA DR, DP and DQ molecules. *BMC Bioinformatics* 2010; 11: 568.
62. Sturniolo T, Bono E, Ding J, Radrizzani L, Tuereci O, Sahin U, Braxenthaler M, Gallazzi F, Protti M, Sinigaglia F, Hammer J. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nature Biotechnology* 1999; 555-561.
63. Guan P, Doytchinova IA, Zygori C, Flower DR. MHCpred: A server for quantitative prediction of peptide-MHC binding. *Nucleic Acid Research* 2003; 31: 3621-3624.

64. Nielsen M, Lund O. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics* 2009; 10: 296.
65. Singh H, Raghava GPS. ProPred: Prediction of HLA-DR binding sites. *Bioinformatics* 2001; 17 (12): 1236-1237.
66. Nielsen M, Lundegaard C, Justesen S, Lund O, Buus S. NetMHCIIpan-2.0 – Improved pan specific HLA-DR predictions using a novel concurrent alignment weight optimization training procedure. *Immunome Res* 2010; 6 (1): 9.
67. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Research* 2006; 2 (2).
68. Deléage G, Combet C, Blanchet C, Geourjon C. ANTHEPROT: An integrated protein sequence analysis software with client/server capabilities. *Computers in Biology and Medicine* 2001; 31 (4): 259-267.
69. Mihășan M. Basic protein structure prediction for the biologist: a review. *Arch. Biol. Sci* 2010; 62 (4): 857-871.
70. Schwede T, Koop J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acid Research* 2003; 31 (13): 3381-3385.
71. Lambert C, Leonard N, De Bolle X, E Depiereux. ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* 2002 ; 18 (9): 1250-1256.
72. Kelley A, Sternberg JE. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 2009; 4(3): 363-371.
73. Zhang Y. I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins: Structure, Function, and Bioinformatics* 2009; 77 (Suppl. 9): 100-113.
74. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acid Research* 2004; 32: W526-W531.
75. Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. *WABI* 2002; 185-200.
76. Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric Docking. *Nucleic Acid Research* 2005; 33: W363-W367.
77. Biosym/MSI. *Discover Forcefield Simulations user guide, part1*. Version 2.9.7. BIOSYM/Molecular Simulations. San Diego, California, USA. 1995.
78. Departamento de salud y servicios humanos. *Bioseguridad en laboratorios de Microbiología y biomedicina*. Centro de Control y prevención de enfermedades. Cuarta edición.

79. Merrifield RB. "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide". *J. Am. Chem. Soc.* 1963; 85 (14): 2149–2154.
80. Chapeton-Montes J, Plaza D, Barrero C, Patarroyo M. Quantitative flow cytometric monitoring of invasion of epithelial cells by *Mycobacterium tuberculosis*. *Frontiers in Bioscience* 2008; 13: 650-656.
81. Parker JM ,Guo D, Hodges RD. New Hydrophilicity Scale Derived from High-Performance Liquid Chromatography Peptide Retention Data: Correlation of Predicted Surface Residues with Antigenicity and X-ray-Derived Accessible Sites. *Biochemistry* 1986; 25: 5425-5432.
82. Allele Frequencies in Worldwide Populations Database. <http://www.allelefrequencies.net/>. Consultado en 2011.
83. Pajon R, Yero D, Lage A, Llanes A, Borroto CJ. Computational identification of beta-barrel outer-membrane proteins in *Mycobacterium tuberculosis* predicted proteomes as putative vaccine candidates. *Tuberculosis* 2006; 86: 290-302.
84. *in vitro* toxicology assay kit Resazurin Based. <http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/Bulletin/tox8bul.Par.0001.File.tmp/tox8bul.pdf>. Consultado en 2012.
85. Microbiology lecture E-text part 1 Unit 2. http://student.ccbcmd.edu/~gkaiser/pdflg/LectureGuide_part1.pdf. Consultado en 2012. p 280.

11. Anexos

11.1 Matriz de todos los epitopes predichos para los 12 HLA DRB1 usando los 5 servidores.

ALELO	POSICION	SECUENCIA	ARBPYTHON	MHCPRED	NETMHCII	NETMHCIIIPAN	PRORED
101	28	YLRLPATFG	X		X	X	X
101	27	FYLRLPATF	X			X	
101	26	IFYLRLPAT	X			X	
101	296	FAAIITSAG	X				
101	263	SFPALAASL	X				
101	295	LFAAIITSA	X				
101	524	YDQLSGAFA	X	X	X		
101	289	LVVFPALFA	X				X
101	267	LAASLANLG	X				
101	287	QLLVVFPAL	X				
101	8	IQLSIFAVI	X				
101	9	QLSIFAVIT	X				
101	207	FLQAQIRAG	X				
101	12	IFAVITVIT	X				
101	81	HMLNSGTGA	X				
101	18	VITLSVMAI	X				X
101	23	VMAIFYLRL	X				
101	142	LLRQAETLL	X				
101	57	NVTYRGVAV	X				
101	177	ARLIESARL	X				
101	193	NYPQVSQLI	X				
101	176	LARLIESAR	X				
101	97	TVRSVSAIG	X				
101	228	FTWQLRAAD	X				
101	46	FVAGGGLYK	X				
101	146	AETLLGSLG	X				
101	6	AKIQLSIFA	X				
101	327	MTGFLPPPL	X				
101	82	MRLNSGTAI	X	X			X
101	260	IRPSFPALA	X				X
101	208	LQAQIRAGG	X				
101	218	IKSLADGLA	X				
101	523	IYDQLSGAF	X	X			
101	231	QLRAADPRL	X				
101	166	FIATNGAGP	X				
101	281	YHKSIEQLL	X				
101	60	YRGVAVGRV	X				X
101	141	DLLRQAETL	X				
101	264	FPALAASLA		X	X		X
101	528	LSGAFADPA		X			
101		YIDLVPEN		X			
101		LLVVFPALF		X			
101	291	FPALFAAII		X			X
101		VITVITLSV		X			
101		LIDQAGPFL		X			
101		MTGASSAEN		X			

101		IPSNVTATV		X		
101	125	FRIQRQNTR				X
101	12	FAVITVITL				X
101	362	VRGARNYPC				X
101	63	VGRVESVGL		X		X
301	179	LIESARLLV	X			
301	232	LRAADPRLR	X		X	
301	135	IGQDVADLL	X			
301	152	SLGDTRLRE	X			
301	185	LLVDEANAN	X			X
301	122	LRNGFRIQR	X			X
301	312	AKLDFKIDL	X			
301	383	LCRDPRGYV				X
301	28	LRLPATFGI				X
301	17	VITLSVMAI				X
301	334	LVRSPADES				X
301	259	IRPSFPALA				X
301	41	VSADFVAGG				X
301	142	LRQAETLLG				X
301	81	MRLNSGTAI				X
301	59	YRGVAVGRV				X
301	83	LNSGTAIPS				X
401	28	YLRLPATFG	X			
401	17	TVITLSVMA	X			
401	185	LLVDEANAN	X			
401	81	HMRLNSGTA	X			
401	524	YDQLSGAFA	X	X		
401	296	FAAIITSAG	X			X
401	26	IFYLRLPAT	X			
401	295	LFAAIITSA	X			
401	10	LSIFAVITV	X			
401	289	LVVFPALFA	X			
401	523	IYDQLSGAF	X	X		
401	52	LYKNANVTY	X			
401		VITVITLSV		X		
401		MTGASSAEN		X		
401		VGRVESVGL		X		
401		FPALAASLA		X		
401		MRLNSGTAI		X		
401		YIDLVPEN		X		
401		LSGAFADPA		X		
401		LIDQAGPFL		X		
401		LLVFPALF		X		
401		IPSNVTATV		X		
401		FPALFAAI		X		
401	278	VIYHKSIEQ				X
401	334	LVRSPADES				X
401	390	YVPVGTNPW				X
401	142	LRQAETLLG				X
401	125	FRIQRQNTR				X
401	405	YGTEVTDGR				X
401	17	VITLSVMAI				X

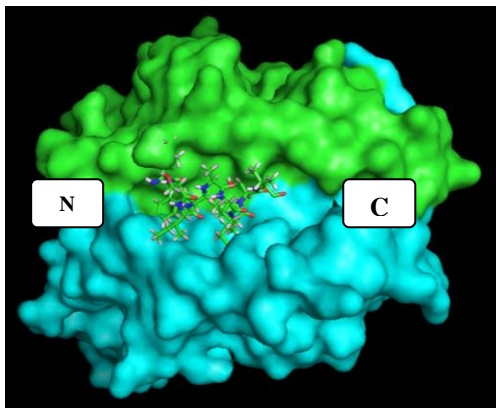
401	259	IRPSFPALA					X
401	72	LNPNGVTAH					X
401	185	LVDEANANY					X
401	362	VRGARNYPC					X
404	295	LFAAIITSA	X				
404	21	LSVMAIFYL	X				
404	10	LSIFAVITV	X				X
404	26	IFYLRLPAT	X				
404	304	LVVFPALFA					X
404	378	VRGARNYPC					X
404	406	YVPVGTNPW					X
404	44	LRLPATFGI					X
404	33	VITLSVMAI					X
404	158	LRQAETLLG					X
404	30	VITVITLSV					X
404	289	LGRVGVIIH					X
404	294	VIYHKSIEQ					X
404	126	LVPPENPSS					X
404	12	IFAVITVIT					X
404	143	IQRQNTRIG					X
404	275	IRPSFPALA					X
405	288	LLVVPALF	X				
405	27	FYLRLPATF	X				
405	53	YKNANVTYR	X				
405	281	YHKSIEQLL	X				
405	7	KIQLSIFAV	X				
405	21	LSVMAIFYL	X				
405	207	FLQAQIRAG	X				
405	390	YVPVGTNPW					X
405	295	FAAIITSAG					X
405	142	LRQAETLLG					X
405	28	YLRLPATFG					X
405	97	VRSVSAIGE					X
405	273	LGRVGVIIH					X
405	287	LVVFPALFA					X
405	28	LRLPATFGI					X
405	17	VITLSVMAI					X
405	12	FAVITVITL					X
405	362	VRGARNYPC					X
405	14	VITVITLSV					X
405	278	VIYHKSIEQ					X
405	59	YRGVAVGRV					X
405	127	IQRQNTRIG					X
701	179	LIESARLLV	X				
701	15	VITVITLSV	X	X			X
701	17	TVITLSVMA	X				
701	23	VMAIFYLRL	X				X
701	10	LSIFAVITV	X				X
701	288	LLVVPALF	X	X			
701	281	YHKSIEQLL	X		X		X
701	296	FAAIITSAG	X				
701		FPALFAAI		X			

701		FPALAASLA		X		
701		IPSNVTATV		X		
701		YIDLVPEN		X		
701		YDQLSGAFA		X		
701	64	VGRVESVGL		X		X
701		IYDQLSGAF		X		
701		LIDQAGPFL		X		
701		LSGAFADPA		X		
701	81	MRLNSGTAI		X		X
701		MTGASSAEN		X		
701	12	FAVITVITL				X
701	28	LRLPATFGI				X
701	390	YVPVGTNPW				X
701	59	YRGVAVGRV				X
701	7	IQLSIFAVI				X
701	18	VITLSVMAI				X
701	97	VRSVSAIGE				X
701	276	VGVIYHCSI				X
802	24	MAIFYLRLP	X			
802	127	RIQRQNTRI	X			
802	10	LSIFAVITV	X			
802	179	LIESARLLV	X			
802	21	LSVMAIFYL	X			
802	258	SGIRPSFPA	X			
802	25	IFYLRLPAT			X	
802	26	FYLRLPATF				X
802	59	YRGVAVGRV				X
802	57	VTYRGVAVG				X
802	520	YTIYDQLSG				X
802	288	LVVFPALFA				X
802	17	VITLSVMAI				X
802	390	YVPVGTNPW				X
802	229	WQLRAADPR				X
901	287	QLLVFPAL	X			
901	60	YRGVAVGRV	X		X	
901	262	PSFPALAAS	X			
901	259	GIRPSFPAL	X			
901	290	VVFPALFAA	X			
901	21	LSVMAIFYL	X			
901	27	FYLRLPATF	X			
901	92	SNVTATVRS	X			
901	523	IYDQLSGAF	X			
1101	27	FYLRLPATF	X		X	
1101	157	RLRELLHEA	X			
1101	276	RVGVIYHKS	X			
1101	26	IFYLRLPAT	X		X	
1101	24	AIFYLRLPA			X	
1101	125	FRIQRQNTR			X	
1101	288	LVVFPALFA				X
1101	17	VITLSVMAI				X
1101	59	YRGVAVGRV				X
1101	273	LGRVGVIIH				X

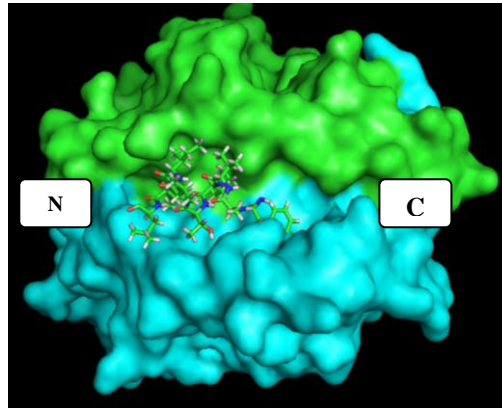
1101	142	LRQAETLLG					X
1101	12	FAVITVITL					X
1101	14	VITVITLSV					X
1101	9	LSIFAVITV					X
1101	77	VTAHMLNS					X
1101	390	YVPVGTNPW					X
1101	278	VIYHKSIEQ					X
1101	259	IRPSFPALA					X
1201	516	ASGPAYTIY	X				
1301	127	IQRQNTRIG					X
1301	231	LRAADPRLR					X
1301	77	VTAHMLNS					X
1301	142	LRQAETLLG					X
1301	288	LVVFPALFA					X
1301	28	LRLPATFGI					X
1301	125	FRIQRQNTR					X
1301	23	MAIFYLRLP					X
1301	17	VITLSVMAI					X
1301	207	LQAQIRAGG					X
1301	9	LSIFAVITV					X
1302	52	LYKNANVTY	X				
1302	8	IQLSIFAVI	X				
1302	70	SVGLNPNGV	X				
1302	128	IQRQNTRIG	X				X
1302	82	MRLNSGTAI	X		X	X	
1302	250	IDEANTAFS	X				
1302	185	LLVDEANAN	X				
1302	53	YKNANVTYR	X				X
1302	125	FRIQRQNTR					X
1302	26	FYLRLPATF					X
1302	231	LRAADPRLR					X
1302	206	FLQAQIRAG					X
1302	77	VTAHMLNS					X
1302	142	LRQAETLLG					X
1302	288	LVVFPALFA					X
1302	28	LRLPATFGI					X
1302	23	MAIFYLRLP					X
1302	520	YTIYDQLSG					X
1302	17	VITLSVMAI					X
1302	207	LQAQIRAGG					X
1501	23	VMAIFYLRL	X				X
1501	287	LLVFPALF	X				
1501	5	LAKIQLSIF	X				
1501	179	LIESARLLV	X				
1501	158	LRELLHEAF	X				
1501	523	IYDQLSGAF	X				
1501	288	LVVFPALFA					X
1501	9	LSIFAVITV					X
1501	28	LRLPATFGI					X
1501	81	MRLNSGTAI					X
1501	121	LRNGFRIQR					X
1501	289	VVFPALFAA					X

1501	362	VRGARNYPC					X
1501	17	VITLSVMAI					X
1501	77	VTAHMLNS					X
1501	278	VIYHKSIEQ					X
1501	14	VITVITLSV					X
1501	520	YTIYDQLSG					X
1501	142	LRQAETLLG					X

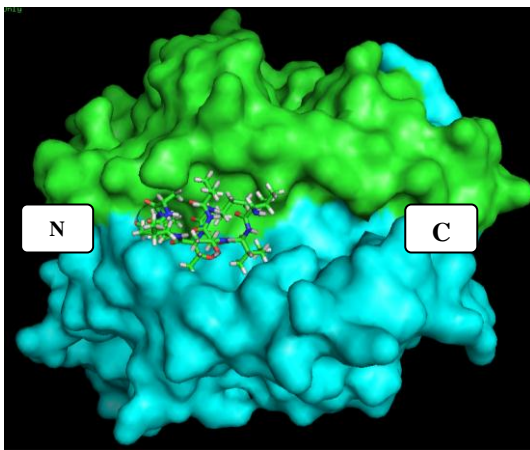
11.2 Docking molecular entre los 21 epítopes T promiscuos (los otros dos epítopes son los enmarcados dentro de la discusión) y el HLA DRB1*0101. Bajo cada imagen se muestra la secuencia de cada epítipo y el aminoácido de inicio de la secuencia.



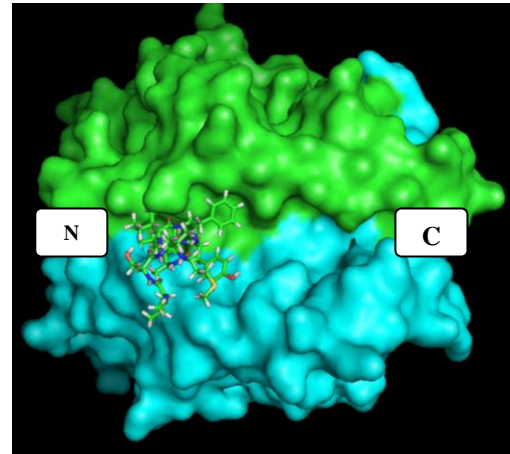
(10) LSIFAVITV



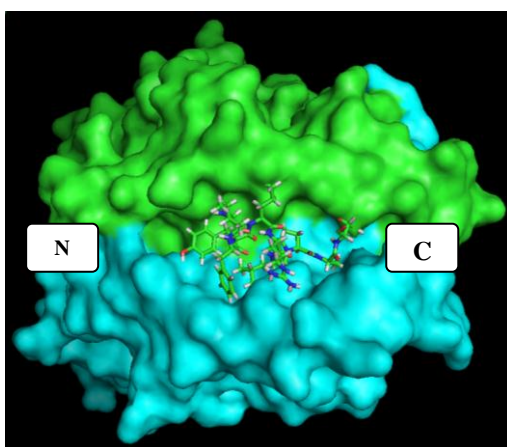
(12) FAVITVITL



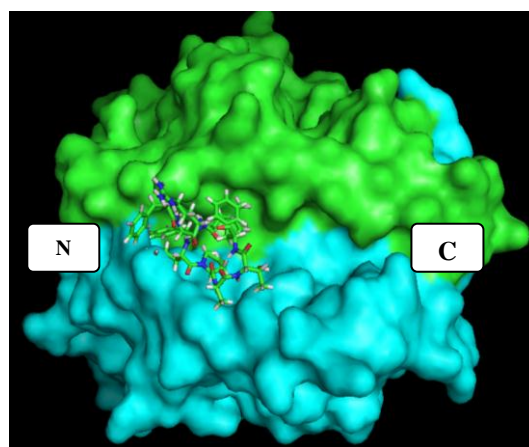
(15) VITVITLSV



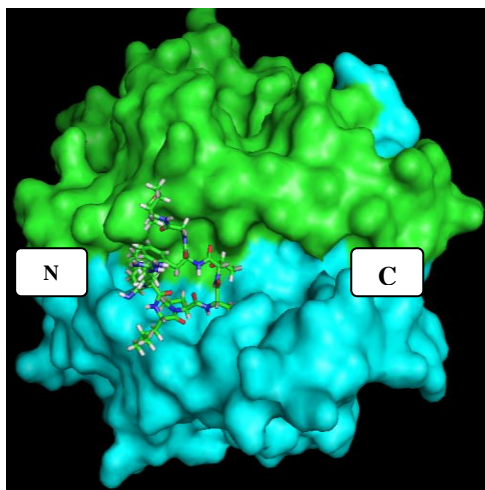
(21) LSVMAIFYL



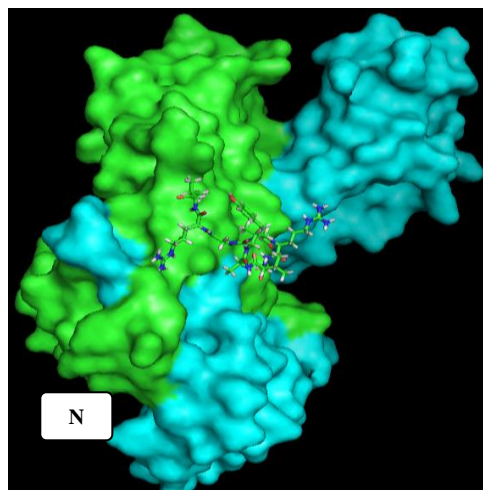
(26) IFYLRLPAT



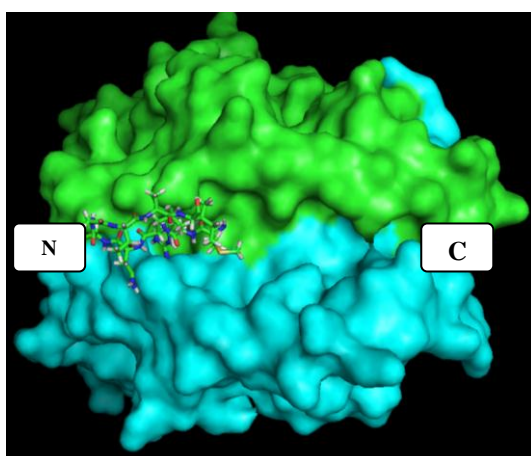
(27) FYLRLPATF



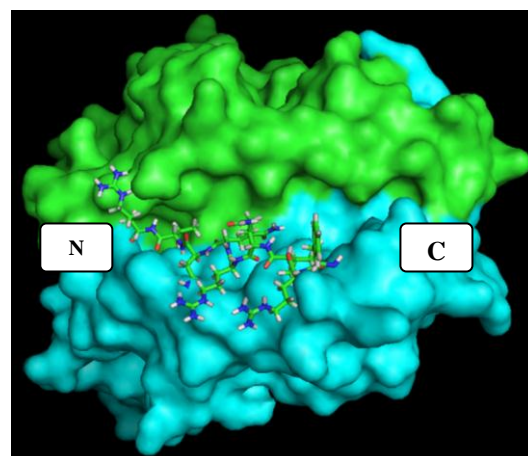
(28) LRLPATFGI



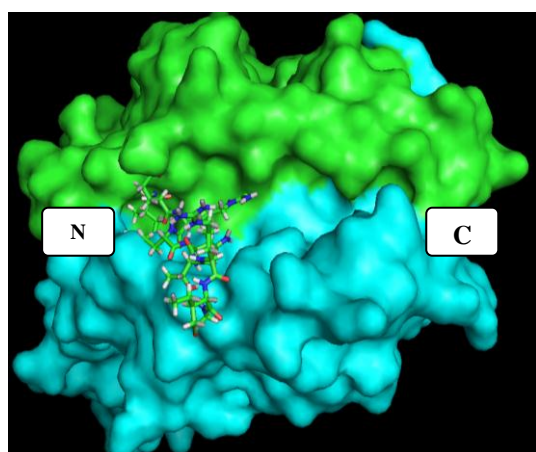
(60) YRGVAVGRV



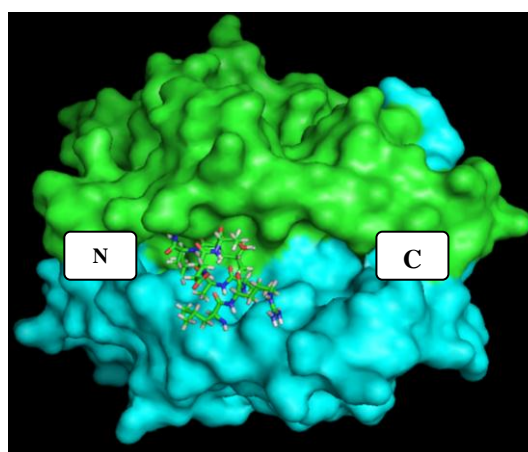
(82) MRLNSGTAI



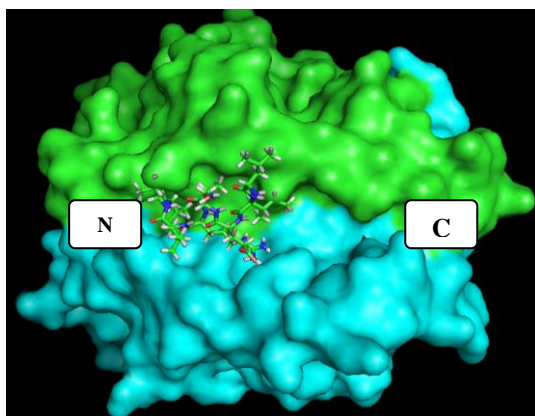
(126) FRIQRQNTR



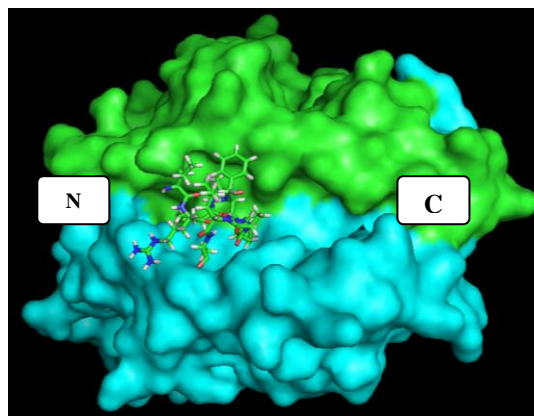
(128) IQRQNTRIG



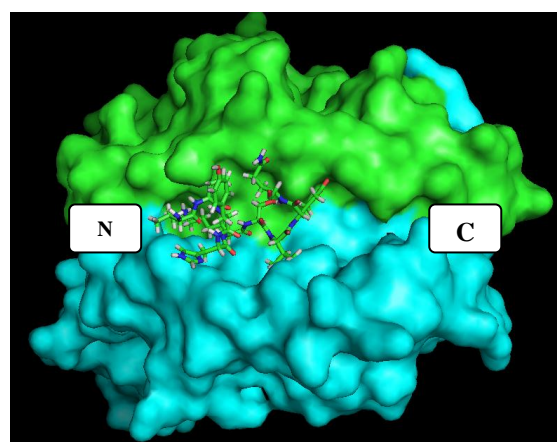
(142) LRQAETLLG



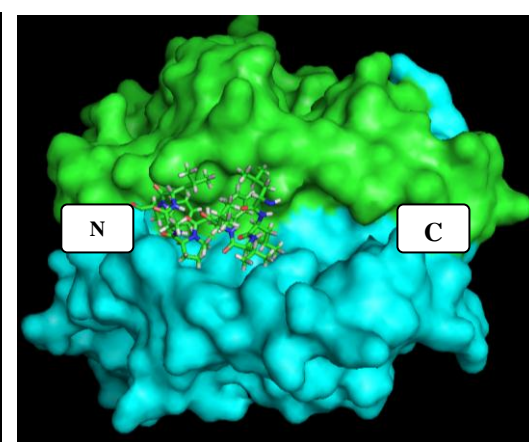
(179) LIESARLLV



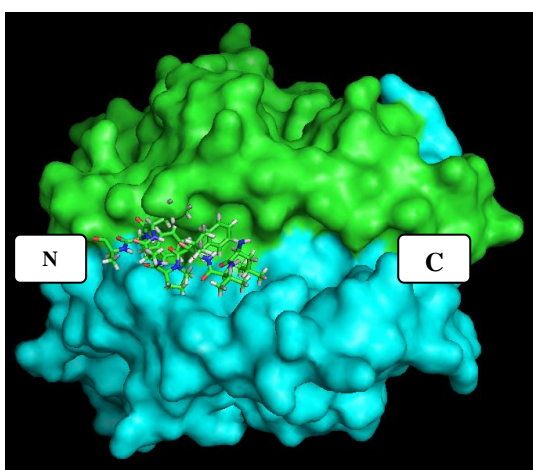
(260) IRPSFPALA



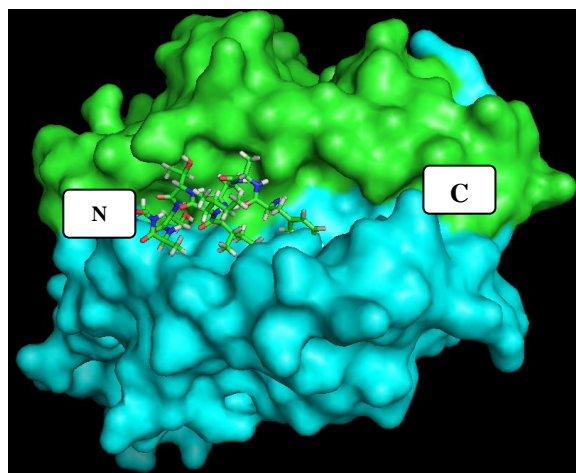
(278) VIYHKSIEQ



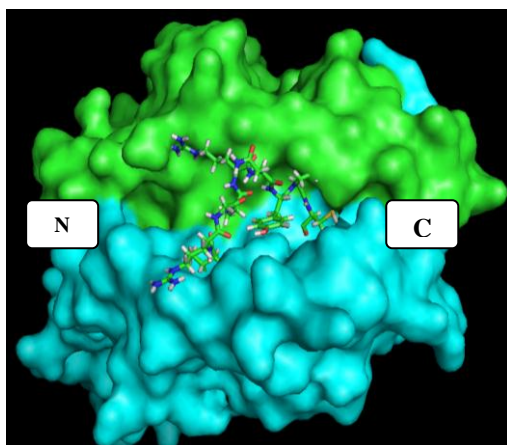
(288) LLVVFALF



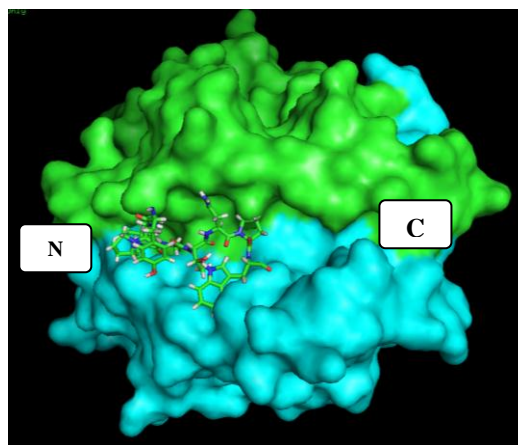
(289) LVVFPALFA



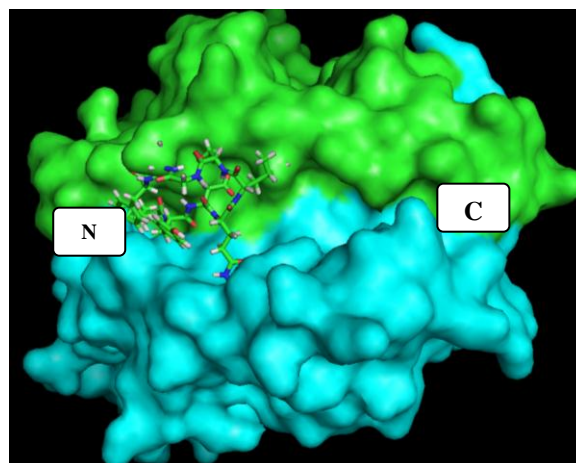
(296) FAAIITSAG



(362) VRGARNYPT



(390) YVPVGTNPW



(523) IYDQLSGAF

11.3 Fluorescencias emitidas por las células con y sin péptido

$\langle \rangle$	1	2	3	4	5	6	7	8	9	10
A	38751	35310	37665	39383	39502	39774	40545	40512	40230	41689
B	37961	35125	37459	39647	40687	39698	40662	43423	40514	40267
C	40611	41376	40138	42737	40536	41591	40894	42710	42364	40335
D	38248	40788	40517	40266	40618	41057	40747	40583	41855	40601
E	44661							37280	2128	2187
F	40664							39028	2101	2096

PEPTIDO	COORDENADA (DUPLICADO)
38369	A1;B1
38370	A2;B2
38371	A3;B3
38373	A4;B4
38374	A5;B5
38376	A6;B6
38377	A7;B7
38378	A8;B8
38379	A9;B9
38380	A10;B10
38382	C1;D1
38385	C2;D2
38386	C3;D3
38387	C4;D4
38388	C5;D5
38389	C6;D6
38390	C7;D7
38394	C8;D8
38395	C9;D9
38396	C10;D10
38397	E1;F1
Control + de bioreducción	E8;F8
Control Medio RPMI	E9;F9
Control HBSS	E10;F10