



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA QUÍMICA Y BIOTECNOLOGÍA

MODELAMIENTO DE SITIOS DE ACOPLAMIENTO E INTERACCIONES PROTEÍNA-PROTEÍNA  
EN RECEPTORES TIPO TOLL

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL EN BIOTECNOLOGÍA

EDUARDO DAVID ROSSEL SALAS

PROFESOR GUÍA:  
J. CRISTIAN SALGADO HERRERA

MIEMBROS DE LA COMISIÓN:  
GERALD ZAPATA TORRES  
ORIANA SALAZAR AGUIRRE

Este trabajo fue financiado por CONICYT mediante  
los Fondos de Iniciación en Investigación  
FONDECYT 11080016 y 1120280

SANTIAGO DE CHILE  
OCTUBRE 2012

## MODELAMIENTO DE SITIOS DE ACOPLAMIENTO E INTERACCIONES PROTEÍNA-PROTEÍNA EN RECEPTORES TIPO TOLL

Enfermedades como la sepsis, artritis reumatoide, tendinitis, entre otras, están ligadas a respuestas pro-inflamatorias anómalas. En ellas participan citoquinas pro-inflamatorias y factores de transcripción activados a partir de la vía de señalización mediada por receptores Tipo Toll y proteínas adaptadoras afines. Estos receptores inician una cascada de señalización a partir de su homodimerización al ser inducidos por señales microbianas. Luego prosigue el reclutamiento de una serie de adaptadores proteicos que terminan finalmente con la transcripción de factores y producción de citoquinas proinflamatorias. En particular, este proyecto se enmarca en el estudio de las vías de señalización en las cuales está involucrado el receptor TLR4.

Las enfermedades mencionadas pueden ser tratadas mediante el uso de fármacos, pero su uso resulta poco efectivo en etapas medio-avanzadas de la enfermedad y la presencia de efectos secundarios serios como osteoporosis, obesidad centrípeta, hipertensión, diabetes mellitus, entre otros. En vista de esto se plantea que la descripción de las interacciones del sistema mediado por TLR4 puede permitir diseñar nuevos tratamientos para estas enfermedades caracterizadas por respuestas inflamatorias anómalas. En función de lo anterior, se define el objetivo general de este trabajo que corresponde al estudio y reconstrucción *in silico* del acoplamiento e interacciones proteína-proteína producidas en la vía pro-inflamatoria activada por TLR4 y proteínas adaptadoras MAL, TRAM y MyD88. Por medio del uso de algoritmos de acoplamiento y herramientas de dinámica molecular se estudia el acoplamiento de los dominios TIR del homodímero de TLR4 y los adaptadores MAL, MyD88 y TRAM. Las interacciones propuestas son caracterizadas determinándose: área de contacto, tipos de interacciones y energía libre de los complejos. De manera adicional, se comparó los resultados con otros modelos presentados en publicaciones anteriores (Nuñez *et al.* (2007), Basith *et al.* (2011)).

Los resultados obtenidos muestran que los adaptadores MAL y TRAM interactúan con el receptor TLR4 mediante la unión a la nueva superficie formada por la homodimerización de los dominios TIR de TLR4. En ambos casos, se observa que tanto MAL como TRAM se acoplan al receptor TLR4 orientando frente a frente sus lazos BB, lo que confirma la relevancia de este sector en las interacciones de esta vía. Otra zona importante de interacción corresponde al lazo CD del adaptador MAL y TRAM, el que según los resultados interactuaría con la Hélice  $\alpha B$  de TLR4. Por su parte, el estudio del complejo TLR4-MAL-MyD88, muestra que el adaptador MyD88 se une preferentemente a la nueva superficie que se produce a partir de la interacción TLR4-MAL. A partir de la estimación de la energía libre de unión vía MM-PBSA, se demuestra que los contactos y el acoplamiento de los modelos propuestos muestran estructuras de mayor estabilidad termodinámica que las presentadas por Nuñez (2007). Los resultados demuestran que zonas como los lazos BB, CD, hélice  $\alpha A$  y  $\alpha C$  son transversalmente relevantes dentro de la cadena de señalización y que presentan características generalizables al tipo de interacciones que se produce en esta vía mediada por el receptor TLR4. Finalmente, se destaca que la metodología planteada en este trabajo se transforma en un procedimiento novedoso y no descrito en trabajos anteriores relacionados a la descripción de interacciones vía TLR. Más aún, el conjunto de herramientas utilizadas permiten modelar las interacciones, probar hipótesis de interacción, brindar explicación a la acción de mutantes y encontrar sectores potenciales de interacción en cualquier tipo de sistema proteína-proteína.

A los que sueñan...

# Agradecimientos

Al finalizar este trabajo, y más aún esta gran etapa de mi vida, quisiera agradecer a todos quienes me brindaron apoyo, dieron fuerzas y alimentaron la llama que me permitió caminar por este largo camino.

Quisiera agradecer el apoyo financiero de CONICYT, quien a través de los Fondos de Iniciación en Investigación FONDECYT 1120280 - *Searching for possible deprotonation sites for substrates and inhibitors of human MAO-A* y FONDECYT 11080016 - *Mathematical modeling of the interaction between cell-penetrating TIR BB loop decoy peptides (BBPs) and toll-like receptor 4 and 2 (TLR4 and TLR2)* apoyaron el desarrollo de este trabajo.

Agradezco a los profesores que me dieron consejo durante esta y otras etapas de mi vida Universitaria. A Gerald Zapata y Oriana Salazar muchas gracias por sus enseñanzas, por la paciencia y el apoyo durante el desarrollo de este proyecto. A Francisco Gracia, muchas gracias por el constante apoyo y las conversaciones acerca de la revolución.

De manera muy especial agradezco a mi profesor guía, Cristián Salgado, por haberme brindado la oportunidad de trabajar en este gran proyecto. Gracias por su apoyo en el desarrollo de esta memoria y por brindarme las herramientas para cruzar este bosque cuando estaba perdido. Más aún, le agradezco el siempre estar ahí en esos momentos difíciles, donde más que ayuda se requiere consejo y una palabra de apoyo. Usted es, sin duda, uno de los mejores docentes de esta Universidad.

Agradezco a mi compañero de batallas, Camilo Patiño, con quién trabajamos durante este proyecto, logrando sacarlo adelante a punta de sueño, café, sudor y lágrimas. A mis compañeros del Laboratorio de Modelamiento de Procesos y Computación Distribuida, gracias por ese espacio de desarrollo, retroalimentación y deliciosas galletas.

Agradezco a la Universidad de Chile, al Departamento de Ingeniería Química y Biotecnología, pero mucho más a su gente, por brindarme una experiencia inolvidable. Gracias Gina, Irene y Erika, el trío de oro del Departamento, por ayudarme a resolver cuanta duda tuve. Gracias Hans, Max, Jorge, Rodrigo y Dani por encender la llama revolucionaria. Gracias Elisa, Carmen, Fernanda, Jose Luis, Valeria, Rhida, Fabián, Francisco, Natalia, Lisset y muchos otros, por tener siempre una palabra de apoyo.

A los amigos que me brindó la Universidad, les agradezco todo; Camila, Cami Bay, Rachel, Gonzalo, Diego, Maria Paz, Reynaldo, Loreto, Gabriela y Nicolás, un gran abrazo y todo el cariño que este pequeño corazón puede brindar.

A los eternos soñadores; Richard, Curotto, Pablo, Esteban y Camilo, gracias infinitas por el oasis brindado en medio de esta ciudad de concreto.

Con mucho cariño, agradezco a mis amigos de siempre. A Cristian, Rodrigo, Gustavo, Rafael, Romy, Diego y Braulio, en quienes encontré un lugar de descanso, escape y apoyo en todo momento.

A los Elefantes. Gracias por la liberación, las luces, el escenario y la música que juntos hemos creado.

A tí, te agradezco lo que vivimos juntos. Aunque nuestros caminos puedan resultar distintos, el alma queda marcada para siempre.

Infinitas gracias a mi familia por darme la libertad de escoger mi camino, permitir mi construcción y mi búsqueda por un lugar en este mundo. Con ustedes las palabras sobran, pues el corazón conoce lo vivido. Los amo.

# Índice general

<b>1. Introducción</b>	<b>8</b>
1.1. Antecedentes Generales . . . . .	8
1.2. Descripción del proyecto . . . . .	12
1.3. Motivación y justificación del proyecto . . . . .	13
<b>2. Objetivos y Limitaciones</b>	<b>14</b>
2.1. Objetivo General . . . . .	14
2.2. Objetivos Específicos . . . . .	14
2.3. Limitaciones . . . . .	14
<b>3. Antecedentes Generales</b>	<b>15</b>
3.1. Estructura del receptor TLR4 y adaptadores . . . . .	15
3.1.1. Receptor TLR4 . . . . .	15
3.1.2. Adaptadores MAL y MyD88 . . . . .	17
3.1.3. Adaptadores TRAM y TRIF . . . . .	19
3.2. Interacciones en la vía de señalización mediada por TLR4 . . . . .	20
3.3. Dinámica Molecular . . . . .	26
3.4. Cálculo de la energía libre . . . . .	28
<b>4. Metodología</b>	<b>30</b>
4.1. Materiales y equipos utilizados . . . . .	30
4.1.1. Proteínas y adaptadores . . . . .	30
4.1.2. Herramientas computacionales . . . . .	31
4.2. Procedimiento . . . . .	32
4.2.1. Selección de los algoritmos de <i>docking</i> . . . . .	32
4.2.2. Preprocesamiento y refinamiento de estructuras . . . . .	33
4.2.3. Docking y Dinámica Molecular . . . . .	33

4.2.3.1.	Aplicación de algoritmos de <i>docking</i> . . . . .	33
4.2.3.2.	Uso de dinámica molecular sobre los complejos obtenidos vía acomplamiento . . . . .	36
4.2.4.	Cálculo de energía libre . . . . .	41
<b>5.</b>	<b>Resultados</b>	<b>43</b>
5.1.	Verificación de la calidad estructural de los modelos utilizados . . . . .	43
5.2.	Modelos de interacción en la vía mediada por TLR4 y adaptadores . . . . .	46
5.2.1.	Efecto del algoritmo de acoplamiento utilizado sobre la obtención de complejos proteína-proteína . . . . .	46
5.2.2.	Uso de acoplamiento sitio dirigido en la búsqueda de modelos de interacción para la vía mediada por TLR4 y adaptadores . . . . .	48
5.2.3.	Descripción de los modelos de interacción obtenidos . . . . .	50
5.2.3.1.	Modelo TLR4-MAL . . . . .	51
5.2.3.2.	Modelo TLR4-MAL-MyD88 . . . . .	52
5.2.3.3.	Modelo TLR4-TRAM . . . . .	53
5.3.	Aplicación de dinámica molecular sobre los modelos de interacción . . . . .	55
5.3.1.	Caracterización de las simulaciones de dinámica molecular . . . . .	55
5.3.2.	Descripción de las interacciones proteína - proteína . . . . .	60
5.3.2.1.	Descripción de los contactos en el homodímero de TLR4 . . . . .	60
5.3.2.2.	Descripción de los contactos en el complejo TLR4-MAL . . . . .	63
5.3.2.3.	Descripción de los contactos en el complejo TLR4-MAL-MyD88 . . . . .	66
5.3.2.4.	Descripción de los contactos en el complejo TLR4-TRAM . . . . .	69
5.3.2.5.	Interacciones conservadas en la vía mediada por TLR4 y adaptadores . . . . .	72
5.3.2.6.	Estimación de la Energía Libre de los complejos proteína-proteína observados en la vía mediada por TLR4. . . . .	76
<b>6.</b>	<b>Discusión</b>	<b>79</b>
6.1.	Validación de los modelos estructurales utilizados . . . . .	79
6.2.	Uso de algoritmos de acoplamiento en el modelamiento de interacciones en la vía mediada por TLR4. . . . .	80
6.3.	Definición de las interacciones proteína-proteína por medio de una estrategia de docking y dinámica molecular . . . . .	84
6.3.1.	Limitaciones de la estrategia de dinámica molecular utilizada . . . . .	85

6.3.2. Consideraciones acerca de la implementación de la dinámica molecular en la definición de interacciones . . . . .	87
6.4. Caracterización y validez de los modelos propuestos para la vía mediada por TLR4 y adaptadores . . . . .	89
6.4.1. Modelo TLR4-TLR4 . . . . .	89
6.4.2. Modelo TLR4-MAL y TLR4-MAL-MyD88 . . . . .	91
6.4.3. Modelo TLR4-TRAM . . . . .	92
6.5. Cálculo de la energía libre de los complejos TLR4-MAL, TLR4-TRAM . . . . .	94
<b>7. Conclusiones</b>	<b>97</b>
7.1. Recomendaciones . . . . .	99
<b>Bibliografía</b>	<b>100</b>
<b>A. Anexo</b>	<b>I</b>
A.1. Secuencias del receptor TLR4 y adaptadores MAL, TRAM y MyD88 . . . . .	I
A.2. Características del software de acoplamiento estudiado . . . . .	I
A.3. Acoplamiento sitio dirigido . . . . .	VI
A.4. Parámetros utilizados en la implementación del protocolo de dinámica molecular en GROMACS . . . . .	VII
A.5. Diferencias entre modelos del receptor TLR4 y adaptadores propuestos por Nuñez <i>et al.</i> (2007) y Salgado (2011) . . . . .	X
A.6. Histogramas RMSD C $\alpha$ durante simulaciones de dinámica molecular . . . . .	XI
A.7. RMSF de los residuos de TLR4 y adaptadores durante las simulaciones de dinámica molecular	XIII
A.8. Otros contactos determinados a partir de la dinámica molecular de los complejos estudiados	XV
A.9. Cálculo de hidrofobicidad superficial promedio de TLR4, adaptadores y complejos . . . . .	XIX
A.10. Detalles de la interacción TLR4-TLR4 . . . . .	XIX
A.11. Estimación de energía libre en función de contactos observados . . . . .	XX
A.12. Representación de las superficies electrostáticas en los complejos TLR4-MAL y TLR4-TRAM.	XXI

# Índice de tablas

3.1. Características de los sitios de interacción TLR4-MAL y TLR4-TRAM . . . . .	22
3.2. Caracterización funcional de residuos de TLR4, MAL y TRAM . . . . .	23
3.3. Resumen del efecto de mutaciones en activación NF- $\kappa$ B e IFN- $\beta$ . . . . .	23
4.1. Resumen de estructuras proteicas utilizadas en el proyecto . . . . .	31
4.2. Herramientas computacionales usadas en el proyecto . . . . .	31
4.3. Software especializado utilizado para el desarrollo del proyecto . . . . .	32
4.4. Criterios para la definición de interacciones durante Dinámica Molecular . . . . .	40
5.1. Evaluación de los modelos estructurales de TLR4, MAL, TRAM y MyD88 . . . . .	43
5.2. Evaluación de los modelos estructurales de TLR4, MAL y MyD88 por medio de ProQ . . . . .	46
5.3. Porcentaje de modelos positivos TLR4-MAL según algoritmo de acoplamiento . . . . .	48
5.4. Tipo de aminoácidos presentes en la zona de interacción TLR4-MAL . . . . .	52
5.5. Tipo de aminoácidos presentes en la zona de interacción TLR4-MAL . . . . .	53
5.6. Tipo de aminoácidos presentes en la zona de interacción TLR4-TRAM . . . . .	55
5.7. Contactos reportados vía MD para la zona de interacción TLR4-TLR4 . . . . .	62
5.8. Contactos encontrados vía MD en la zona de interacción TLR4-MAL . . . . .	66
5.9. Contactos encontrados en la zona de interacción TLR4-MAL-MyD88 vía MD . . . . .	69
5.10. Contactos encontrados en la zona de interacción TLR4-TRAM vía MD . . . . .	72
5.11. Estimación de la energía libre de los complejos . . . . .	77
5.12. Detalle de componente energética por estructura protéica para la estimación de energía libre del complejo TLR4-MAL . . . . .	77
5.13. Detalle de componente energética por estructura protéica para la estimación de energía libre del complejo TLR4-TRAM . . . . .	78
A.1. Secuencias de aminoácidos de los modelos utilizados para TLR4, MAL, TRAM y MyD88 . . . . .	I
A.2. Principales características de los algoritmos de acoplamiento estudiados . . . . .	II
A.3. Aplicaciones de los algoritmos de acoplamiento utilizados . . . . .	III



A.4. Número de citas por algoritmo de acoplamiento . . . . .	IV
A.5. Tiempo promedio de ejecución de algoritmos de acoplamiento . . . . .	V
A.6. Parámetros del archivo de configuración (.mdp) para cada uno de las etapas de MD . . . . .	VII
A.7. RMSD entre estructuras del dominio TIR de TLR4 y adaptadores propuestas por Nuñez (2007) y Salgado (2010) . . . . .	X
A.8. Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-MAL . . . . .	XV
A.9. Contactos con prevalencia media en la zona de interacción TLR4-MAL vía MD . . . . .	XVI
A.10. Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-MAL-MyD88 . . . . .	XVII
A.11. Contactos vía MD con prevalencia media en la zona de interacción TLR4-MAL-MyD88 . . . . .	XVII
A.12. Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-TRAM . . . . .	XVII
A.13. Contactos con prevalencia media en la zona de interacción TLR4-TRAM vía MD . . . . .	XVIII
A.14. Hidrofobicidad superficial promedio según estructura para diversas escalas de hidrofobicidad	XIX
A.15. Estimación de energía libre para complejo TLR4-MAL de acuerdo a los contactos presentados	XX
A.16. Energía de unión según tipo de interacción para complejo TLR4-TRAM de acuerdo a los contactos presentados . . . . .	XX

# Índice de figuras

1.1. Vías de señalización del sistema inmune innato en las que participan los receptores tipo Toll	10
3.1. Estructura del receptor TLR4	16
3.2. Estructura del dímero receptor TLR4	17
3.3. Esquema de la vía de señalización del receptor TLR4 en el sistema inmune innato	18
3.4. Estructura tridimensional de los adaptadores MAL y MyD88	19
3.5. Estructura tridimensional del adaptador TRAM	20
3.6. Complementariedad en superficies de potencial electrostático para interacciones TLR4 y MAL	21
3.7. Resultados obtenidos mediante docking para interacción TLR4 (dímero)/MAL y TLR4 (dímero)/TRAM	22
3.8. Sitios del dominio TIR de MyD88 relevantes para la interacción de acuerdo a Ohnishi <i>et al.</i> 2009	24
3.9. Modelo estructural del complejo de señalización formado por MyD88 y MAL	24
3.10. Diagrama de bloques de una simulación de MD	27
4.1. Diagrama de bloques para la metodología planteada para el proyecto	32
4.2. Estructura del homodímero receptor TLR4	35
4.3. Diagrama de bloques con etapas del protocolo MD utilizado	37
4.4. Ciclo termodinámico utilizado en el cálculo de energía libre vía MM-PBSA	41
5.1. Superposición modelos de Nuñez <i>et al.</i> (2007) vs modelos utilizados.	45
5.2. Modelos de interacción según distintos algoritmos de acoplamiento	47
5.3. Energía para el mejor modelo según sector de acoplamiento	48
5.4. Resultados del acoplamiento sitio dirigido de TLR4-MAL y TLR4-TRAM vía Autodock 4.	49
5.5. Modelos de interacción TLR4-MAL mediante acoplamiento sitio dirigido	50
5.6. Modelo de la interacción TLR4-MAL	51
5.7. Modelo de la interacción TLR4-MAL-MyD88	52
5.8. Modelo de la interacción TLR4-TRAM	54

5.9. Perfiles de energía, temperatura, presión y RMSD ( $C\alpha$ ) de la dinámica molecular de TLR4-MAL . . . . .	56
5.10. Cambios conformacionales en la estructura de MAL durante dinámica molecular de complejo TLR4-MAL (P) . . . . .	57
5.11. Perfiles de energía, temperatura, presión y RMSD ( $C\alpha$ ) de la dinámica molecular de TLR4-MAL-MyD88 . . . . .	59
5.12. Perfiles de energía, temperatura, presión y RMSD ( $C\alpha$ ) de la dinámica molecular de TLR4-TRAM . . . . .	60
5.13. Zonas de contacto observadas vía MD para el complejo TLR4-TLR4 . . . . .	61
5.14. Detalle de la interacción en el complejo TLR4-TLR4 . . . . .	62
5.15. Zonas de contacto observadas vía MD para el complejo TLR4-MAL . . . . .	64
5.16. Detalle de la interacción en el complejo TLR4-MAL . . . . .	65
5.17. Contactos validados experimentalmente para el modelo TLR4-MAL-MyD88 . . . . .	67
5.18. Zonas de contacto observadas mediante dinámica molecular para el complejo TLR4-MAL-MyD88 . . . . .	68
5.19. Detalle de la interacción en el complejo TLR4-TRAM . . . . .	70
5.20. Zonas de contacto observadas mediante dinámica molecular para el complejo TLR4-TRAM . . . . .	70
5.21. Alineamiento de las secuencias para el estudio de conservación de residuos . . . . .	74
5.22. Alineamiento de las secuencias y análisis de la hidrofobicidad de los residuos . . . . .	75
A.1. Esquema de la estrategia sitio dirigido utilizada con Autodock y Vina . . . . .	VI
A.2. Histograma de RMSD $C\alpha$ de TLR4 y MAL durante la simulación del complejo tetramérico TLR4-MAL . . . . .	XI
A.3. Histograma de RMSD $C\alpha$ de TLR4, MAL y MyD88 generado a partir de la simulación del complejo TLR4-MAL-MyD88 (Modelo 1) . . . . .	XII
A.4. Histograma de RMSD $C\alpha$ de TLR4 y TRAM durante simulación del complejo TLR4-TRAM . . . . .	XII
A.5. Fluctuación cuadrática media de los residuos de TLR4 y MAL durante dinámica molecular del complejo TLR4-MAL . . . . .	XIII
A.6. Fluctuación cuadrática media de residuos de TLR4, MAL y MyD88 en dinámica molecular dle complejo TLR4-MAL-MyD88 . . . . .	XIV
A.7. Fluctuación cuadrática media de los residuos de TLR4 y TRAM durante dinámica molecular del complejo TLR4-TRAM . . . . .	XV
A.8. Interacciones hidrofóbicas en el complejo TLR4-TLR4 . . . . .	XIX
A.9. Representación de las superficies electrostáticas de TLR4-MAL y TLR4-TRAM . . . . .	XXI

# Capítulo 1

## Introducción

### 1.1. Antecedentes Generales

El profundo estudio de las vías de nuestro sistema inmune ha permitido determinar la existencia de dos tipos de acciones distintas dentro de éste. Una de estas líneas corresponde al sistema inmune innato, relacionado a una respuesta de tipo inmediato. Esta se caracteriza por utilizar como mecanismo una intrincada red de efectores con el objetivo de eliminar o controlar la replicación de patógenos. La segunda línea corresponde al sistema inmune adaptativo, activado varios días luego de la infección, cuya función es permitir al organismo mantener una respuesta inmune permanente y estable [1, 2].

Ha sido propuesto, para el caso del sistema inmune innato, que el mecanismo de red o cascada de efectores se inicia en base a la detección microbiana vía receptores que son esenciales para gatillar una respuesta inmediata sistemática que permita rápidamente controlar la infección [2]. Este tipo de moléculas se conoce como receptores reconocedores de patrones (PRRs), y cumplen la función de reconocer estructuras microbianas altamente conservadas a lo largo de los microorganismos, conocidos como patrones moleculares patógeno asociados (PAMPs), los que mayoritariamente cumplen un rol crítico en la replicación y supervivencia de los microorganismos. Algunos ejemplos de este tipo de compuestos microbianos son lipopolisacáridos (LPS), lipoproteínas y ácidos nucleicos, virales o bacterianos. La señalización a partir de este tipo de compuestos produce finalmente una cascada de eventos, tales como la inducción de la producción de quimioquinas y citoquinas pro-inflamatorias, el reclutamiento de células fagocíticas y la movilización de células que presenten antígenos [3]. Este tipo de respuestas inflamatorias es normal y beneficiosa en nuestro organismo, dado que es un mecanismo defensivo en contra de la agresión patógena, permitiendo la reparación de los tejidos y la detención de la acción patógena. Sin embargo, en ocasiones se generan respuestas pro-inflamatorias exageradas dando lugar a enfermedades tales como la sepsis, artritis reumatoide, aterosclerosis, asma, tendinitis, laringitis, enfermedad celíaca, entre otras, todas ellas caracterizadas por la pérdida de control de la homeostasis de las vías pro-inflamatorias de nuestro organismo.

Los receptores tipo Toll (TLRs) son receptores transmembrana, de la familia del receptor de interleuquina (IL-1R), caracterizados por ser parte clave en el proceso de reconocimiento de estructuras microbiológicas dentro del sistema inmune innato [4]. Estas moléculas pertenecen al conjunto de los PRRs usados por el sistema inmune para reconocer estructuras PAMPs. Su estructura se caracteriza por poseer un dominio N-terminal con repeticiones ricas en leucina (LRR) para la unión del ligando, un dominio transmembrana y un dominio C-terminal relacionado a la señalización intracelular, denominado dominio TIR [3]. Los receptores TLR se expresan en una gran variedad de tipos celulares, incluyendo células no-hematopoyéticas, epiteliales y endoteliales. En la actualidad se han descrito 13 tipos distintos de TLRs en mamíferos, 10 en humanos y 13 en ratas. Sin embargo, solo los TLRs 1 al 9 son conservados tanto en humanos como en ratas, TLR10 por su parte se presenta solo en humanos y TLR11 exclusivamente en ratas [5]. Numerosos trabajos señalan que la diversidad estructural de los TLR se encuentra en una región *loop*, lo que explica la especificidad molecular en las interacciones de la cascada de señalización. De manera más específica se ha señalado que el *loop* BB es un punto clave en la funcionalidad de los dominios TIR de

los receptores tipo Toll [4].

La transducción de la señal post receptor requiere, en el caso de los receptores tipo Toll, de la participación de varias proteínas adaptadoras conservadas, ya que al contrario de otra clase de receptores, los TLRs no poseen actividad enzimática propia. En lugar de esto, se produce una dimerización u oligomerización que se cree tiene la capacidad de reacomodar el dominio TIR, de tal manera que actúen como un escalón de reclutamiento de proteínas adaptadoras río abajo [6]. La dimerización de monómeros de TLR es considerada ampliamente como un evento necesario en la transducción de la señal intracelular y se cree que es el resultado del contacto a nivel tanto extracelular como citoplasmático [7]. Los receptores tipo Toll, al detectar ciertos compuestos microbianos generan una cascada de señalización que parte con este evento de dimerización y continúa con el reclutamiento de una serie de adaptadores que se van acoplando de manera de transmitir una señal que finalmente permite la transcripción de factores proinflamatorios o generar otras respuestas ya mencionadas. Con mayor detalle se aprecia, en la figura 1.1, las distintas vías de señalización mediadas por receptores TLR. En todos los casos se tiene que PAMPs unen a los receptores TLR para provocar su homodimerización en la membrana plasmática (TLRs 1-6) o dentro de la membrana de compartimientos endosomales (TLRs 7-9) [3]. A partir de ese evento se produce la unión de una serie de adaptadores. Los principales adaptadores en esta cascada de señalización se caracterizan por ser reclutados vía interacciones de los dominios TIR, ellos son: MyD88, MAL, TRIF, TRAM y SARM [8]. En el caso de la señalización vía TLR 1/2/5/6/7-9 esta se lleva a cabo exclusivamente por medio de la intervención del adaptador MyD88. Por su lado el receptor TLR3 utiliza de manera exclusiva el adaptador TRIF para su señalización, mientras que el receptor TLR4 utiliza las vías mediadas por TRIF y MyD88. La cascada prosigue mediante el reclutamiento de una serie de adaptadores proteicos (IRAK-1, IRAK-4, RIP1, TRAF5, IRF7, TAK1 y otros dependiendo de la vía) que terminan finalmente con la transcripción de factores como NF- $\kappa$ B, IRF-3/7/5 y AP-1, lo que finalmente converge en la producción de citoquinas proinflamatorias e interferones tipo I [3].

En particular, este proyecto se enmarca en el estudio de las vías de señalización en las cuales está involucrado el receptor TLR4. El TLR4 reconoce principalmente el lipolisacárido LPS, un componente de la membrana externa de las bacterias Gram negativas [9] [10], del que se ha documentado que, incluso en cantidades trazas de LPS, permite la activación del sistema inmune vía TLR4 llevando a la producción de mediadores proinflamatorios. Este receptor actúa como un homodímero inducido por señales microbianas, a partir del cual se reconstruye una cadena de adaptadores tal como se aprecia en la figura 1.1. Tal como se menciona con anterioridad, el receptor es capaz de establecer una vía de señalización por medio de su unión a los adaptadores MyD88 y TRIF [3]. Esta unión se produce, previa intervención de un adaptador secundario, respectivamente MAL o TRAM para las vías mediadas por MyD88 o TRIF, que cumple una función de mediador en la unión de MyD88 y TRIF [11, 12].

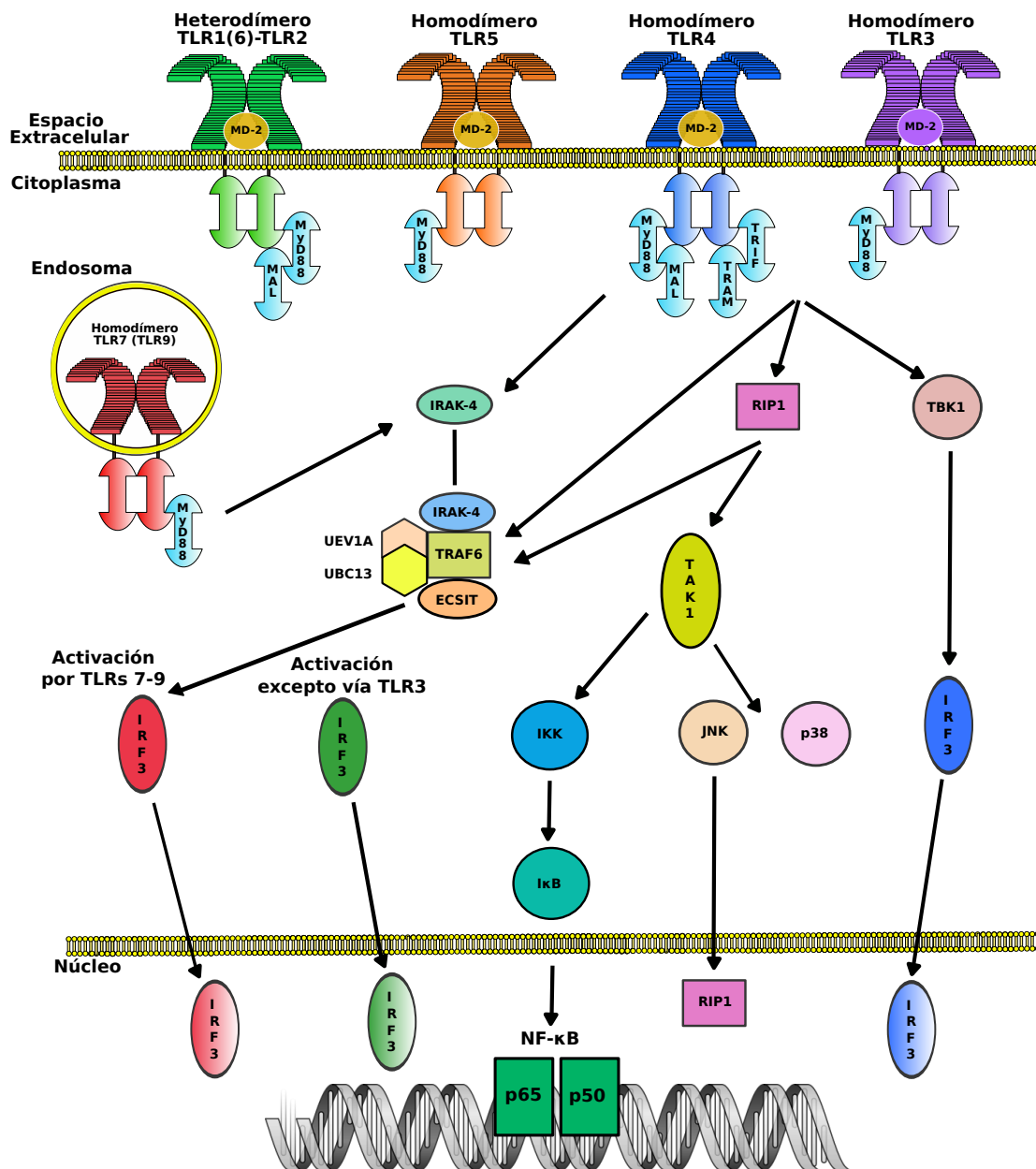


Figura 1.1: Vías de señalización del sistema inmune innato en las que participan los receptores tipo Toll. Adaptado de West et al. 2006. Vías de señalización mediadas por TLR4. Los TLR 1-6 se encuentran ubicados en la membrana plasmática de la mayoría de las células, mientras que los TLRs 7-9 se encuentran en compartimentos endosomales. La unión de factores microbianos conservados produce cambios conformacionales y la dimerización de los receptores. A partir de este evento, se unen adaptadores como MAL (TIRAP en la figura), TRAM, MyD88 y TRIF por medio de interacciones que se producen entre los dominios TIR de las distintas proteínas. Los receptores que se asocian a MyD88 reclutan otros adaptadores como IRAK-1 e IRAK-4, finalmente y luego de la activación de otras proteínas como TRAF6, TAK1, IKK y JNK se produce la activación de NF-κB y AP-1 y la transcripción de efectores del sistema inmune innato. De manera similar, la activación del complejo formado por MyD88, IRAK y TRAF6 activa los factores de transcripción IRF5 y IRF7 para la producción de citoquinas. En el caso de la vía mediada por TRIF, se promueve la activación de TRAF6 y RIP1. Esto finalmente resulta en la actividad de NF-κB y AP-1, produciendo efectores del sistema inmune innato [3].

Para estudiar las vías de señalización compuestas por TLRs, se vuelve necesario entender las interacciones proteína-proteína que dominan la cascada de señales. En general, conocer las estructuras tridimensionales de los complejos proteína-proteína es de suma importancia en el entendimiento de las funciones biológicas y comprensión de las interacciones entre proteínas. Experimentalmente, es posible determinar la estructura de un complejo proteico mediante cristalografía de rayos X [13], sin embargo esto requiere de etapas de purificación de proteínas y de una gran habilidad para cristalizar cualquier tipo de complejo proteico, los que en muchos casos tienen un carácter débil o transitorio. Otra posibilidad es estudiar la estructura de un complejo proteico y sus partes vía homología con alguna otra proteína cuya estructura haya sido ya determinada, de manera de generar modelos moleculares de cada uno de los miembros del complejo o del complejo mismo. En el caso de contar con la estructura tridimensional de los miembros de un complejo proteico, es posible también determinar su interacción a través de métodos de *docking* o acoplamiento.

*Docking* es el término utilizado para los esquemas computacionales que intentan encontrar el mejor acoplamiento entre dos moléculas: un receptor y un ligando. El problema puede ser definido como el de la predicción de una correcta asociación entre los sets de coordenadas atómicas de dos moléculas, tales como proteínas, ácidos nucleicos, péptidos, etc [14]. Dentro del esquema de *docking* existen 2 tipos particulares de problemas, llamados *bound* y *unbound*. En el primero de los casos, el problema consiste en que a partir de un complejo cristalizado de dos o más moléculas, se obtenga las estructuras del ligando y complejo. Luego, el objetivo final del problema es reconstruir el complejo original. Por otro lado el problema *unbound*, consiste en la predicción ciega de un complejo a partir de 2 moléculas en sus conformaciones nativas. Este último problema resulta difícil para la mayoría de los algoritmos de *docking*, ya que la mayoría de los casos requiere la consideración de cambios conformacionales [15]. Sin embargo, existe una gran cantidad de algoritmos y trabajos dedicados a resolver, con distintos grados de éxito, problemas de predicción ciega. Algunos casos de éxito se detallan en los reviews publicados [14, 16, 17] y en los informes de las iniciativas CAPRI (*Critical Assessment of PRediction of Interactions*) [18] y CASP (*Critical Assessment of Structure Prediction*) [19]. La propuesta de estas dos últimas iniciativas, CAPRI y CASP, apunta al constante mejoramiento de herramientas que permitan llevar a cabo la predicción ciega de interacciones biológicas, destacando aquellos desafíos futuros que la técnica debe esforzarse en resolver [18, 19]. Adicionalmente, cabe destacar que los algoritmos de *docking* han sido utilizados para describir parte de las interacciones observadas en la vía del sistema inmune innato mediada por TLR4 [20–23].

Las principales limitaciones de este tipo estas técnicas están ligadas a la incorporación de flexibilidad en los algoritmos. Este tipo de consideraciones dificulta la resolución y el desempeño de los algoritmos, por un lado aumenta el número de grados de libertad y adicionalmente incrementa la razón de falsos positivos dentro de las soluciones. Adicionalmente, Lo anterior implica la necesidad de implementar un sistema preciso de puntuación de modo de seleccionar conformaciones de relevancia biológica desde el conjunto de soluciones obtenidas. En vista de lo anterior es que generalmente, se limita la flexibilidad a solo algunos tipos de movimiento y solo uno de los componentes de la interacción [14, 24].

Se debe tener en cuenta que cada algoritmo de *docking* puede en principio generar su propia jerarquía de resultados en base a la función de puntuación que utilice. Por lo mismo, se hace necesario calcular una magnitud independiente a los métodos de *docking* utilizados, que permita dar con una estimación real de la afinidad de cada complejo. En este contexto se utiliza la energía libre de unión que corresponde a la función termodinámica de estado que entrega la información acerca del estado de equilibrio del sistema, considerándose que un sistema, en el que la temperatura, número de partículas y volumen son constantes, está en equilibrio si el valor de su función de energía libre es mínimo. [25]

El cálculo de la energía libre puede llevarse a cabo mediante los métodos basados en principios termodinámicos, que corresponden a la forma más exacta de calcular la energía libre considerando la teoría de la energía libre de perturbación y la integración termodinámica [26]. Sin embargo, estos métodos terminan

siendo poco prácticos por el tiempo que toman en el cálculo directo de la energía libre. Otras formas de calcular la energía libre son los métodos que aprovechan el hecho de que la energía libre es una función de estado y por tanto en principio solo es necesario evaluar los estados finales e iniciales del complejo para calcular la energía. Dentro de este grupo se encuentran métodos como el MM-PBSA (Molecular Mechanics Poisson – Boltzmann Surface Area), que basan el cálculo en el análisis de distintas configuraciones del complejo proteico obtenidas a partir de simulaciones de dinámica molecular con solvente explícito; Un segundo método a considerar dentro de este tipo de metodologías es el Análisis de Interacción Linear (LIE), que corresponde a un esquema basado en la idea de que la energía libre puede ser descompuesta en contribuciones polares y no polares, donde las contribuciones polares se estiman mediante la teoría de respuesta lineal [24]. Además de las metodologías ya mencionadas, cabe destacar que también es posible utilizar funciones empíricas para el cálculo de la energía libre, estas funciones incorporan distintos parámetros a la función de energía (electrostática, fuerzas Van der Waals, puentes hidrógeno) y usan un set de entrenamiento para calibrar la función empírica. Estos métodos resultan rápidos en el cálculo, pero muchas veces presentan estimaciones insuficientes [27].

## 1.2. Descripción del proyecto

En vista de los antecedentes anteriores es que nace el interés por precisar la descripción de las vías de señalización en las que participa el receptor TLR4. Algunas de las interacciones relevantes en la vía mediada por el receptor TLR4 han sido descritas con anterioridad en los trabajos de Núñez *et al.* (2007) [20], Dunne *et al.* (2003) [23], Ohnishi *et al.* (2008) [22] y Basith *et al.* (2011) [21], entre otros. Estos estudios han permitido obtener datos relevantes acerca de las interacciones entre TLR4 y los adaptadores MAL y MyD88, demostrando que efectivamente se han producido grandes avances en la caracterización de las vías protagonizadas por TLR4. Aún así, el enfoque presentado en este trabajo contempla una serie de propuestas que hacen atractivos y de gran relevancia los resultados esperados de este proyecto.

De manera más precisa, este trabajo considera la utilización de seis algoritmos de *docking* en la determinación de los sitios de interacción en la cadena de señalización. El objetivo de esto es no sesgar la búsqueda de los sitios de interacción en base al algoritmo de *docking* utilizado y su particular método de búsqueda, sino más bien aprovechar las características de cada uno y determinar un sitio de consenso a partir de los resultados que cada uno de ellos entregue. Además, la metodología planteada en este trabajo contempla el uso de la dinámica molecular para el estudio de las estructuras obtenidas mediante el *docking*, lo que permitirá analizar dinámicamente los contactos encontrados brindando mayor validez y precisión a los resultados encontrados. La dinámica molecular ha sido utilizada con anterioridad en el contexto del estudio de las vías mediadas por TLR4, pero su implementación se ha limitado al refinamiento de estructuras previo al uso de algoritmos de *docking* [20, 21]. En el caso de este trabajo se plantea un enfoque nuevo para el uso de la dinámica molecular, que consiste en la aplicación de esta técnica para determinar los contactos y cambios conformacionales que se pueden producir, a lo largo del tiempo, sobre los complejos obtenidos por *docking*. Se cree que este enfoque presenta claras ventajas por sobre las metodologías utilizadas hasta el momento, dado que permite observar los cambios fisicoquímicos y la dinámica de interacción de los complejos estudiados. Finalmente, la metodología considera el cálculo de la energía libre de los complejos determinados. Este es un punto muy relevante ya que brinda información inequívoca acerca de la afinidad y estabilidad termodinámica de los complejos propuestos. Cabe destacar que la estimación de energía libre es un dato que no ha sido considerado en ninguna de las publicaciones revisadas, por lo que se convierte en una contribución importante al estudio de las vías mediadas por el receptor TLR4.

En vista del nuevo enfoque planteado es que se pretende estudiar el acoplamiento e interacciones proteína – proteína de la cadena de señalización en la que participan los receptores tipo Toll y adaptadores. De manera más concreta, la propuesta se traduce en llevar a cabo la reconstrucción de la vía de señalización



en base a las interacciones de sus partes, los TLRs y proteínas adaptadoras, de modo de dar con los factores críticos de cada interacción. Adicionalmente, el proyecto contempla en una etapa posterior, la minería de datos aplicada a los resultados obtenidos desde el acoplamiento y el estudio de interacciones. La intención que hay detrás de esta última etapa es la de tratar de identificar patrones de interacción que sean característicos a la vía de interacción mediada por TLR4 y a los que además pueda asignarse una relevancia biológica. Finalmente, la caracterización de estas vías permitirá entender los factores más relevantes en la interacción de los receptores TLR con sus proteínas adaptadoras, de modo que, a partir de esta información, sea posible la generación de nuevas estrategias que permitan evitar los eventos que llevan a la transcripción de factores proinflamatorios, sobre todo en casos anómalos.

### **1.3. Motivación y justificación del proyecto**

Se debe destacar que el proyecto resulta de interés dadas las posibles aplicaciones que permitirá la caracterización de las vías de señalización del sistema inmune innato. A partir de esta descripción será posible el desarrollo de nuevas estrategias que permitan predecir la susceptibilidad a enfermedades, monitorear terapias y en última instancia desarrollar nuevas estrategias para la prevención y tratamiento de enfermedades crónicas asociadas con una fuerte respuesta pro-inflamatoria. Actualmente a este tipo de enfermedades se les da un tratamiento mediante medicamentos antipalúdicos, antiinflamatorios, antirreumáticos modificadores de la enfermedad y corticosteroides [28]. En general este tipo de tratamiento resulta poco efectivo en etapas medio-avanzadas de la enfermedad, y en particular los dos últimos tipos de medicamentos suelen presentar efectos secundarios serios, tales como osteoporosis, obesidad centrípeta, alteraciones de la cicatrización de heridas, hipertensión, diabetes mellitus, entre otros [29].

Por lo mismo, este proyecto busca brindar nuevos antecedentes que permitan diseñar nuevos tratamientos para estas enfermedades caracterizadas por respuestas inflamatorias anómalas. A modo de ejemplo se tiene que en base a la descripción que se busca obtener será posible diseñar, optimizar y testear péptidos específicos que puedan detener las interacciones de la cascada de señales permitiendo detener las respuestas inflamatorias crónicas típicas, mediante un tratamiento específico que evite la aparición de efectos secundarios provocados por drogas altamente invasivas.

Finalmente, es de suma importancia destacar el aporte que se brinda a este proyecto desde la perspectiva de la Ingeniería Civil en Biotecnología, ya que se entiende de manera completa las variables que están determinando el fenómeno biológico que se está estudiando, al mismo tiempo que se maneja los conceptos y las habilidades necesarias para la resolución del problema de optimización que se está tratando. Por lo mismo es posible llevar a cabo un análisis completo de la metodología y de los resultados obtenidos, de modo de generar un propuesta atractiva de modelamiento, diseño y optimización que pueda en el largo plazo, permitir generar estrategias concretas para combatir enfermedades caracterizadas por una respuesta proinflamatoria no controlada.

## Capítulo 2

# Objetivos y Limitaciones

### 2.1. Objetivo General

*Estudiar el acoplamiento e interacciones proteína-proteína de la cadena de señalización en la que participan el receptor tipo Toll TLR4 y los adaptadores MAL, TRAM y MyD88.*

### 2.2. Objetivos Específicos

Con la idea de alcanzar el principal objetivo planteado, es que se han definido los siguientes objetivos específicos.

- Estudio y comprensión de vías proinflamatorias activadas por TLR4.
- Identificación de las variables que caracterizan las interacciones proteína – proteína en vías activadas por TLR4.
- Selección de algoritmos para el estudio de interacciones proteínas, en función de las ventajas comparativas y limitaciones que presenten.
- Determinación sitios de unión entre proteínas TLR4 y sus adaptadores mediante el uso de algoritmos de docking de modo de reconstruir el segmento de la vía de señalización que involucra al receptor TLR4 y los adaptadores.
- Caracterización de los sitios de unión y las interacciones asociadas al acoplamiento del receptor TLR4 y los adaptadores, obtenidas a través de los algoritmos de docking.
- Análisis de las interacciones obtenidas por docking y reconocimiento de patrones de interacción que puedan ser generalizables dentro de la vía mediada por el receptor TLR4 y las proteínas adaptadoras involucradas.
- Aplicación de minería de datos para determinar, a partir de los datos recolectados, patrones generalizables dentro de las interacciones mediadas por el receptor TLR4.

### 2.3. Limitaciones

El proyecto se centra en la descripción de las vías de señalización en la que participa el receptor TLR4 y considera solamente el estudio de las interacciones con los adaptadores MAL, TRAM y MyD88. Otros

factores que pudieran ser parte de la vía de señalización no serán considerados. Aun así se plantea la necesidad de modelar la estructura del adaptador TRIF, que también interactúa directamente en esta vía, pero esto último escapa los alcances de este trabajo.

El estudio desarrollado se dedica exclusivamente a estudiar aquellas interacciones basadas en la premisa de que el reclutamiento en la vía del sistema inmune parte con la homodimerización del receptor TLR4, tal como es planteado por Nuñez et al. [20] Si bien otros autores han estudiado las interacciones en esta vía a partir de un monómero de TLR4 [23], la mayor parte de las publicaciones se inclinan por la primera hipótesis mencionada. [21, 30, 31]

Una de las principales limitaciones del estudio realizado, es que los contactos definidos son determinados exclusivamente por medio de algoritmos de acoplamiento en conjunto con dinámica molecular, y no mediante desarrollo experimental *in vivo*. Por tanto, si bien los protocolos para la predicción de los complejos fueron cuidadosamente preparados, los resultados pueden estar sesgados por los métodos de búsqueda y funciones de puntuación utilizadas. Finalmente, la propuesta de este trabajo es la producción de modelos de interacción *in silico*. Como tales, estos modelos presentan evidencia que nunca puede ser mejor que la generada experimentalmente en el laboratorio.

## Capítulo 3

# Antecedentes Generales

### 3.1. Estructura del receptor TLR4 y adaptadores

#### 3.1.1. Receptor TLR4

La familia de receptores TLR, posee una estructura común altamente conservada que consiste un dominio N-terminal con repeticiones ricas en leucina (LRR) para la unión del ligando, un dominio transmembrana y un dominio C-terminal relacionado a la señalización intracelular, denominado dominio TIR. La disposición de estos 3 dominios puede se observa en la Figura 3.1 [3].

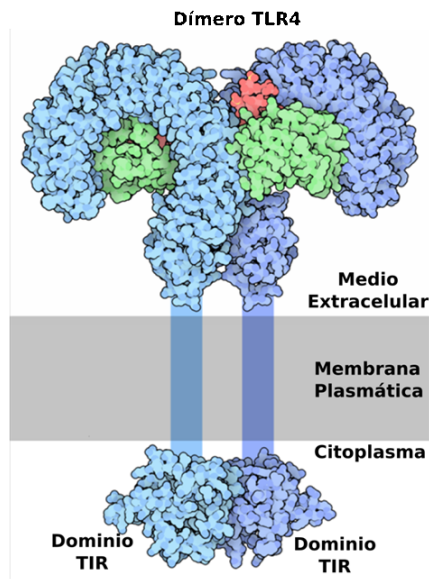


Figura 3.1: Estructura del receptor TLR4. En color azul y morado se aprecia la estructura de los receptores TLR4 formando un homodímero. Se observan los 3 dominios característicos de los receptores TLR: N-terminal, Transmembrana y C-Terminal o TIR. Adicionalmente, se muestra en verde el co-receptor MD-2 necesario para el reconocimiento de algunos factores microbianos como LPS, que se observa en rojo en la figura [32].

El dominio TIR del receptor TLR4 tiene un largo de 139 aminoácidos. Su estructura no ha sido resuelta por medio de cristalografía de rayos X, pero sí ha sido reconstruida por medio de modelamiento como dímero [20] y como monómero [23], en base a las estructuras de los dominios TIR de receptores TLR1, TLR2 y TLR10 cuya estructura sí ha sido resuelta vía cristalografía. [4] [33] Estructuralmente el dominio TIR está formado principalmente por un centro de 5 hojas  $\beta$  paralelas (BA-BE, denominadas alfabéticamente desde el extremo N-terminal) rodeadas por 5  $\alpha$ -hélices (AA-AE). La estructura recién descrita puede observarse en cada uno de los dominios TIR que muestra la figura 3.2, los que en este caso se encuentran formando el homodímero de TLR4. Se ha demostrado que las cadenas de aminoácidos o *loops* que unen estas estructuras tiene una importante función y son fundamentales para las interacciones TIR-TIR. [4] [23] [20] [34].

En particular, el BB *loop* es un segmento, de alrededor de 10 aminoácidos, que conecta la hoja  $\beta$ B con la hélice  $\alpha$ B dentro de los dominios TIR y que forma una protrusión en la superficie de la proteína [4]. Una representación atómica de los residuos que componen el BB *loop* se observa en la figura 3.2. Se ha señalado que este *loop* es clave en la funcionalidad de los dominios TIR. Más aún, algunos de sus aminoácidos altamente conservados, como el residuo de prolina, han sido reportados como factores clave en la señalización [4] [34] [35]. Como antecedente adicional, se demostró en el homodímero de TLR10, que el BB *loop* es parte de la superficie de interacción [33]. Además, estudios de modelamiento estructural muestran que el BB *loop* es parte de la superficie de interacción del dímero TLR4 [20,21]. A continuación se muestra el dímero formado, en el interior celular, por los dominios TIR de los receptores TLR4, destacándose los BB *loops* en la figura 3.2.

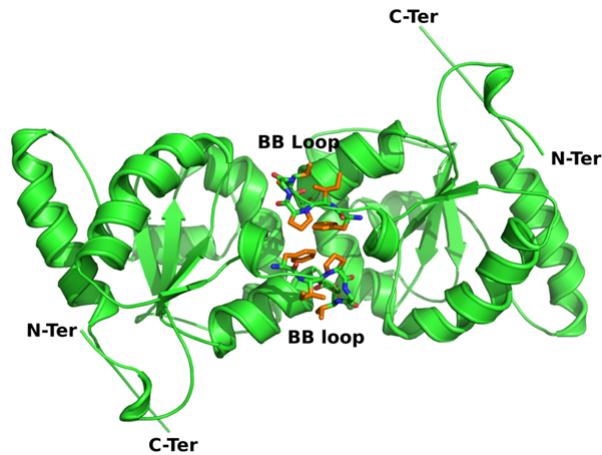


Figura 3.2: Estructura del dímero receptor TLR4. Los dominios TIR de los receptores TLR4, ubicados en el interior celular, forman una estructura homodimérica en la que se enfrentan los loops BB de cada dominio TIR. Se destaca la estructura atómica de los residuos que conforman los BB loops de cada uno de los dominios TIR. Los modelos utilizados fueron generados mediante el software MODELLER (Salgado 2011, Resultados no publicados).

### 3.1.2. Adaptadores MAL y MyD88

Las proteínas adaptadores MAL (*MyD88 adaptor like*) y MyD88 (*Myeloid differentiation primary response protein 88*) son parte de la vía de señalización mediada por TLR4 que lleva a la producción de citoquinas proinflamatorias NF- $\kappa$ B (ver Figuras 1.1 y 3.3) La funcionalidad de ambos adaptadores, está asociada a la estructura del dominio TIR de cada una de estas proteínas. En particular, el dominio TIR del adaptador MAL tiene un rol conector que permite la interacción entre TLR4 y MyD88. Lo anterior implica, que el dominio TIR del adaptador MyD88 no une directamente a TLR4, sino que lo hace mediante la unión a MAL. Esto es consistente con el hecho que la superficie electrostática del dominio TIR de MAL es altamente electronegativa, mientras que la superficie de MyD88 es electropositiva al igual que la del dominio TIR del receptor TLR4 [23, 36].

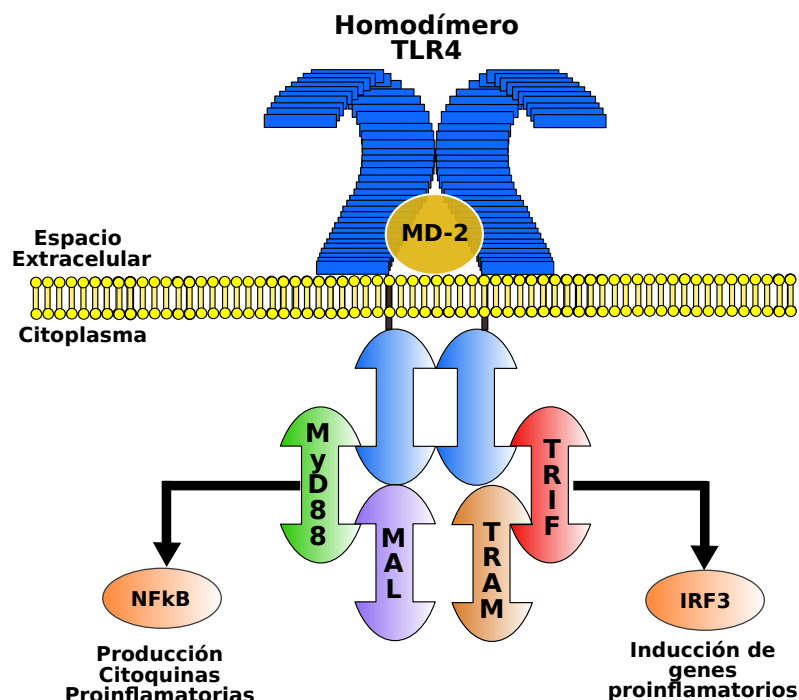


Figura 3.3: Esquema de la vía de señalización del receptor TLR4 en el sistema inmune innato. *El receptor TLR4, acoplado al co-factor MD-2, es capaz de reconocer factores microbianos. Cuando esto ocurre se produce la dimerización del dominio TIR (intracelular) de los receptores TLR4. Las proteínas adaptadoras MAL y MyD88 pueden unirse al homodímero del receptor TLR4, por medio de sus dominios TIR, lo que finalmente produce la síntesis de citoquinas pro-inflamatorias NF-κB. En este caso, el adaptador MAL cumple una función de conector entre TLR4 y MyD88. Análogamente los dominios TIR de TRIF y TRAM pueden unirse al receptor TLR4 para inducir la transcripción de genes pro-inflamatorios por medio de IRF3 [3, 36]*

En la figura 3.4 se observa, en el lado izquierdo, la estructura del dominio TIR de MAL. En ella se observa que está formada por 4 hojas  $4\beta$  rodeadas por 4  $\alpha$ -hélices, en las que el BB loop adopta una conformación fija similar a la observada en receptores TLR [20]. El dominio TIR de este adaptador tiene un largo de 150 aminoácidos. Además, el receptor MAL tiene un dominio en su extremo N-Terminal que une a fosfatidilinositol-4,5-bisfosfato, lo que permite la unión de MAL a la membrana plasmática [36].

Por su lado, el adaptador MyD88 posee una estructura similar a la de los receptores TLR, lo que se observa en la estructura del dominio TIR de MyD88 presente en la figura 3.4 (lado derecho de la imagen). Es una proteína modular con un dominio TIR y un dominio de interacción proteico denominado dominio Death [23]. El dominio TIR de MyD88, que tiene un largo de 149 aminoácidos, le permite interactuar con TLR4 mediante su unión MAL, mientras que el dominio Death funciona como punto de interacción río abajo con el dominio Death del extremo N-terminal de IRAK [11, 37–39].

La figura 3.4 muestra la estructura del dominio TIR de los adaptadores MAL y MyD88, destacándose la representación de la estructura atómica de los residuos que forman parte del BB loop.

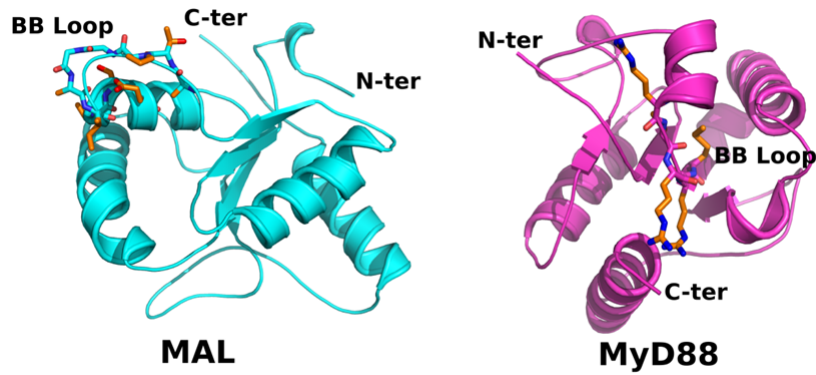


Figura 3.4: Estructura tridimensional de los adaptadores MAL y MyD88. Se muestra la estructura terciaria de los dominios TIR correspondientes a los adaptadores MAL y MyD88. Los sectores correspondientes a los lazos BB han sido destacados mostrando la estructura atómica de sus residuos. El receptor MAL fue generado mediante el software MODELLER (Salgado 2011, Resultados no publicados). El adaptador MyD88 fue obtenido de Protein Data Bank, Código 2JS7, y luego refinado por Salgado et al. 2011 (Resultados no publicados)

### 3.1.3. Adaptadores TRAM y TRIF

Los adaptadores TRAM (*TRIF-related adaptor protein*) y TRIF (*TIR-domain containing adaptor protein inducing interferon- $\beta$* ) forman parte de la vía de señalización mediada por TLR4 no dependiente de MyD88, estimulando una cascada de señalización que lleva a la activación del factor interferón de respuesta IRF-3 (ver Figura 3.3) [20]. En este caso, el dominio TIR del adaptador TRAM actúa interactuando con el dímero del receptor TLR4, formando una nueva superficie de interacción mediante la cual el dominio TIR del adaptador TRIF es reclutado para generar una señal río abajo [36].

La estructura del dominio TIR del adaptador TRAM ha sido modelada en base a las estructuras de los receptores TLR1 y TLR2, encontrando que tiene un largo de 148 aminoácidos y que está formada por 4 hojas  $\beta$  rodeadas por 5  $\alpha$ -hélice [20]. La estructura recién descrita se observa en la representación tridimensional presentada en la figura 3.5, en la que se destaca el BB *loop* por medio del modelo atómico de sus residuos. El extremo N-terminal de TRAM sufre de miristoilación constitutiva, un evento requerido para su asociación con la membrana plasmática [40]. Por esto es funcionalmente equivalente al adaptador MAL, el que se une a fosfatidilinositol-4,5-bisfosfato para unir a la membrana.

Por su parte, el dominio TIR del adaptador TRIF, que cumple una función similar a la de MyD88, tiene en este caso una estructura más larga, en la que el dominio TIR está localizado en el medio de su secuencia. Además posee en su extremo N-terminal una región, denominada dominio *death* (DD) que le permite la activación río abajo del factor regulador IRF-3 [39].

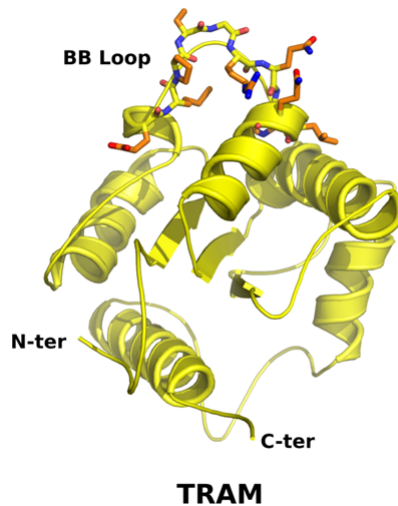


Figura 3.5: Estructura tridimensional del adaptador TRAM. Se observa la estructura terciaria del dominio TIR del adaptador TRAM. La estructura del lazo BB se resalta al mostrar sus residuos con una representación atómica. El receptor TRAM fue generado mediante el software MODELLER (Salgado 2011, Resultados no publicados).

### 3.2. Interacciones en la vía de señalización mediada por TLR4

La vía de señalización del sistema inmune innato en la que participa el TLR4, ha sido parcialmente caracterizada en una serie de publicaciones. En particular, el trabajo de Dunne *et al.* (2003) [23] estudió, experimentalmente y vía modelamiento, las interacciones entre heterodímeros TLR4/MyD88 y TLR4/MAL, partiendo de la base de que el dominio TIR es esencial para señalización y que éste se encuentra presente en los receptores TLR como también en las proteínas adaptadoras MAL y MyD88. El trabajo muestra que algunos residuos conservados de TIR están expuestos al solvente y pueden ser considerados como posibles puntos de interacción. Se plantea además que los dominios TIR purificados provenientes de MAL y MyD88 pueden formar heterodímeros estables con TLR4 y TLR2, sugiriendo que MAL y MyD88 unen a regiones de los receptores tipo toll. Esto esta respaldado por la descripción de las superficies de potenciales electrostáticos de las proteínas, la que sugiere que existe complementariedad de cargas en las superficies en contacto. Lo anterior se puede observar en la Figura 3.6, en la que se muestran representaciones para las superficies del potencial electrostático para los dominios TIR del receptor TLR4 y el adaptador MAL.



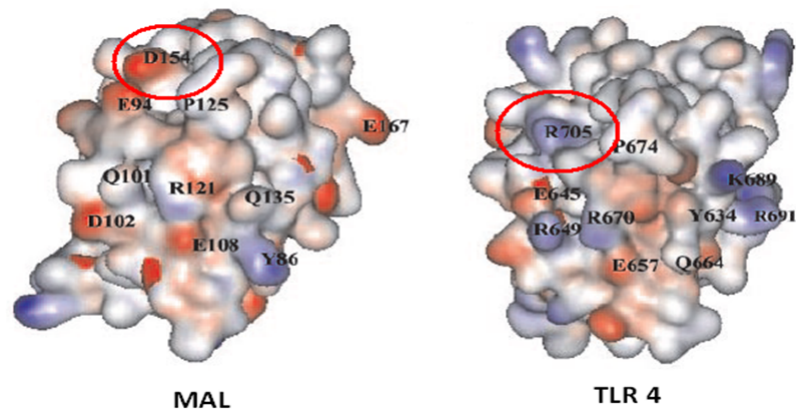


Figura 3.6: Complementariedad en superficies de potencial electrostático para interacciones TLR4 y MAL. (Adaptado de Dunne *et al.* 2003). La figura muestra representaciones de las superficies del potencial electrostático para los dominios TIR de las proteínas MAL y TLR4. El color rojo en las superficies indica que éstas son áreas de carga negativa y el color azul indica las zonas que tienen carga positiva. Los residuos que contribuyen a la carga de la superficie se nombran y numeran de acuerdo a las secuencias utilizadas. En un círculo rojo se han enmarcado algunas zonas que presentan complementariedad de cargas electrostáticas de acuerdo al trabajo desarrollado por Dunne *et al.* [23]. En ambos casos, las superficies están orientadas de modo de que la zona de los BB loop enfrenta al lector.

El estudio plantea, en base a datos obtenidos mediante el uso del algoritmo de docking GRAMM, que no existen sitios de unión coincidentes para TLR4/MyD88 y TLR4/MAL apoyando la tesis del rol cooperativo de los adaptadores en la cadena de señalización. Cabe destacar que en este estudio el algoritmo de *docking* es aplicado sobre para dar con los acoplamiento del dominio TIR de TLR4 (como monómero) con los dominios TIR de MAL y MyD88 por separado. Además, se señala que existen sitios discretos diferentes en el dominio TIR de TLR4 para la unión de MAL y MyD88 y un tercer modo unión para la formación de heterodímeros MAL/MyD88, haciendo hincapié en que ninguno de los sitios de interacción determinado involucra al *loop* BB que ha sido señalado en numerosas estudios como factor determinante de las interacciones. El trabajo describe que las interacciones entre TLR4 y MAL se generan respectivamente a partir del contacto de los *loops* DE y DD, mientras que las interacciones entre TLR4 y MyD88 se generan por el contacto de los *loops* AA y DD, pertenecientes a MyD88, al *loop* CD de TLR4. El estudio muestra además que se repiten las mismas interacciones en el caso de TLR2. Para ahondar más en el rol que tiene el *loop* BB en las interacciones, el trabajo de Dunne realiza una mutación sitio dirigida en el *loop* BB, cambiando el aminoácido prolina, de gran relevancia en la interacción [41], por una histidina, no observándose cambios relevantes en la interacción, por lo que Dunne respalda su hipótesis de la baja relevancia de este *loop* en las interacciones. Cabe destacar que si bien esta última fue una de las conclusiones de su trabajo, Dunne hace hincapié en que puede estar sesgada por el método rígido de docking utilizado [27]. Este último comentario justifica la necesidad de contrastar los resultados que pueden ofrecer los distintos algoritmos de docking.

La publicación de Núñez *et al.* (2007) brinda mas información relevante acerca de las interacciones en la vía TLR4. En particular, plantea una hipótesis diferente a la señalada por Dunne, dando cuenta de que los TLR4 se activan por un ligando que induce su dimerización, evento a partir del cual se provocarían cambios en la superficie del dominio TIR del homodímero permitiendo la interacción de otras proteínas adaptadoras. El trabajo de Núñez se dedica a analizar las interacciones del homodímero TLR4 con los adaptadores MAL y TRAM, pertenecientes a las dos principales vías en las que está involucrado el TLR4. A partir de los resultados del estudio se determina que si bien se ha modelado TLR4 como monómero, resulta más interesante el estudio de esta molécula como dímero ya que en general la gran mayoría de las conformaciones de docking determinadas con MAL y TRAM dan como resultado la interfase del homodímero. Más aún,

el trabajo recalca las uniones obtenidas para MAL y TRAM con la interfase del homodímero TLR4, dando cuenta que en ellas está involucrado fuertemente el elemento estructural del *loop* BB de cada una de las proteínas comprometidas. [20]

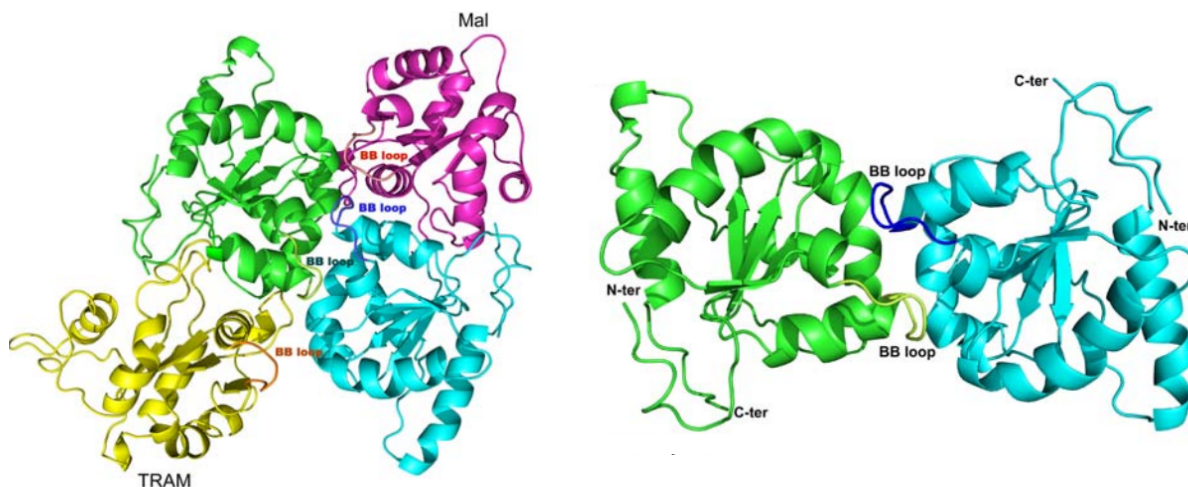


Figura 3.7: Resultados obtenidos mediante docking para interacción TLR4 (dímero)/MAL y TLR4 (dímero)/TRAM. (Núñez *et al.*, 2007). La figura de la izq. corresponde al dímero TLR4 (color verde y cian) ligado a los adaptadores MAL (violeta) y TRAM (amarillo), mientras que la figura de la der. muestra una vista superior del dímero TLR4 aislado.

Los resultados de Núñez van más allá, señalando que la interacción del homodímero TLR4 con MAL y TRAM, crea en ambos casos nuevas superficies de interacción que permiten la respectiva unión del nuevo complejo con MyD88 y TRIF, los siguientes adaptadores de la cadena. En este caso, el autor plantea además que la superficie de interacción del adaptador MAL y TRAM podría ser la misma, por lo cual las dos vías serían excluyentes. Es importante recalcar que en este caso el trabajo de Núñez aporta con una descripción detallada de los contactos TLR4-MAL y TLR4-TRAM, la que se aprecia en la siguiente tabla.

Tabla 3.1: Características de los sitios de interacción TLR4-MAL y TLR4-TRAM. (Adaptación Núñez *et al.* 2007)

Tipo de interacción	TLR4-MAL	TLR4-TRAM
Superficie escondida	TLR4 = 1.824 [ $\text{\AA}^2$ ]	TLR4 = 1.729 [ $\text{\AA}^2$ ]
	MAL = 1.761 [ $\text{\AA}^2$ ]	TRAM = 1.805 [ $\text{\AA}^2$ ]
Contactos	49	43
Interacción VdW Fuerte	TRP757(TLR4)-GLU167(MAL)	TRP757(TLR4)-ASN159(MAL)
Puente Salino	ARG763(TLR4)-GLU167(MAL)	GLU824(TLR4)-ARG231(MAL)
		GLU685(TLR4)-ARG119(MAL)
		ARG780(TLR4)-GLU75(MAL)
Enlaces de hidrógeno	7	3

Se definen contactos aminoácidos tal que  $\Delta\text{ASA} > 1 [\text{\AA}^2]$

Se definen como interacciones fuertes aquellas en que  $\Delta\text{ASA} > 40[\text{\AA}^2]$ .

La nomenclatura utilizada para la descripción de los contactos corresponde a la numeración de los aminoácidos de acuerdo a las secuencias que se muestran en el Anexo A.1

Además el trabajo mencionado, a través de la utilización del software FPSPD, predice la funcionalidad de una serie de sectores de las proteínas TLR4, MAL y TRAM (Tabla 3.2), complementando aún más

las teorías y resultados propuestos. Finalmente y a través de un amplio estudio de posibles mutaciones en la superficie del homodímero TLR4, Núñez logra determinar los efectos de una serie de mutaciones (Ver tabla 3.3) que van desde el impedimento de la homodimerización o la formación de dímeros inactivos, al impedimento de la unión de los adaptadores y hasta una posible pérdida de estabilidad en los complejo. El mismo estudio de mutaciones permite afirmar que al contrario de lo planteado por Dunne, se tiene que el *loop* BB y en particular el residuo de prolina que posee es fundamental en la interacción homodímero/adaptador.

Tabla 3.2: Caracterización funcional de residuos de TLR4, MAL y TRAM. (*Adaptación de Núñez et al. (2007)*)

Proteína	Residuos	Comentarios
TLR4	677-698	Participación en unión MAL
	705-730	Participación en homodimerización y unión TRAM
	734-739	Participación en homodimerización y unión MAL
	795-812	Posible sitio de unión MyD88
MAL	89-104	Sitio unión TLR4
	108-122	Sector cercano a la unión con MyD88
	121-139	Sitio unión TLR4
	146-169	Sitio unión TLR4
TRAM	79-94	Sitio unión TLR4
	117-127	Residuos que presentan unión con TLR4
	136-164	Sitio unión TLR4
	171-181	Potencial sitio de unión con TRIF
	218-226	Potencial sitio de unión con TRIF

La nomenclatura utilizada para la descripción de los contactos corresponde a la numeración de los aminoácidos de acuerdo a las secuencias que se muestran en el Anexo A.1

Tabla 3.3: Resumen del efecto de mutaciones en activación NF- $\kappa$ B e IFN- $\beta$  (*Adaptación Núñez et al. 2007*).

Mutación TLR4	Activación NF- $\kappa$ B	Activación IFN- $\beta$
E689K	S	S
C706S	RS	nd
P714H	NS	NS
G726C	NS	nd
H728D	RS	RS
Q743A	S	S
C747S	RS	nd
W757A	S	RS
K779D	S	S
C706S-C747S	NS	nd

S: Señalización normal, RS: Señalización reducida, NS: Sn señalización, nd: No se pudo determinar variaciones.

Otro de los estudios considerados relevantes corresponde al realizado por Ohnishi et al. (2009) [22]. Este se desarrolla a partir de las primicias de Kagan y Medzhitov (2006) quienes plantean que la molécula MyD88 sirve como adaptador en la señalización río abajo, mientras que MAL serviría como un adaptador que facilitaría el contacto entre MyD88 y TLR4. Ohnishi parte además del hecho de que a la fecha de publicación de su trabajo no se ha observado unión homodimérica de TLR4 en solución, y que además a partir de experimentos in vitro ha sido posible demostrar que el dominio TIR de MyD88 no une directamente con TLR4, mientras que el adaptador MAL si lo hace. En base a lo anterior es que Ohnishi et al., a partir de

ensayos con genes reporteros, da cuenta de que existen 3 sitios en la superficie del dominio TIR de MyD88 que son relevantes en la vía de señalización: el sitio I (Arg217/Box1), el sitio II (Arg196, Arg197/Box2-BB loop) y el sitio III (Arg288/Box3), donde los Box son motivos conservados dentro de los dominios TIR (ver Figura 3.8)

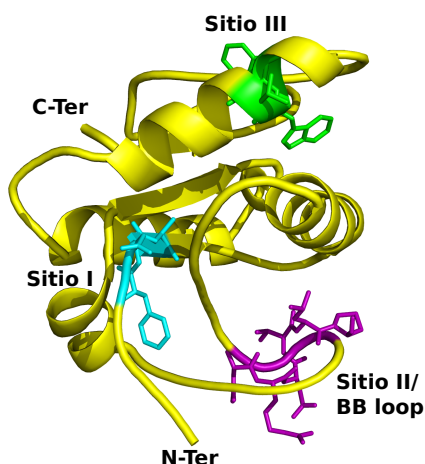


Figura 3.8: Sitios del dominio TIR de MyD88 relevantes para la interacción de acuerdo a Ohnishi *et al.* 2009. Se destacan en el dominio TIR de MyD88 los sitios: I (Arg217/Box1) [cian], II (Arg196, Arg197/Box2-BB loop) [púrpura] y el sitio III (Arg288/Box3) [verde].

Adicionalmente, se identifica que los sitios II y III, ubicados en extremos opuestos, cumplen un rol importante en la interacción directa con el dominio TIR de MAL, por lo que sería la proteína MAL, y no TLR4, la que une directamente con MyD88. En base a esto último es que este trabajo plantea un modelo de interacción en el que dos moléculas de MAL unen al adaptador MyD88, en el que además se propone que MAL une simultáneamente a TLR4 y MyD88, mediando la asociación.

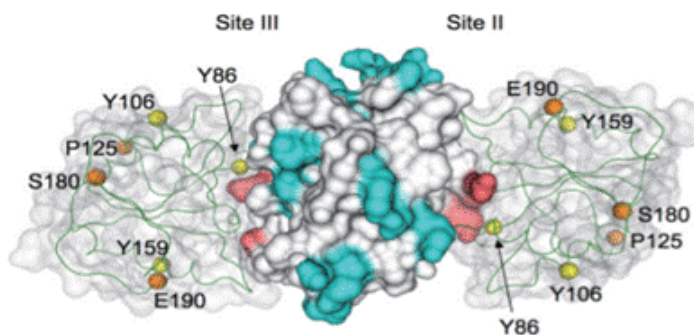


Figura 3.9: Modelo estructural del complejo de señalización formado por MyD88 y MAL. (Ohnishi *et al.*, 2009). La figura muestra el modelo de acoplamiento propuesto por Ohnishi para los dominios TIR de MAL y MyD88. El modelo muestra como dos moléculas de MAL se acoplan a los costados del dominio TIR de MyD88 (centro de la figura). Los sitios de unión II y III de MAL se muestran en rojo, mientras que los sitios de residuos no críticos para la señalización y el sitio I se muestran en celeste.

Existe otro modelo estructural descrito dentro de la revisión de antecedentes. El trabajo de Gong *et al.* (2010) [42] argumenta que el modelo planteado por Núñez *et al.* supone la dimerización del receptor TLR4, mientras que hasta el momento no se ha reportado evidencia directa de esto. El estudio plantea, a través del

modelamiento y el uso de software de *docking*, la formación de un tetrámero entre los dímeros de TLR4 y de MyD88, en la que el dímero TLR4 tiene una superficie cargada negativamente en la interface  $\alpha$ C (GLN755) de uno de los monómeros de TLR4, como también  $\alpha$ B (ALA719-HIS724) y  $\alpha$ C (TYR751 – THR756) del segundo monómero. Se plantea que la hélice  $\alpha$ E (CYS280-ARG288) altamente conservada de MyD88, con su carga positiva se acopla al bolsillo interactuando con los residuos de TLR4. Además, se detalla que la interacción es estabilizada por los siguientes acoplamientos: MyD88/DE *loop* (ILE271) – TLR4/CD *loop* (ARG763-ARG764); MyD88/EE *loop* (ASP275-THR277) – TLR4/CD *loop* (THR 756-GLN758); y MyD88/  $\alpha$ A (GLN181-ASN186) – TLR4/CD *loop* (TRP757-LEU760). Como defensa al modelo planteado, se considera el trabajo de Poltorak et al. (1998), en el que se reporta que la mutación P712H del dominio TIR de TLR4, termina con la respuesta de TLR4 a LPS, lo que concuerda con el modelo que postulan. Si bien estudios indican que MAL promueve el reclutamiento de MyD88 a TLR4, y que por tanto no habría interacción directa entre ambas proteínas, se ha mostrado que se puede dispensar del adaptador MAL cuando MyD88 está unido a un dominio PIP2 [43].

La importancia del BB *loop* en los contactos también es revisada por el trabajo de Toshchakov et al., [34], quien diseña un set de péptidos bloqueadores a partir de la secuencia del BB *loop* de los dominios TIR de las proteínas adaptadoras MyD88, TRIF, TRAM y MAL. A partir del uso de los péptidos de bloqueo diseñados, logra interrumpir la expresión, inducida por LPS, de genes mediada por TLR4 asociada con vías dependientes e independientes a MyD88. La hipótesis propuesta para el método de acción de los péptidos, es que estos secuestran y bloquean las proteínas por medio de su acoplamiento a los BB *loop* del receptor TLR4 o adaptadores, interrumpiendo las interacciones naturales entre el receptor y los adaptadores.

De manera adicional al BB *loop*, el trabajo de Jiang et al. (2006) [35] define otro sitio que resulta importante para la señalización mediada por TLR4. Mediante un estudio de mutagénesis, se determinó que la mutación del residuo V693N en el receptor TLR4 (equivalente a I179 en MyD88) abolió completamente la activación de NF- $\kappa$ B inducida por LPS. En cambio, cuando se revirtió la mutación a su estado nativo, se recuperó por completo la activación de NF- $\kappa$ B. Por lo que este residuo, ubicado en el medio de la A  $\alpha$ -hélice rodeado de residuos hidrofóbicos, tendría un rol crítico en la funcionalidad del dominio TIR de las proteínas TLR y adaptadores.

De los antecedentes aquí expuestos, se desprende que existen una serie de modelos que han buscado responder a como se desarrolla el mecanismo de señalización mediada por TLR4 y sus adaptadores dentro del sistema inmune innato. Si bien cada modelo puede responder al cómo se desarrollan los contactos, se opta por considerar el modelo en el cual el receptor TLR4 actúa como un dímero como punto inicial de la cascada de reclutamiento, ya que este es el modelo que actualmente está siendo considerado en publicaciones más recientes [21, 30, 31, 44]. Esto implica que los modelos de interacción que se van a establecer están basados en el concepto clave de que la homodimerización del receptor TLR4 es el primer paso dentro de la cascada de señalización en la vía pro-inflamatoria en la que participan los receptores estudiados [20, 21, 30, 31]. En vista de esto se descarta un modelo de interacción monomérica entre TLR4-MAL y MAL-MyD88 como el planteado en otros trabajos [23]. Por lo tanto todos los modelos parten de la base de la interacción dimérica del receptor TLR4.

Adicionalmente, dentro de los modelos basados en la homodimerización de TLR4, el modelo planteado por Nuñez et al. 2007 permite la interacción simultánea de las proteínas adaptadoras MAL y TRAM. De manera similar los modelos planteados por este trabajo permiten la interacción simultánea de los adaptadores aprovechando la simetría axial que provee el dímero de TLR4, similar a la del dímero de TLR10 [33]. Cabe destacar que al contrario de lo planteado por Nuñez, los modelos que plantea este trabajo consideran como base la unión exclusiva de MAL o TRAM a ambos costados del dímero de TLR4. Este supuesto se ha adoptado en vista de la evidencia que indica que las interacciones entre TLR4 y MAL ocurren principalmente en la superficie de la membrana plasmática, mientras que la interacción entre TLR4 y TRAM se produciría luego de la internalización en endosomas o fagosomas [45], por lo que resulta improbable la interacción de

simultánea de TRAM y MAL con el mismo receptor TLR4. En función de lo anterior los modelos propuestos en este caso, se acercan a los modelos generados por Basith *et al.* 2011 [21], considerando la formación de un tetrámero entre el homodímero de TLR4 y dos moléculas de MAL o TRAM según corresponda. Adicionalmente, el modelo propuesto para la interacción TLR4-MAL-MyD88 considera la formación de un hexámero, en el que la proteína adaptadora MyD88 se acopla en sectores simétricos a ambos extremos del tetrámero de TLR4-MAL. Este tipo de modelos complejos está en acuerdo con el concepto de Mydosoma que han estudiado y respaldado algunas publicaciones [44, 46, 47]. Este tipo de modelos plantea que en esta vía se generan interacciones altamente complejas que involucren un gran número de receptores como el caso de los complejos MyD88 y IRAK-4 que forman complejos en razones estequiométricas de 7:4 u 8:4 [44].

Sin perjuicio de lo anterior, debe mencionarse que a la fecha no se cuenta con evidencia de la formación de complejos del receptor TLR4 y sus adaptadores TIR desde componentes purificados. Esto sugiere que el proceso de acoplamiento de las proteínas es de alta complejidad, y que la cooperatividad y la acción de múltiples interacciones débiles o inestables actúan de manera conjunta para dar lugar a la unión transiente y estable de los complejos de señalización [44]. De la misma manera, a esta altura no ha sido posible determinar si la unión de MAL o TRAM al receptor TLR4 es excluyente. En vista de lo anterior, la pregunta acerca de si solo una molécula de TRAM o MAL puede unirse al receptor TLR4 impidiendo la acción de la otra, o si la unión de un adaptador genera un aumento en la afinidad en un segundo sitio quedan aún sin responder.

### 3.3. Dinámica Molecular

El estudio de las proteínas y otros sistemas moleculares biológicos puede realizarse mediante modelamiento, de manera de reducir la estructura de una macromolécula a un modelo de átomos esféricos conectados por resortes que representan los enlaces y funciones de potencial que den cuenta de la energía del sistema. La dinámica de este modelo puede estudiarse mediante la resolución numérica de las ecuaciones de movimiento de Newton para cada uno de los átomos, de modo de obtener las trayectorias de cada átomo en un periodo de tiempo que en la actualidad llega usualmente al orden de 100 [ns] [48, 49].

Mediante las simulaciones de dinámica molecular es posible caracterizar aspectos específicos de la estructura de proteínas, tal como las características de su cinética enzimática, estabilidad macromolecular, variaciones termodinámicas, etc. En escalas de tiempo, que no es posible manejar en el laboratorio, del orden de femtosegundos a nanosegundos [49].

En general, las simulaciones de dinámica molecular (MD, por sus siglas en inglés) tienen una estructura típica que involucra una serie de etapas como: preparación del sistema, minimización inicial de la energía del sistema, determinación de condiciones de simulación, adición de iones y solvatación del sistema, minimización del sistema solvatado, simulaciones de equilibrio y una simulación final de producción, a partir de la cual se obtienen los resultados a analizar [50]. La siguiente figura muestra un diagrama con las principales etapas de una simulación MD.

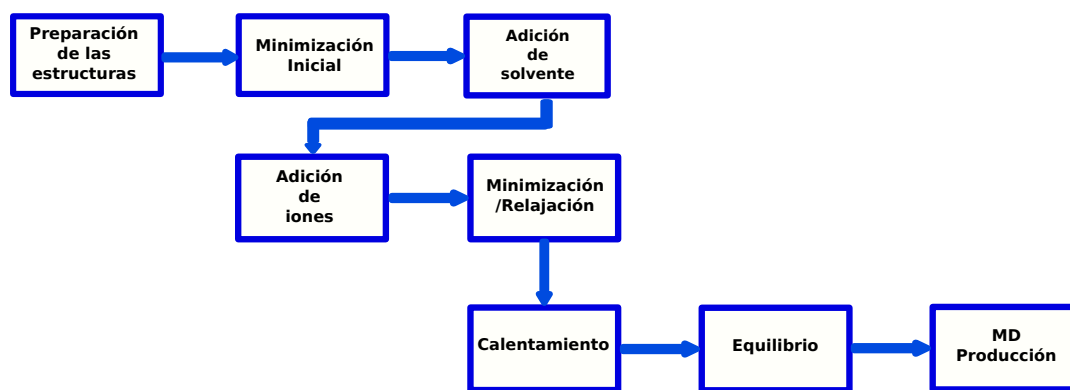


Figura 3.10: Diagrama de bloques de una simulación de MD.

Cada etapa de un protocolo de MD tiene claros objetivos definidos, así por ejemplo la primera etapa, denominada preparación inicial, busca preparar las proteínas a estudiar, identificando los átomos junto con definir sus cargas, posición inicial, velocidad inicial, etc.

La etapa de minimización inicial, y en general cualquier etapa de minimización, tiene por objetivo el corregir las fallas que puedan presentarse en una estructura molecular no refinada que pueda presentar sus ángulos de enlace o largos de enlace distorsionados. En el caso de un modelo de proteína, el problema se traduce a encontrar las coordenadas de los átomos “r” tal que se minimice una función de potencial  $V(r)$ . Los métodos más populares para resolver este tipo de problemas son aquellos derivativos de varios ordenes, incluyendo primer orden como método del gradiente y gradiente conjugado, además de métodos de segundo orden como el método de Newton Raphson [49]. Usualmente una etapa de minimización contempla unas 500 – 1000 pasos de minimización usando el método del gradiente o gradiente conjugado usando alguna función potencial basada en CHARMM, GROMACS o AMBER [51] [52].

La adición de moléculas de agua responde a tratar de generar una representación razonable del ambiente en que se encuentra las proteínas. Por esto es que usualmente se agregan moléculas de agua las que pueden ser representadas explícitamente usando alguno de los modelos como SPC, SPC/E, TIP3P, TIP4P, TIP5P entre otros, los que difieren en el tratamiento de los ángulos de la molécula del agua, difusividad, densidad, constante dieléctrica, entre otros [53]. Dado el costo computacional que involucra el simular explícitamente miles de moléculas de agua, es que a veces se opta por el tratamiento del solvente de manera implícita. En estos casos el solvente es representado por una función continua, esto reduce los tiempos de cálculo pero produce pérdida de precisión al calcular efectos electrostáticos entre el solvente y el soluto [49].

La etapa de adición de iones, tiene por objetivo obligar al sistema a tener una carga neta neutra ya que algunas proteínas pueden ser sensibles a la fuerza iónica del solvente. Además es común utilizar el método PME para el cálculo de energías electrostáticas, y este método requiere que el sistema sea eléctricamente neutro [54]. Por lo mismo es usual reemplazar moléculas de solvente por iones  $Cl^-$  y  $Na^+$  hasta neutralizar el sistema [50].

La etapa de calentamiento dentro del protocolo de MD tiene por objetivo el llevar al sistema a la temperatura de simulación, usualmente unos 298 - 300 K, de modo de establecer la correcta estructura de las moléculas del sistema a estas condiciones. Es usual además usar algún tipo de controlador de temperatura para evitar grandes desviaciones durante las siguientes etapas, los algoritmos de control más usados son los de Berendsen y Nose Hoover [55]. Posteriormente se utiliza una etapa de equilibrio, que busca equilibrar el solvente y los iones alrededor del sistema de manera de evitar distorsiones innecesarias del sistema. Durante esta etapa se realiza una simulación que busca que la energía del sistema converja [51, 52, 56].

Finalmente se lleva a cabo la simulación de producción, que busca ser lo suficientemente larga como

para obtener datos de la cinética, estructura y termodinámica del sistema, por lo cual el sistema no es restringido y se monitorea constantemente todas las variables [51–53].

### 3.4. Cálculo de la energía libre

La energía libre es considerada como la magnitud que refleja el grado de equilibrio de un sistema, y en este caso particular también la afinidad de un complejo proteico, en tanto consideramos que el grado de afinidad de un complejo está directamente relacionado con que la estructura sea termodinámicamente favorable y sea por tanto de mínima energía. Además el conocer el valor de la energía libre del sistema permite calcular explícitamente la constante de equilibrio del sistema según la ecuación [25, 27],

$$(3.1) \quad \Delta G_b^\circ = -k_b T \ln(C^\circ K_b)$$

Donde  $\Delta G_b$  es la energía de unión,  $k_b$ , es la constante de Boltzman,  $C^\circ$  es la concentración estándar y  $K_b$  es la constante de equilibrio.

Como función de estado, es posible considerar que para calcular el valor de la energía libre solo es necesario conocer los estados inicial y final del sistema, esta es la primicia que considera el método de MM-PBSA (Molecular Mechanics Poisson – Boltzmann Surface Area) para calcular esta magnitud. Esta técnica, revisada por Swanson et al. (2004) [57], se basa en el cálculo de la energía de un complejo como la suma de términos de entalpía y entropía, usando aproximaciones con solvente explícito para evaluar los conjuntos de equilibrio iniciales y finales generados usando simulaciones de dinámica molecular. La metodología MM-PBSA plantea considerar la unión de dos proteínas A (receptor) y B (ligando) como una reacción reversible que da lugar al complejo AB. Dentro de esta reacción es que se considera que cada molécula (A y B) posee movilidad traslacional, rotacional y a la vez posee modos internos de movimiento. Al unirse el ligando al receptor, este pierde su movilidad traslacional y rotacional pasando esos grados de libertad a convertirse en modos internos de movimiento para el complejo. Considerando lo anterior es que se calcula la energía libre de afinidad como [57, 58]:

$$(3.2) \quad \Delta G_{AB}^\circ = -RT \ln \left( \frac{C^\circ}{8\pi^2} \right) \left( \frac{Z_{N,AB} Z_{N,O}}{Z_{N,A} Z_{N,B}} \right) + P^\circ \langle \Delta V_{AB} \rangle$$

$\Delta G_{AB}^\circ$  es la energía libre de afinidad del complejo, R es la constante de los gases, T la temperatura,  $C^\circ$  la concentración molar estándar (1 molécula/1660 Å<sup>3</sup>),  $Z_{N,X}$  es la integral de configuración de la molécula X (O en el caso del solvente), y el término  $P^\circ \langle \Delta V_{AB} \rangle$  corresponde al trabajo asociado al cambio de tamaño del sistema al pasar de estar formado por dos moléculas libres a convertirse en un solo complejo. Por integral de configuración de una proteína se hace referencia a las expresiones [57, 58]:

$$(3.3) \quad Z_{N,A} = \int e^{\beta U(r_A, r_s)} dr_A dr_s$$

$$(3.4) \quad Z_{N,O} = \int e^{\beta U(r_s)} dr_s$$

En las que  $\beta = \frac{1}{k_b T}$ ,  $k_b$  es la constante de Boltzmann, U(r) la energía del sistema o proteína,  $r_A$  y  $r_s$  son las coordenadas de las moléculas de la proteína A y el solvente respectivamente. Posteriormente y en



base a una serie de aproximaciones, que detalla Swanson et al. (2004), en las que se considera los grados de libertad que representan los movimientos de traslación y rotación del ligando unido, se obtiene la siguiente expresión

$$(3.5) \quad \Delta G_{AB}^{\circ} = -RT \ln \left( \frac{C^{\circ} Z_B^{trans} Z_B^{rot}}{8\pi^2} \right) + (\langle E_{AB} \rangle - \langle E_A \rangle - \langle E_B \rangle)$$

En la que las expresiones  $\langle E_X \rangle = \langle U(r_X) + W(r_X) \rangle$  representan el promedio de la energía a lo largo de una simulación de dinámica molecular para la proteína X, donde  $U(r_X)$  es la energía mecánica del sistema y  $W(r_X)$  es la energía de solvatación. El término de solvatación es descompuesto en dos términos, un primer término que considera los aportes no polares [57]

$$(3.6) \quad \Delta W_{SA}^{np} = \gamma(SASA) + \beta$$

Donde el término SASA corresponde el área de superficie accesible por el solvente, y los valores  $\beta$  y  $\gamma$  son determinados experimentalmente. Un segundo término brinda la contribución polar a la solvatación, factor que es calculado a partir de la ecuación de Poisson-Boltzmann [59].

$$(3.7) \quad \nabla \epsilon(r) \nabla \phi(r) + 4\pi \rho(r) = 0$$

En el que  $\epsilon(r)$  representa la constante dieléctrica,  $\phi(r)$  el potencial electrostático y  $\rho(r)$  la distribución de la carga eléctrica.

Por su parte el término logarítmico representa los cambios en la entropía del sistema en base a  $Z_B^{trans}$  y  $Z_B^{rot}$  que son el movimiento traslacional y rotacional del ligando. En la práctica estos términos se calculan a partir del uso de un modelo quasi armónico, basado en la suposición de que tanto el movimiento traslacional como rotacional pueden ser descritos mediante un función de probabilidad gaussiana de multivariable. La idea general consiste en superponer una serie de capturas instantáneas de la proteína receptora, en particular de los carbonos-alfa, de manera de evaluar la matriz de covarianza de la posición del centro de masa del ligando, cuyos valores propios son utilizados para calcular  $Z_B^{trans}$  de acuerdo a la siguiente ecuación (análoga para  $Z_B^{rot}$ ) [57, 58]

$$(3.8) \quad Z_B^{trans} = \int e^{-\frac{\kappa_1 \Delta x_1^2}{2k_b T}} dx_1 \int e^{-\frac{\kappa_2 \Delta x_2^2}{2k_b T}} dx_2 \int e^{-\frac{\kappa_3 \Delta x_3^2}{2k_b T}} dx_3 \\ = (2\pi)^{3/2} (\langle \Delta x_1^2 \rangle \langle \Delta x_2^2 \rangle \langle \Delta x_3^2 \rangle)^{1/2}$$

Se ha estudiado las características (precisión, desviación estándar) y la dependencia de los resultados de este método en función del método de simulación utilizado (Weis et al., 2006). Se ha determinado que pueden ser utilizados varios tipos de simulaciones MD considerando diversas formas de delimitar el espacio conformacional (Cubo, Esfera, Esfera centrada en ligando, solvente implícito) obteniéndose resultados similares aunque con diferencias en el tiempo de cálculo y costo operacional. Lo esencial en todos los casos resulta ser el representar el solvente explícitamente en las simulaciones. Otras consideraciones que se tienen es que pueden ser usados distintos campos de fuerzas en el cálculo de energía, entregando resultados

similares siempre y cuando exista coherencia entre el campo de fuerzas utilizado en la simulación y el cálculo de energías; en el caso del cálculo de la energía de solvatación se muestra que alternativamente a la resolución de la ecuación de Poisson-Boltzmann (PB) es posible utilizar la ecuación generalizada de Born, lo que si bien entrega resultados mucho más rápido, estos resultan ser mucho peores que los obtenidos por PB. Además, el trabajo de Weis et al. da cuenta de que el efecto que tienen las cargas de los aminoácidos en el cálculo de la energía es pequeño, ya que al neutralizar las cargas del complejo los resultados varían en no más de 3 [kJ/mol], lo que indica adicionalmente que las cargas que están lejos del sitio de unión tienen poca relevancia en la determinación de la energía libre. Por último se plantea que las simulaciones deben sobrepasar los 450 [ps] para que los resultados obtenidos sean estables [60].

En general, el principal problema del método MM-PBSA es la alta desviación estándar que presenta el cálculo de la entropía. A modo de ejemplo, los resultados de un estudio del sistema de avidina mostraron que la desviación estándar del término de entropía es altamente superior a la de los otros términos. En este caso particular, el efecto de la entropía produjo finalmente una desviación estándar del 47–62 [kJ/mol] en la energía libre de unión del sistema de avidina, lo que es considerable dado que la estimación de energía libre arrojó un valor de 83 [kJ/mol] [61]. En vista de estos casos, se han propuesto alternativas para el cálculo de la entropía como la metodología propuesta por Kongsted et al. (2008) que disminuye la desviación estándar del término de la entropía en un factor de 2 a 4, permitiendo que deje de ser el término que limita la precisión del método. Esto último se logra incorporando una región buffer en la simulación de dinámica molecular, utilizada por el método MM-PBSA, de modo de impedir que las estructuras del complejo capturadas difieran una de otra de manera exagerada [61].

## Capítulo 4

# Metodología

### 4.1. Materiales y equipos utilizados

Los materiales y equipos utilizados para realizar el trabajo descrito se describen a continuación.

#### 4.1.1. Proteínas y adaptadores

Se contó con las estructuras de las siguientes proteínas y adaptadores proteicos: MyD88 (2JS7, código PDB) [62], homodímero TLR4, MAL y TRAM (Tabla 4.1). Estas 3 últimas estructuras fueron obtenidas mediante homología y refinadas en MODELLER [63], un software que permite comparar estructuras proteicas y generar modelos estables. Finalmente, las estructuras fueron refinadas mediante dinámica molecular utilizando el software NAMD. El detalle de este procedimiento no se incluye en este informe, ya que forma parte del trabajo previamente realizado por Salgado (Resultados no publicados) en el contexto del proyecto FONDECYT titulado “Mathematical modeling of the interaction between cell-penetrating TIR BB loop decoy peptides (BBPs) and toll-like receptor 4 and 2 (TLR4 and TLR2)”. Las secuencias de aminoácidos

de cada una de las estructuras se muestran en el Anexo A.1

Tabla 4.1: Resumen de estructuras proteicas utilizadas en el proyecto

Proteína	Características	N° Residuos	Origen	Referencia
TLR4	Dominio TIR-Estructura	326	Modeller	Salgado*
MAL	Dominio TIR-Estructura	150	Modeller	Salgado*
MyD88	Dominio TIR-Estructura	149	NMR	Rossi <i>et al.</i> * Salgado**
TRAM	Dominio TIR-Estructura	171	Modeller	Salgado*
TLR4-MAL	Estructura Complejo	476	<i>Docking</i>	Núñez (2007) [20]
TLR4-TRAM	Estructura Complejo	497	<i>Docking</i>	Núñez (2007) [20]

\*:Resultados no publicados

\*\* :Estructura fue publicada por Rossi, pero refinada por Salgado (Resultados no publicados)

La calidad estructural de los modelos utilizados fue estimada a partir del uso de las herramientas disponibles en el servidor SAVES: PROCHECK, WHAT CHECK, ERRAT, VERIFY 3D y PROVE [64–68]. Se prestó particular atención a las diferencias que mostraron los indicadores estimados para los modelos propuestos por Salgado y Núñez, de manera de evaluar la validez de los modelos generados a partir de cada estructura. De la misma manera se evaluó los modelos usando las herramientas ProQ y ModFOLD [69, 70], de manera de poder comparar la calidad de los modelos utilizados con los presentados en el trabajo de Basith *et al.* (2011)

#### 4.1.2. Herramientas computacionales

El proyecto contó con el siguiente equipamiento para su desarrollo:

Tabla 4.2: Herramientas computacionales usadas en el proyecto

Equipo	Tipo	Descripción
Clúster de producción <i>Eowyn</i>	Hardware	80 Núcleos, CPU Quad Core AMD Opteron <sup>TM</sup> , 16 x 10 GBi RAM, 1 TBi de espacio en disco RAID1, conectado a través de una red de 1 Gbs
Ordenador	Hardware	Procesador AMD Athlon <sup>TM</sup> 64 X2 Dual Core 4600+, 4 GBi RAM
Linux Centos OS 5.5 [71]	Software	Sistema Operativo implementado en Clúster de producción
Linux Ubuntu OS 11.04 [72]	Software	Sistema Operativo implementado en Ordenador
Condor 7.4.4 [73]	Software	Sistema de gestión de recursos para cálculo en paralelo implementado en Clúster de producción

La importancia de contar con estos recursos yace en la reducción del tiempo promedio que requieren los cálculos realizados en los algoritmos de *docking*, el refinamiento mediante simulaciones de dinámica molecular y el cálculo de la energía libre mediante la metodología MM-PBSA. A modo de ejemplo es posible citar los resultados preliminares del proyecto FONDECYT en que se enmarca este trabajo. En ellos se muestra que el uso del clúster, con el objeto de paralelizar los cálculos en una simulación de dinámica molecular, permite reducir a un 20,8 % de lo que toma un cálculo habitual.

A lo largo del proyecto se utilizó además el siguiente software especializado para la visualización de las moléculas, resultados de acoplamiento proteína-proteína, simulaciones de dinámica molecular y cálculos de energía libre (Tabla 4.3).

Tabla 4.3: Software especializado utilizado para el desarrollo del proyecto

Software	Versión	Propósito	Sitio Web	Referencia
GRAMM	1.0.3	<i>Docking</i>	<a href="http://vakser.bioinformatics.ku.edu">http://vakser.bioinformatics.ku.edu</a>	Vakser <i>et al.</i> (1991) [74]
HEX	6.3	<i>Docking</i>	<a href="http://www.hex.loria.fr">http://www.hex.loria.fr</a>	Ritchie (2006) [75]
DOCK	6.4	<i>Docking</i>	<a href="http://dock.compbio.ucsf.edu">http://dock.compbio.ucsf.edu</a>	Kuntz <i>et al.</i> (1982) [76]
ZDOCK	3.0.2	<i>Docking</i>	<a href="http://zdock.bu.edu">http://zdock.bu.edu</a>	Chen, Weng (2002) [77]
pyDOCK	Servidor web	<i>Docking</i>	<a href="http://life.bsc.es/servlet/pydock/">http://life.bsc.es/servlet/pydock/</a>	Fernández-Recio <i>et al.</i> (2007) [78]
Pymol	1.3	Visualización	<a href="http://pymol.org">http://pymol.org</a>	Schrodinger, LLC [79]
Chimera	1.5.3	Visualización	<a href="http://www.cgl.ucsf.edu/chimera/">http://www.cgl.ucsf.edu/chimera/</a>	Ferrin <i>et al.</i> (2004) [80]
VMD	1.9.1	Visualización/Análisis	<a href="http://www.ks.uiuc.edu/Research/vmd/">http://www.ks.uiuc.edu/Research/vmd/</a>	Humphrey <i>et al.</i> (1996) [81]
GROMACS	4.5.5	Dinámica Molecular	<a href="http://www.gromacs.org">http://www.gromacs.org</a>	Berendsen <i>et al.</i> (1995) [55]

## 4.2. Procedimiento

A continuación se muestra la metodología utilizada durante el proyecto, detallándose los protocolos y parámetros usados en cada uno de los algoritmos utilizados. Previo a la presentación en detalle de la metodología, se muestra en la figura 4.1 un diagrama de bloques que resume las distintas etapas involucradas en el trabajo.

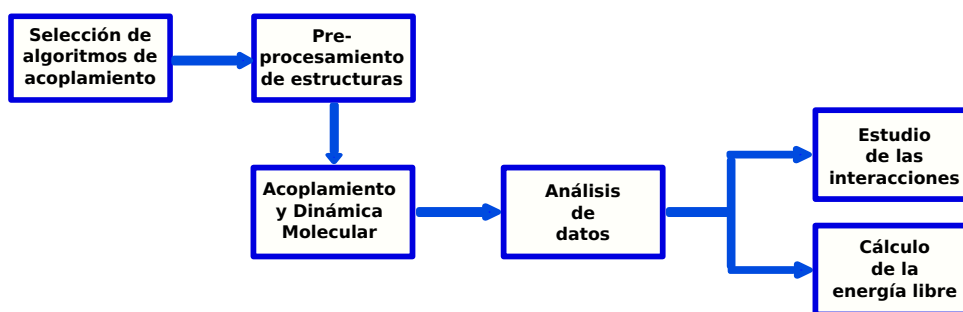


Figura 4.1: Diagrama de bloques para la metodología planteada para el proyecto

### 4.2.1. Selección de los algoritmos de *docking*

La primera etapa de la metodología considera el estudio de distintos algoritmos de acoplamiento o *docking* de modo de identificar las principales ventajas y desventajas de cada software. En base a una extensa búsqueda bibliográfica se determinó algunas de las principales características de cada uno de los algoritmos, en particular: tipo de búsqueda (guiada o sistemática), tipo de representación de las proteínas (atómica, granos, discreta o superficial), tipo de licencia del software (libre, libre académica o comercial), tipo de plataforma de implementación del algoritmo (descargable, servidor web, web mail), año de la última versión, disponibilidad del soporte técnico y referencia original como respaldo al algoritmo. El anexo A.2 muestra el software estudiado junto con sus principales características. De manera adicional es importante destacar la disponibilidad de soporte técnico, ya que es un indicador del status de actualización del software, y de si se cuenta con algún respaldo para la pronta resolución de problemas de ejecución.

Como indicadores adicionales se determinó el número de veces que ha sido citado cada algoritmo en otras publicaciones, los tipos de aplicaciones para los que se ha usado algoritmo y tiempos de ejecución promedio reportados en el uso de los algoritmos. Esta información permite establecer cuál es el software más utilizado y cuales cuentan con un mayor respaldo de la comunidad científica. Por otro lado, el conocer

los tiempos promedio de ejecución de los algoritmos facilita la planificación inicial del proyecto. Las Tablas A.3, A.4 y A.5 del Anexo A.2 muestran la información comentada.

Es importante mencionar que la selección de algoritmos se limitó a aquellos utilizables en la plataforma LINUX. Este plataforma se utilizó a lo largo del proyecto por considerarse que cumplía con los estándares de calidad, estabilidad y seguridad requeridos, dadas las largas simulaciones y series de cálculo características de los métodos de *docking*, dinámica molecular y cálculo de la energía libre.

En base a la revisión bibliográfica y al estudio particular de aquellos algoritmos de *docking* que resultaron más interesantes, se determinó finalmente estudiar el siguiente software de acoplamiento: GRAMM, HEX, DOCK, zDOCK, pyMOL, Vina y Autodock. Cabe destacar que cada uno de los algoritmos utilizados realiza *docking* rígido, con excepción del software DOCK (que posee un módulo flexible de *docking*), Vina y Autodock. Todos estos algoritmos utilizan una representación discreta de las proteínas, describiendo de acuerdo a una grilla cada uno de los átomos que componen las proteínas lo que brinda una mayor exactitud en los contactos generados. Los algoritmos GRAMM y DOCK utilizan un sistema de búsqueda sistemático explorando exhaustivamente el espacio geométrico definido para la predicción de complejos. Por su parte, los algoritmos Autodock, Hex, Zdock, pyMOL y Vina, utilizan un sistema de búsqueda guiada. Esto implica que acotan la cantidad de resultados generados en cada iteración, mediante la minimización de energía de los complejos obtenidos en cada caso. Cabe destacar que cada uno de los algoritmos escogidos ha sido citado extensamente en publicaciones, lo que se considera como un indicador del impacto que generaron en el rubro.

#### **4.2.2. Preprocesamiento y refinamiento de estructuras**

Las estructuras proteicas contempladas para este estudio fueron pretratadas de acuerdo a lo requerido por cada uno de los algoritmos de acoplamiento. Dependiendo del algoritmo de *docking*, es necesario determinar las posiciones de los hidrógenos en la estructura tridimensional de cada proteína o eliminar residuos de aminoácidos no reconocidos.

En la mayoría de los casos, la eliminación o adición de hidrógenos se llevo a cabo usando las herramientas estándar del software CHIMERA. En particular, para el uso del software DOCK se utilizó la herramienta Dock Prep [82] presente en CHIMERA. Esta opción permite eliminar moléculas de solvente presentes, añade hidrógenos para la determinación de los estados de protonación de las proteínas, asigna cargas a los átomos en residuos aminoacídicos estándar y escribe las moléculas en formato Mol2 requerido por el algoritmo DOCK.

En ocasiones se realiza simulaciones de dinámica molecular mediante software como AMBER, NAMD o GROMACS para preparar las estructuras y remediar anomalías que puedan presentar los modelos de las proteínas, como pueden ser la superposición de átomos o plegamientos inadecuados. En el caso de este trabajo, en vista de la inspección de los modelos a utilizar no se consideró necesario *a priori* llevar a cabo esta etapa.

#### **4.2.3. Docking y Dinámica Molecular**

##### **4.2.3.1. Aplicación de algoritmos de *docking***

Una vez pretratadas las estructuras tridimensionales de las proteínas, se utilizó los modelos para determinar los sitios de acoplamiento y estudiar las interacciones que se generan en la cascada de señalización mediada por TLR4 mediante el uso de los algoritmos de *docking* seleccionados. Para cada uno

de los algoritmos se predijo las estructuras de los complejos TLR4-MAL, TLR4-TRAM y posteriormente TLR4-MAL-MyD88. De acuerdo a lo planteado en antecedentes generales, los modelos de interacción planteados están basados en el concepto clave de que la homodimerización del receptor TLR4 es el primer paso dentro de la cascada de señalización en la que participan los receptores estudiados [20,21,30,31,44]. En vista de esto, todos los modelos parten de la base de la interacción dimérica del receptor TLR4. Esto implica que, para los efectos de la aplicación de los algoritmos, la estructura homodimérica de TLR4 se utiliza como receptor y por tanto el espacio de búsqueda utilizado por cada algoritmo es definido a partir de ella. Por su parte, los adaptadores son considerados como el ligando por lo que su estructura es la que es rotada o trasladada alrededor del receptor en el espacio de búsqueda según cada algoritmo. Finalmente, el complejo TLR4-MAL-MyD88 fue construido, de manera separada, a partir del complejo TLR4-MAL presentado por Nuñez y el determinado de manera independiente en este trabajo. En este último caso, la molécula de TLR4-MAL la que es considerada como el receptor por los algoritmos, mientras que la molécula de MyD88 es designada como el ligando.

Se debe mencionar que el uso de los distintos tipos de algoritmos implica per se la determinación de una serie de parámetros para la ejecución de cada uno de los algoritmos. Previo al uso de cada software se llevó a cabo una serie de pruebas, usando la proteína TLR4 y sus adaptadores, de modo de determinar parámetros para las simulaciones definitivas.

Para poder definir un modelo revelante en cada caso se generaron al menos 100 predicciones por modelo, las que se obtuvieron a partir de varias ejecuciones en la aplicación de cada algoritmo. Durante la producción de los modelos se utilizó principalmente dos estrategias. La primera de estas considera el uso de una metodología de *docking* ciego. En este caso se explora la totalidad del espacio conformacional sobre el receptor TLR4 de modo de determinar los contactos posibles de cada adaptador sobre el receptor. Adicionalmente se utilizó una segunda estrategia de acoplamiento sitio dirigido de modo de acotar el espacio de búsqueda. Este tipo de búsqueda, basado en la búsqueda sistemática de posibles sitios de acoplamiento en la superficie, simplifica el problema de optimización y a la vez permite concentrarse en sectores que han sido definidos como relevantes biológicamente por publicaciones anteriores. Para el acoplamiento sitio dirigido se diseñó una grilla de 26 cuadrantes alrededor del receptor TLR4, de modo de abarcar la totalidad de la estructura según se muestra en el Anexo A.3. En cada uno de los cuadrantes se realizó *docking* utilizando los software Vina y Autodock mediante los cuales es posible demarcar el espacio de búsqueda. Cabe destacar que el software mencionado posee una función empírica para el cálculo de la energía libre de unión, la que fue utilizada para determinar cuáles sitios de la grilla tienden a mostrar complejos de mayor afinidad.

La selección de un modelo representativo en cada caso se lleva a cabo utilizando una serie de reglas o criterios para guiar la toma de decisiones. Dentro de los criterios considerados se distinguen:

- Ausencia de interacciones en los extremos terminales del dominio TIR de TLR4, ya que tales extremos están conectados al dominio transmembrana lo que impediría físicamente cualquier interacción (ver Figura 4.2).
- Presencia en el modelo de interacciones relacionadas a elementos estructurales conservados.
- Existencia de evidencia biológica experimental publicada que valide el modelo propuesto. [20–23, 34, 35, 42, 46, 83–85]
- Ranking del modelo dentro de los resultados del algoritmo de búsqueda.

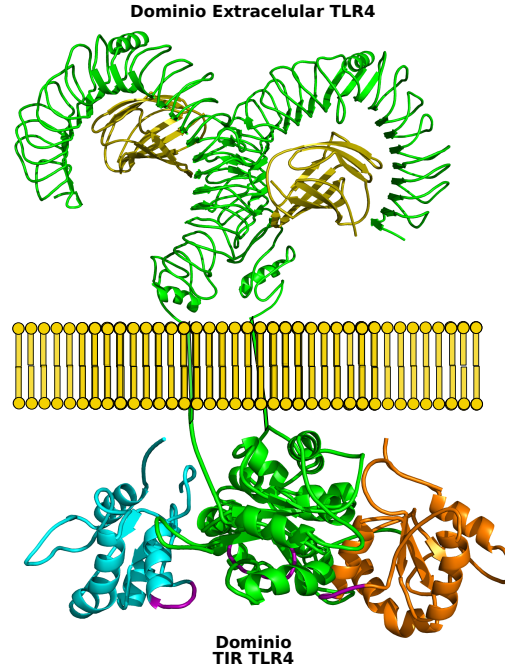


Figura 4.2: Estructura del homodímero receptor TLR4. Cada uno de los receptores TLR4 que forma parte del homodímero cuenta con 3 dominios: extracelular, transmembrana e intracelular (dominio TIR). Los extremos N- y C-terminal del dominio TIR del receptor TLR4 están unidos a su dominio transmembrana.

Finalmente, a partir de los complejos seleccionados se procede a determinar algunos de los aspectos más relevantes de la interacción tales como: tamaño de la superficie de interacción, aminoácidos en contacto, nivel de hidrofobicidad, interacciones con el solvente, etc. Lo que se logra mediante el uso de alguna herramienta de visualización de moléculas como CHIMERA o Pymol.

El tamaño de la superficie de interacción fue estimado por medio de un script implementado en pymol. Esta herramienta permite calcular el área escondida por cada proteína producto de la interacción, mediante el cálculo del área superficial del complejo y las áreas superficiales del receptor y ligando. Finalmente, el área escondida por los contactos se calcula según la siguiente ecuación.

$$(4.1) \quad A_{Escondida} = A_{Superficie,Complejo} - (A_{Superficie,Receptor} + A_{Superficie,Ligando})$$

La hidrofobicidad superficial promedio tanto de los adaptadores como de los complejos formados fue calculada de acuerdo a la metodología planteada en los trabajos de Berggren *et al.* (2002) y Salgado *et al.* (2005) [86, 87]. El cálculo se realiza asumiendo que cada aminoácido en la proteína contribuye, proporcionalmente a su abundancia, a las propiedades asociadas en la superficie de la proteína [87]. De esta manera, la hidrofobicidad superficial promedio se calcula de acuerdo a la siguiente ecuación:

$$(4.2) \quad \langle \Phi_{superficie} \rangle = \sum_{i \in A} \frac{S_i}{\sum_{j \in A} S_j} \Phi_i$$

donde  $\langle \Phi_{superficie} \rangle$  corresponde a la hidrofobicidad superficial promedio de la proteína estudiada, A es el conjunto de los 20 aminoácidos,  $S_i$  es la suma del área accesible (ASA) de todos los aminoácidos

de tipo  $i$  (calculada mediante NACCESS) [88], y  $\Phi_i$  es la hidrofobicidad del aminoácido  $i$ . Este último valor fue asignado de acuerdo a la escala de Kyte-Doolittle [89]. La escala propuesta por Kite-Doolittle esta basada en una serie de observaciones experimentales derivadas de la literatura. En ella se asigna un número a cada aminoácido para representar las propiedades hidrofílicas o hidrofóbicas de su cadena lateral, de modo que a mayor índice, mayor hidrofobicidad tiene el aminoácido. Dentro de la escala los aminoácidos más hidrofóbicos son Isoleucina (4.5) y Valina (4.2), mientras que los más hidrofílicos son la Arginina (-4.5) y Lisina (-3.9). La hidrofobicidad de los aminoácidos es una característica importante a la hora de estudiar la estructura de una proteína; los aminoácidos hidrofóbicos tienden a ubicarse en sectores internos, mientras que los aminoácidos hidrofílicos se ubican con mayor frecuencia en los sectores superficiales de la proteína [89].

Se puede obtener una estimación de los residuos que están participando en las interacciones, y/o forman parte del área de interacción, mediante el cálculo de la diferencia entre el área accesible de los residuos de una proteína asociada en un complejo y la proteína aislada de acuerdo a la ecuación:

$$(4.3) \quad \Delta ASA = ASA_{aislada} - ASA_{complejo}$$

donde el ASA fue calculado usando el software NACCESS. Esta definición permite describir de manera rápida aquellos aminoácidos que están en la superficie de interacción que son aquellos que muestran un  $\Delta ASA > 1$  [ $\text{\AA}^2$ ] [90]. El trabajo de Nuñez *et al.* (2007) considera además la definición de interacciones fuertes, como aquellas formadas por residuos que muestran un  $\Delta ASA > 40$  [ $\text{\AA}^2$ ]. Al utilizar esta definición es necesario dar cuenta que ella solo considera el cambio absoluto en el ASA de los residuos, por esta razón subestima (sobrestima) los cambios que se presentan en aquellos residuos que originalmente muestran valores pequeños (grandes) de ASA. De manera de resolver este inconveniente, se modificó la definición para que considerara la magnitud original del valor de ASA de los residuos. En la práctica esto se tradujo a calcular la variación relativa del ASA de cada residuo, de acuerdo a la siguiente definición:

$$(4.4) \quad \Delta ASA_{relativa} = \frac{ASA_{aislada} - ASA_{complejo}}{ASA_{aislada}} * 100$$

de acuerdo a esta nueva definición se definió a los residuos que están dentro del área de interacción como aquellos que presentan un  $\Delta ASA_{relativa} > 50$  %.

#### 4.2.3.2. Uso de dinámica molecular sobre los complejos obtenidos vía acomplamiento

De acuerdo a los resultados obtenidos en la sección anteriores se consideró, de acuerdo a los criterios descritos, la mejor de las estructuras producidas para cada complejo estudiado. Las estructuras seleccionadas fueron sometidas a un protocolo de dinámica molecular con solvente explícito de manera de permitir la movilidad de la cadena principal y las laterales, dando espacio para que ocurran cambios conformacionales a nivel local en la superficie de interacción.

El protocolo utilizado para la dinámica molecular se explicita en la Figura 4.3 y fue implementado en GROMACS v4.5.4 compilado en precisión doble. Los parámetros utilizados en cada etapa y una descripción de su uso se describen en el Anexo A.4. Este software se caracteriza por contar con una serie de módulos ejecutables que facilitan la preparación, ejecución y análisis de los datos generados. Además esto permite el efectivo seguimiento por etapas del protocolo de dinámica molecular, de modo que se puede identificar claramente el efecto que tienen los parámetros escogidos sobre la dinámica molecular del sistema estudiado. El software cuenta con amplio soporte técnico y cuenta con una licencia GNU. Adicionalmente el software cuenta con módulos paralelizables vía el uso de comunicación estándar MPI. Este software ya ha sido utilizado para el estudio de dinámica molecular de complejos proteicos y para el cálculo de energía libre,



esto finalmente permitió validar los protocolos de dinámica molecular y la metodología MM-PBSA planteada [56].

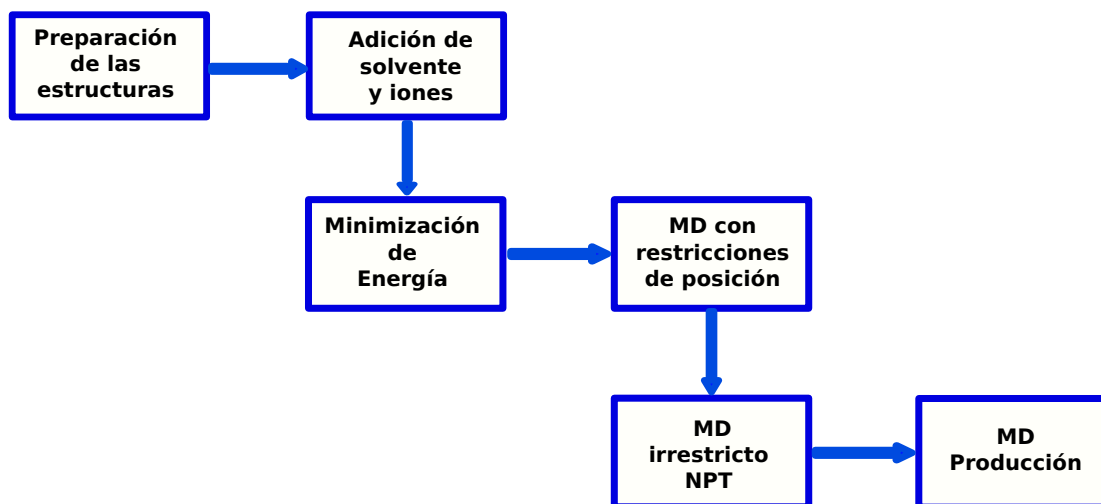


Figura 4.3: Diagrama de bloques con etapas del protocolo MD utilizado. Los protocolos de dinámica molecular (MD por sus siglas en inglés) cuentan con una serie de etapas. En primer lugar se prepara el sistema definiendo la topología del sistema y limitando el espacio de simulación. Posteriormente se procede a solvatar el sistema mediante el uso de un modelo explícito para las moléculas de agua. Una vez solvatado el sistema, se agregan iones de modo de neutralizar la carga total del sistema. El sistema luego para por una etapa de minimización de energía para eliminar interacciones desfavorables, para luego dar espacio al desarrollo de una etapa de MD con restricciones de posición. Después se procede a liberar la temperatura y presión del sistema bajo una simulación NPT ( $N^\circ$  moles,  $T$  y  $P$  constantes). Finalmente se desarrolla la simulación de producción que corresponde a aquella de la cuál se extraerán los datos para su posterior análisis.

La primera etapa del protocolo de dinámica molecular consiste en la preparación del sistema. El punto de partida de la dinámica molecular corresponde a un archivo PDB que define las coordenadas de los átomos que definen la estructura de una proteína. A partir de las coordenadas del átomo y por medio de la utilización del paquete *pdb2gmx* de GROMACS, se construye la topología del sistema. Esto corresponde a la descripción del sistema en términos de los tipos de átomos, cargas, enlaces, torsiones, etc. La definición de la topología del sistema está ligada al uso de un campo de fuerza el que parametriza cada uno de los componentes de las estructuras de proteínas. De manera adicional, durante la preparación de la topología del sistema se definen los estados de protonación de cada uno de los residuos del sistema mediante la opción *-ignh*. De este modo GROMACS define los estados de protonación, que serán considerados durante la simulación, una única vez. La definición considera un pH constante de 7.0, y se lleva a cabo mediante un proceso de optimización de los puentes de hidrógeno que presenta la estructura [55]. Esto implica que GROMACS ignora las variaciones en el pKa de los residuos que pueden producirse por efecto de la proximidad de otros residuos. Con el objetivo de verificar que la aproximación de los estados de protonación realizada por GROMACS fuera correcta, se corroboró los estados de protonación por medio del servidor especializado H++ [92].

La exactitud de la función de potencial es crucial para la validación y la estabilidad de las simulaciones de sistemas proteicos. La mayor parte de los campos de fuerza comparte una serie de similitudes describiendo el largo y el ángulo de los enlaces atómicos mediante términos armónicos, usando series de Fourier para describir las torsiones, y usando funciones de Lennard-Jones y Coulomb para la descripción de las interacciones por pares de átomos. En el caso de este trabajo se utilizó el campo de fuerza AMBER99SB.

La elección de este campo de fuerza responde a una serie de consideraciones. En primer lugar, los

campos de fuerza de AMBER son ampliamente utilizados en el desarrollo de dinámica molecular [49]. En particular, dentro de los campos de fuerza de AMBER, el campo de fuerza AMBER99SB es fruto del mejoramiento de los términos de dihedros  $\Phi/\Psi$  en la función de energía de ff99. Los parámetros actualizados en ff99SB de estos términos muestran un mejor balance de los elementos de la estructura secundaria de la proteínas según lo observado en un set de estructuras PDB y demuestran la corrección en el cómputo de energía interna [93]. Adicionalmente se observó que este campo de fuerza a sido utilizado en sistemas similares [60, 94, 95]. Otros campos de fuerza disponibles son las funciones GROMOS y OPLS. El uso de los campos GROMOS no fue considerado dado que las versiones del campo de fuerza se encuentran desactualizadas en comparación a los campos de fuerza de AMBER que han sido constantemente modificados. Por su parte el campo de fuerza OPLS, si bien esta basado en AMBER, fue desarrollado específicamente para la descripción de la dinámica molecular de soluciones orgánicas y otros líquidos [96].

Una segunda parte de la preparación de las estructuras corresponde a la definición del espacio en el cuál se desarrollará la simulación. Como en la mayor parte de las simulaciones de dinámica molecular se definieron condiciones periódicas de borde. esto implica la definición de una celda unitaria la que se replica en las tres dimensiones de modo de evitar la presencia de efectos de borde. En este caso se definió como unidad un dodecahedro rómbico, ya que permite ocupar un volumen de alrededor de un 71 % de un cubo. Esto reduce notablemente el número de moléculas de agua que son añadidas en el proceso de solvatación, permitiendo disminuir el tiempo de cálculo [55]. Las condiciones periódicas de borde fueron definidas mediante el paquete *editconf* de GROMACS considerando una distancia mínima de 1 [nm] entre la proteína y el borde de la celda, esto evita que se produzcan interacciones entre las imágenes periódicas.

Una vez preparado el sistema, se procede a solvatar el sistema. Existen una serie de modelos distintos para representar la molécula de agua. En este caso se utilizó el modelo de agua TIP3P, el que es ampliamente utilizado en el estudio de dinámica molecular con solvente explícito al usar el campo de fuerza de AMBER. Otros modelos populares y utilizados en conjunto con las funciones de energía de AMBER son TIP4P y TIP5P. La principal diferencia que existe entre estos modelos, es que el modelo TIP3P solo considera los átomos que constituyen la molécula de agua. Por su parte los modelos TIP4P y TIP5P consideran sitios adicionales sin masa que representan la deslocalización de las cargas del átomo de oxígeno. En vista de que no se requiere estudiar en detalle la localización de las cargas de los átomos de hidrógeno, se decidió simplificar el modelo de agua y utilizar el modelo TIP3P, mediante el cuál es posible describir de manera general el efecto de la solvatación. En la práctica la solvatación del sistema se lleva a cabo utilizando la herramienta *genbox* que ubica las moléculas de agua en un espacio cercano a la proteína utilizando un archivo que agrega regiones cúbicas de 216 moléculas en una región cúbica hasta completar el número de moléculas que requiere el sistema.

Luego de la solvatación del sistema es necesario agregar iones al sistema de modo de alcanzar un sistema con carga neta neutra. Esta adición de iones es una etapa comúnmente utilizada durante los protocolos de dinámica molecular con condiciones periódicas de borde. Esto se debe a que usualmente la energía de interacción electrostática se calcula mediante PME (*Particle-Mesh Ewald*), metodología que requiere que el sistema presente carga neta neutra. La adición de iones se lleva a cabo utilizando la herramienta *genion* y los archivos de estructura y topología previamente generados. Se especificó una concentración de 0.1 M de iones de  $\text{Na}^+$  y  $\text{Cl}^-$  [51, 52], además se utilizó la instrucción *-neutral* para indicar la adición en exceso de uno de las dos especie de modo de neutralizar el sistema. Finalmente, se debe considerar que los iones son localizados en reemplazo de otras moléculas del sistema, por lo que se seleccionó reemplazar las moléculas de solvente de acuerdo a lo especificado en el manual de GROMACS.

Una vez definido el sistema de la simulación se procede a desarrollar la etapa de minimización de energía. Esta etapa busca eliminar las interacciones desfavorables que puede haber generado la adición de solvente e iones al sistema, de modo que se lleva a cabo la relajación controlada del sistema. La minimización de energía se desarrolla por medio de los paquetes *grompp* y *mdrun* de GROMACS. El protocolo de

minimización considera 1000 pasos del método del gradiente conjugado, además cada 100 pasos se realiza una iteración utilizando el método del gradiente. El número de pasos fue definido de modo de lograr el objetivo de esta etapa evitando llegar a sobreminimizar el sistema, lo que puede llevar a que la estructura alcance una estructura "inmovilizada" que no presente un comportamiento natural. La tolerancia considerada para el método del gradiente o gradiente conjugado es de  $10.0 \text{ [kJ mol}^{-1} \text{ nm}^{-1}]$ . El componente electrostático del sistema es calculada mediante PME-Switch utilizando una distancia de corte de 1.2 [nm]. Esto implica que se calculan las interacciones electrostáticas usando el método *Fast Particle-Mesh Ewald* acoplado a una función Switch lo que permite el decaimiento suave de la función electrostática sin generar discontinuidades en la función de energía, por lo que esta orientado a evitar artefactos asociados a los metodos cut-off convencionales. Como parámetros para PME-Switch se consideró un tamaño de grilla para FFT de 0.12 [nm], interpolación de orden 4 y una tolerancia de  $10^{-5}$  [52]. Para el cálculo de la energía de no enlace se utilizó la opción *Shift* de GROMACS, que corresponde a una función de potencial de tipo Lennard-Jones modificada de modo de que al alcanzar la distancia de corte no se produzcan discontinuidades. Se definió una distancia de corte para las interacciones de Van der Waals de 1.2 [nm].

Luego de la etapa de minimización de energía se desarrolla una etapa de dinámica molecular con restricciones de posición. El objetivo de esta etapa es permitir que el solvente se mueva libremente alrededor de la estructura de la proteína, cuyos átomos (de cadena principal) son fijados alrededor de una posición de referencia. Esto implica que los átomos pueden moverse, pero existe una gran penalización por hacerlo. El uso de restricciones para la posición de los átomos se implementa mediante el uso de la opción -DPOSRES de GROMACS [55]. En la práctica las restricciones se establecen en el archivo .itp para los átomos distintos de hidrógeno. La simulación tuvo una duración de 50 [ps] y consideró un paso de integración de 1 [fs]. Para el cálculo de las componentes electrostáticas y no enlace se utilizó respectivamente PME-Switch y Shift con una distancia de corte de 1.2 [nm]. Durante la simulación la temperatura promedio del sistema es mantenida mediante el uso de un termostato. Se utiliza el algoritmo V-rescale con una temperatura de 200 [K] y una constante  $\tau_T = 0.1$  [51,52]. Por su parte, la presión del sistema se mantiene mediante el algoritmo de Berendsen con una presión de referencia de 1 [bar] y un  $\tau_P = 1$  [ps] [51,52]. Durante esta etapa, y por ser la primera etapa de dinámica molecular se asignaron al azar velocidades a cada uno de los átomos utilizando una distribución de Maxwell a una temperatura de 200 [K]. Durante la simulación se extrajeron datos para la trayectoria y energía del sistema cada 1 [ps].

Una vez que las moléculas de agua y los iones fueron acomodados alrededor del sistema de estudio, se procedió a liberar la temperatura y presión del sistema hasta alcanzar la condiciones en que se llevará a cabo la simulación de producción. La simulación en condiciones NPT ( $N^\circ$  moles constante, Presión constante, Temperatura constante) se realizó durante 100 [ps] con un paso de 2 [fs]. Se utilizó condiciones de borde periódicas y un corte de 1.2 [nm] para las interacciones de Coulomb y Van der Waals. La temperatura se mantuvo mediante el algoritmo V-rescale, con una temperatura de referencia de 300 [K] y un  $\tau_T$  de 0,1 [ps]. Cabe destacar que es en este punto donde se produce el calentamiento del sistema desde una temperatura de 200 a 300 [K]. El cambio de temperatura se produce de manera que la desviación sigue una curva exponencial que decae de acuerdo a la constante de tiempo  $\tau_T$ . Por otro lado, el control de la presión se mantiene mediante el algoritmo de Berendsen, con una presión de referencia de 1 [bar] y un  $\tau_P = 1$  [ps]. Al contrario de la simulación anterior en este casos no se generaron nuevas condiciones iniciales para la simulación sino que se utilizaron los datos de salida de la etapa anterior. La extracción de datos se realizó cada 1 [ps]. Cabe destacar que esta etapa busca, además de estabilizar la temperatura y presión del sistema, idealmente estabilizar las variaciones del sistema. Esta estabilidad es medida a través del RMSD de los C- $\alpha$  del sistema que debiese mostrar convergencia en valores menores a 0.3 [nm].

Una vez acondicionado el sistema se prosiguió con la etapa de producción que corresponde a aquella a partir de la cuál se extraen los datos finales tanto para el análisis de los contactos como para el cálculo de la energía libre. La simulación de producción se llevó a cabo durante 6 [ns] con un paso de tiempo de 2 [fs], guardando las coordenadas del sistema y variables energéticas cada 20 [ps]. Se utilizó condiciones de borde

periódicas y un corte de 1.2 [nm] para las interacciones de Coulomb y Van der Waals. Para el control de la temperatura se usa el algoritmo de V-rescale, con una temperatura de referencia de 300 [K] y un  $\tau_T$  de 0,1 [ps]. Por su parte el control de la presión se realiza mediante el algoritmo de Berendsen, con una presión de referencia de 1 [bar] y un  $\tau_P$  de 0,5 [ps]. Al igual que en el caso anterior, se utilizó los datos de salida de la etapa anterior como punto de partida para esta etapa. Durante la simulación se usa el algoritmo LINCS para restringir el largo de todos los enlaces.

Una vez desarrollada la etapa de producción de la dinámica molecular, y como etapa previa al análisis de simulaciones, se requiere eliminar los efectos de las condiciones periódicas de borde en el sistema. Esto tiene por fin eliminar una serie de artefactos que se observan al visualizar la simulación (moléculas no centradas, moléculas rotas, efectos de difusión fuera del área de simulación, etc) y al realizar cálculos de RMSD. Para eliminar estos efectos se siguió el protocolo sugerido por el manual de GROMACS y que utiliza el paquete *trjconv* del software.

Finalmente, a partir de las trayectoria de los complejos es posible estudiar las interacciones que se producen entre receptor y ligando. Las interacciones son definidas a partir de criterios geométricos que consideran la distancia entre los átomos de los residuos según se dispone en la Tabla 4.4. Los criterios recién mencionados fueron implementados en VMD mediante la modificación de otros scripts similares (Patiño 2012, Resultados no publicados). Las interacciones que fueron contempladas dentro del análisis son:

- Cation- $\pi$ : Corresponden a interacciones que se pueden producir entre un átomo ionizable con carga positiva y un anillo aromático. Se producen cuando la distancia entre el catión y el centroide de un anillo aromático es menor o igual a 6 [Å] [97].
- Aromáticas: Interacciones que se producen por el apilamiento de grupo aromáticos. Las interacciones entre anillos son consideradas aromáticas si la distancia entre los centroides es menor o igual a 4.5 [Å] [98].
- Puentes de Hidrógeno: Interacciones entre átomos clasificados como aceptores y donadores de puentes de hidrógeno si la distancia inter-atómica es menor o igual a 3.5 [Å] y el ángulo entre el átomo de hidrógeno, el átomo dador y aceptor presenta una desviación menor  $\pm 30^\circ$  del ángulo ideal [99].
- Hidrofóbicas: Corresponden a interacciones entre residuos clasificados como hidrofóbicos. Se consideran aquellas interacciones en que los átomos de los residuos tienen una distancia menor o igual a 5 [Å] [89].
- Iónicas (Puentes Salinos): Se definen como aquellas interacciones entre residuos que son clasificados como iónicos. Se consideran como contactos los que se producen entre átomos ionizables positiva o negativamente (independiente de su estado de protonación) si su distancia es menor a 4 [Å] [100, 101].

Tabla 4.4: Criterios para la definición de interacciones durante Dinámica Molecular

Tipo de interacción	Distancia [Å]	Ángulo [°]	Referencia
Puentes Salinos	4	—	Barlow <i>et al.</i> (1983), Donald <i>et al.</i> (2011) [100, 101]
Puentes de Hidrógeno	3.5	$\pm 30^\circ$ <sup>1</sup>	Torshin <i>et al.</i> (2002) [99]
Hidrofóbica	5	—	Kyte <i>et al.</i> (1982) [89]
Cation- $\pi$	6	—	Dougherty <i>et al.</i> [97]
Aromática	4.5	—	Chakrabarti <i>et al.</i> (2007) [98]

El ángulo corresponde a la desviación angular desde un Puente de Hidrógeno perfecto que tiene un ángulo H-Donor-Aceptor de valor 0 [°]

Adicionalmente se utilizan algunas herramientas complementarias presentes en GROMACS de modo de establecer las variaciones en el RMSD C- $\alpha$  del sistema (mediante *g\_rms* o *g\_rmsf*), estudiar las variables de estado (T,P), analizar las componentes energéticas del sistema (vía *g\_energy*) y chequear la calidad de la simulación (por medio de *g\_gmxcheck*).

#### 4.2.4. Cálculo de energía libre

Una vez determinados y refinados los complejos que se obtienen como resultado de los algoritmos de *docking*, se procede a calcular la energía libre de cada complejo mediante la técnica de MM-PBSA. El protocolo MM-PBSA permite calcular la energía libre entre dos estados termodinámicos. Idealmente el valor de la energía libre podría ser estimado directamente a partir de un sistema solvatado. El problema radica en que en estos sistemas, la mayor parte de las contribuciones a la energía proviene de las interacciones solvente-solvente por lo que las fluctuaciones de la energía serían de una magnitud mayor a la energía de unión haciendo que los cálculos tomen un largo tiempo en converger. Por lo mismo es que se utiliza el ciclo termodinámico que muestra la figura 4.4 de manera de evitar el cálculo directo de la energía libre.

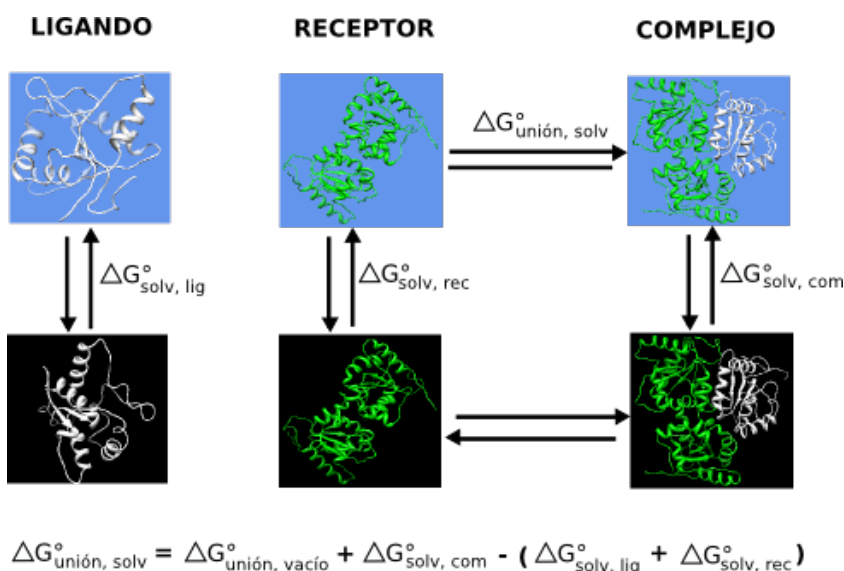


Figura 4.4: Ciclo termodinámico utilizado en el cálculo de energía libre vía MM-PBSA

La estimación de energía libre de unión a través del método MM-PBSA (*Molecular mechanics - Poisson Boltzmann Surface Area*) incluye términos asociados a la mecánica molecular ( $E_{MM}$ ) tales como la energía interna ( $E_{int}$ ), energía electrostática ( $E_{elec}$ ) y energía de Van der Waals ( $E_{vdW}$ ). Junto con esto, se calcula la variación de energía libre asociada a la solvatación, la cual tiene una componente polar ( $\Delta G_{solv, polar}$ ) y otra apolar ( $\Delta G_{solv, apolar}$ ). Por último, se incorpora la entropía del sistema ( $-T\Delta S$ ). A partir de estos términos se estima la energía libre de acuerdo a las siguientes ecuaciones.

$$(4.5) \quad \Delta G_{\text{unión, vacío}}^{\circ} = \langle E_{MM} \rangle - TS$$

$$(4.6) \quad \Delta G_{\text{solv, com}}^{\circ} = \langle G_{\text{polar, solv}} \rangle + \langle G_{\text{nonpolar, solv}} \rangle$$

$$(4.7) \quad \Delta E_{MM}^{\circ} = \langle E_{int} \rangle + \langle E_{\text{electrostatics}} \rangle + \langle E_{vdW} \rangle$$

Los paréntesis angulares  $\langle \rangle$  indican que la magnitud se calcula como un promedio a lo largo de la trayectoria MD específica del complejo, receptor o ligando según sea el caso. [55, 102] En la práctica se utilizarán las estructuras obtenidas en 300 frames provenientes de los 6 nanosegundos de la simulación de producción realizada en cada caso.

De los términos mencionados se tiene que  $E_{MM}$  corresponde a la energía interna del sistema. Como tal, está asociada a la energía involucrada en los enlaces covalentes y los ángulos que dan forma a la estructura proteica (planos, dihedros propios y dihedros impropios). El detalle acerca de cómo calcula tales valores puede encontrarse en el manual de GROMACS versión 4.5.5.

La energía electrostática ( $E_{elec}$ ) es calculada de manera separada considerando por un lado las interacciones de corto alcance (SR por sus siglas en inglés) y por otro las de largo alcance (LR por sus siglas en inglés). Durante la simulación se utilizó el método PME para calcular las interacciones de Coulomb. En el caso de las interacciones de largo alcance denominadas como Coul.-Recip en GROMACS, se considera para su cálculo un sistema homogéneo en el cual no es posible desglosar el aporte energético de cada grupo de átomos (solvente, iones, proteína) sobre las interacciones de largo alcance. En vista de esto, y de manera de calcular exclusivamente el aporte de las proteína-proteína en las interacciones de largo alcance, se extrajo la trayectoria de las proteínas desde el sistema solvatado y se recalculó las energías y fuerzas de la trayectoria de producción. De manera similar, el cálculo de la energía de las interacciones de Van der Waals ( $E_{VDW}$ ) considera la diferenciación entre la componente de corto y largo alcance de acuerdo a un radio de corte denominado  $r_{VDW}$ . Los términos  $E_{elec}$ ,  $E_{VDW}$  y  $E_{int}$  se calculan mediante el paquete *g energy* de GROMACS.

Por otro lado, la componente polar de la variación de energía libre de solvatación es estimada por medio del uso del software APBS. Dentro del archivo de configuración de APBS, se definen varios parámetros entre los cuales destaca el tamaño de la molécula de solvente y las constantes dieléctricas del solvente y el soluto. El primero define la superficie del soluto al recorrer toda su estructura, mientras que los últimos corresponden a la capacidad que tienen para polarizarse bajo la acción de un campo magnético y, por lo tanto, de conducir electricidad. En forma estándar, se considera que el vacío posee una constante dieléctrica de 1.0 y que el agua tiene un valor de 78.54. Si bien no es correcta la suposición de que una proteína pueda ser representada como un medio continuo como es el caso del vacío y del agua, se suelen utilizar valores entre 2.0 y 4.0 para calcular la contribución electrostática a la solvatación. En este caso se considera un valor de 2.0 como constante dieléctrica de las proteínas. El cálculo de APBS requiere la definición de una grilla de puntos que discretice el sistema, se designó un tamaño de grilla de 0.5 Å una solución de 100 mM NaCl para la resolución numérica de la ecuación no lineal de PB. Cabe destacar que APBS recibe como archivo de entrada no un archivo *.pdb* o *.gro*, si no un archivo *.pqr* que contiene además de las coordenadas atómicas contiene las cargas y radios atómicos. Los archivos *.pqr* fueron generados a partir de VMD. Posteriormente se utilizó el script *psize.py* para definir la grilla óptima que requiere el tamaño del sistema. Por su parte, el término  $\langle G_{polar,solv} \rangle$  se determina calculando la SASA mediante la aplicación del software GROMACS, el que implementa el método de lattice en una malla de 1.4 [Å], usando  $\gamma = 2.2$  [kJ mol<sup>-1</sup> nm<sup>-2</sup>] y  $\beta = 3.84$  [kJ mol<sup>-1</sup>] [56].

Finalmente la entropía se calcula mediante análisis de modos normales (NMA por sus siglas en inglés). Para esto, se obtuvieron las conformaciones del complejo, el receptor y ligando desde el último cuadro de cada trayectoria independiente. Las tres estructuras son sometidas a una rigurosa minimización de energía hasta que la fuerza máxima aplicada sobre un átomo del sistema alcance una magnitud menor a 10<sup>-6</sup> [kJ mol<sup>-1</sup> nm<sup>-1</sup>]. Una vez alcanzado tal grado de minimización, se procede a desarrollar la simulación para el análisis de modos normales. Luego se obtiene la matriz Hessiana a partir de la simulación de modos normales. Posteriormente, se utilizó los valores propios obtenidos para calcular la entropía vibratoria del sistema de acuerdo a la siguiente ecuación [56]:

$$(4.8) \quad S_{vib} = \sum_{i \in A} -R \ln(1 - \exp^{-h\nu_i/kT}) + \frac{N_A h \nu_i \exp^{-h\nu_i/kT}}{T(1 - \exp^{-h\nu_i/kT})}$$

donde  $S_{vib}$  es la entropía vibracional,  $h$  es la constante reducida de Planck,  $v_i$  son los valores propios del modo normal,  $k$  es la constante de Boltzmann,  $T$  es la temperatura absoluta del sistema y  $N_A$  es el número de Avogadro. Cabe destacar que los primeros 6 valores propios se descartan por representar efectos rotacionales y traslacionales. Todos los cálculos fueron realizados en GROMACS compilado con doble precisión.

r

## Capítulo 5

# Resultados

### 5.1. Verificación de la calidad estructural de los modelos utilizados

La calidad de los modelos estructurales utilizados (Nuñez *et al.* (2007) y Salgado (Resultados no publicados)), entendida como la coherencia fisicoquímica y biológica entre la secuencia de aminoácidos y la estructura terciaria de una proteína, fue estimada con las herramientas disponibles en el servidor SAVES: PROCHECK, What Check, Verify 3D, Errat y Prove (Ver Tabla 5.1).

Proteína	Evaluación Modelo Salgado	Evaluación Modelo Nuñez
TLR4	96.9% C, 3.1% P, 0.0% PH, 0.0% NP <sup>1</sup> 94.8% V3D <sup>2</sup> 61.1% Errat <sup>3</sup>	93.9% C, 5.7% P, 0.3% PH, 0.0% NP 100% V3D 71.3% Errat
MAL	93.8% C, 6.2% P, 0.0% PH, 0.0% NP 82.7% V3D 95.3% Errat	90.6% C, 9.4% P, 0.0% PH, 0.0% NP 74.8% V3D 72.1% Errat
TRAM	96.2% C, 3.8% P, 0.0% PH, 0.0% NP 71.1% V3D 92.1% Errat	88.5% C, 7.1% P, 3.8% PH, 0.6% NP 50.6% V3D 30.3% Errat
MyD88	95.4% C, 4.6% P, 0.0% PH, 0.0% NP 100% V3D 94.9% Errat	

<sup>1</sup> Porcentaje de residuos según su ubicación en las distintas zonas del gráfico de Ramachandran. C: Core, P: Permitidos, PH: Permitidos con holgura, NP: No permitidos

<sup>2</sup> Verify 3D: Porcentaje de residuos con un score promedio superior a 0.2. El score mide la coherencia entre la secuencia de aminoácidos y la estructura terciaria de la proteína [66].

<sup>3</sup> Errat: Factor de calidad general. El factor es resultado de la comparación estadística de las interacciones de no enlace entre los diferentes tipos de átomos y los resultados que presentan estructuras altamente refinadas, cuyos errores han sido calibrados para dar altos niveles de confianza [67].

Tabla 5.1: Evaluación de los modelos estructurales de TLR4, MAL, TRAM y MyD88

Los datos indican que el modelo para TLR4 producido por Nuñez muestra mejores indicadores, bajo los tests Verify 3D y Errat, que los generados por Salgado. Sin embargo, bajo el test PROCHECK el modelo para TLR4 de Salgado muestra un mejor porcentaje de residuos ubicados en las zonas más favorables (CORE) del gráfico de Ramachandran. En particular el residuo HIS144 del modelo de TLR4 de Nuñez presenta problemas según el gráfico de Ramachandran encontrándose en una zona límite del gráfico muy cercana a zonas prohibitivas. Este residuo pertenece a la hélice  $\alpha E$  de la cadena A de TLR4, en el trabajo de Nuñez no se le atribuye interacción alguna, pero es mencionado por Basith como un residuo que interactuaría con el adaptador MyD88 en el complejo TLR4-MAL-MyD88. Además, a partir de los test de PROCHECK es posible dar cuenta que, para el modelo TLR4 de Salgado, los residuos SER26, SER104, GLU97 presentan scores menores a -3.0 lo que puede indicar conformaciones no favorables. Por su parte los residuos GLU64, ARG84, GLY104, GLU227, ARG247 y ALA266 presentan altas desviaciones en el ángulo  $\Omega$  de torsión lo que también es un indicador de conformaciones no óptimas. En el caso del modelo de Nuñez, los residuos GLU97, SER101, HIS144 y SER263 presentan scores menores a -3.0, mientras que los residuos GLU64, ARG84, GLY104, LEU154, ARG247 y ALA266 presentan altas desviaciones del ángulo promedio de torsión  $\Omega$ . De lo anterior se desprende que algunas de las imperfecciones de los modelos están presentes en ambos modelos, lo que resalta su similitud. Por otro lado los indicadores de los test Verify 3D y Errat muestran que bajo tales pruebas, el modelo de Nuñez es levemente mejor al generado por Salgado.

En el caso de los modelos para el adaptador MAL, los indicadores bajo el test PROCHECK muestran un leve aumento del porcentaje de residuos presentes en los sectores favorables del gráfico de Ramachandran. El mismo test permite identificar algunos sectores que presentan conformaciones poco favorables, como los residuos GLY97, PRO104, TYR114 y GLU117 del modelo de Salgado cuyo score es menor a -3.0. Además los residuos ILE18, PHE42, TYR48, VAL55, GLU64, PHE80, THR95, ALA103, GLY109 presentan una desviación mayor a la promedio en ángulo  $\Omega$ . Por su parte los residuos ALA98, ALA112, GLU94, GLY108, TYR122 del modelo de Nuñez también presentan un score menor a -3.0, mientras que los residuos GLY38, GLU82 presentan desviaciones angulares para  $\Omega$  mayores al promedio de los ángulos. Por otro lado, los indicadores de los test Verify 3D y Errat muestran que la calidad del modelo de Salgado es superior al utilizado por Nuñez.

Los resultados del test PROCHECK para los modelos de TRAM arrojan que el modelo de Salgado presenta porcentajes superiores de residuos en áreas altamente favorables del gráfico de Ramachandran. Por su parte el modelo de Nuñez presenta 6 residuos en áreas poco favorables (ASN34, ASN114, GLU119, GLU132, GLU143, GLN155) y 1 residuo (TYR103) en áreas prohibidas del gráfico. En base a este indicador, el modelo de Salgado parece ser ampliamente superior. Adicionalmente se observa que los residuos GLU2, ASP82 y PRO122 poseen un score menor a -3.0 en el modelo de Salgado, mientras que lo mismo ocurre en los residuos PHE89, SER92, SER96, VAL97, HIS101 y PRO122 del modelo de Nuñez. Además, en el caso del modelo de Salgado se tienen cerca de 15 residuos con desviaciones angulares mayores al promedio del ángulo observado, mientras que en el caso del modelo de Nuñez son cerca de 6 los residuos que presentan ese comportamiento. En este caso particular los indicadores entregados por los test Verify 3D y Errat para el modelo de Salgado son al menos un 25 % superiores a los que muestran el modelo de Nuñez, lo que confirma la calidad del primero sobre el último.

Por otro lado, en el caso de los modelos MAL y TRAM los indicadores proporcionados por los test PROCHECK, Verify 3D y Errat muestran que los modelos propuestos por Salgado son mejor evaluados que los de Nuñez. En particular el modelo de TRAM presentado por Salgado elimina la presencia de residuos de las zonas no permitidas del gráfico de Ramachandran. Finalmente, a partir de los indicadores presentados se confirma la calidad estructural del modelo de MyD88.

Las diferencias estructurales que existen entre los modelos de Salgado y Nuñez también pueden ser apreciadas a partir de la inspección visual y la superposición de las estructuras de los diferentes modelos (Ver Figura 5.1). A partir de la figura se observa los modelos del dominio TIR de TLR4 y MAL presentados



por Salgado y Nuñez no presentan grandes diferencias en su estructura secundaria, aunque sí se observa que presentan discrepancias en la orientación espacial de sus cadenas laterales. Distinto es el caso de los modelos del dominio TIR de TRAM en los que se observan algunas discrepancias, que van desde variaciones en el largo de algunos segmentos  $\alpha$ -hélice hasta la aparición de un segmento  $\alpha$ -hélice en el modelo de Nuñez que abarca desde MET4 a GLU7 (Ver Figura 5.1). De manera más precisa se puede cuantificar las diferencias existentes entre las estructuras de los modelos, propuestos por Nuñez y Salgado, por medio del cálculo de RMSD. Estos resultados, que pueden apreciarse en el Anexo A.5, confirman las observaciones sobre los modelos del dominio TIR de TLR4 mostrando que efectivamente no existen diferencias importantes entre las estructuras planteadas por ambos autores. Por el contrario, los resultados de RMSD de los modelos de los dominios TIR de MAL y TRAM, muestran que en estos casos, si existen diferencias en las estructuras de los modelos de Nuñez y Salgado, y esta es de magnitud similar en ambos casos.

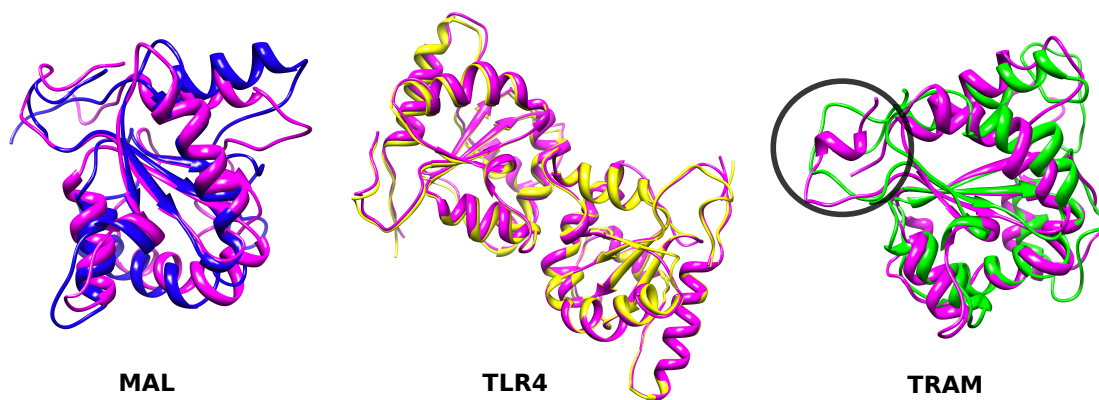


Figura 5.1: Superposición modelos de Nuñez *et al.* (2007) vs modelos Salgado (Resultados no publicados, 2010). La figura muestra la superposición de las representaciones de la estructura terciaria, de los modelos de Nuñez y Salgado, para los dominios TIR del receptor TLR4 y los adaptadores MAL y MyD88. Los modelos de Nuñez se muestran en rosado. Se ha destacado en un círculo la zona perteneciente a los residuos MET4-GLU7 del dominio TIR de TRAM, en la que se observan diferentes estructuras secundarias al comparar los modelos de Nuñez y Salgado.

Adicionalmente, se buscó comparar la calidad estructural de los modelos propuestos en este trabajo con los utilizados por Basith *et al.* (2011) en su estudio de las interacciones entre el receptor TLR4 y los adaptadores MAL y MyD88. En este caso, no se tuvo acceso a las estructuras de Basith, pero si se contaba con la evaluación de los modelos provista en su trabajo. En vista de estos datos, se utilizó la herramienta ProQ para evaluar, bajo tales indicadores, los modelos propuestos por Salgado y utilizados en este trabajo. Los resultados de la evaluación se muestran en la tabla 5.2. En general estos resultados indican que en ambos casos los modelos alcanzan las calificaciones de muy buenos o extremadamente buenos, de acuerdo a la escala utilizada por la herramienta. Las diferencias más importantes se observan en el caso del receptor TLR4 y MyD88, cuyos indicadores alcanzan valores mayores para los modelos propuestos por Basith. En este caso no se puede establecer una relación entre los indicadores y la estructura de los modelos dado que no se dispone de las estructuras de Basith para su inspección.

Tabla 5.2: Evaluación de los modelos estructurales de TLR4, MAL y MyD88 por medio de ProQ

Estructura	ProQ		ProQ - Basith (2011)	
	LGscore	MaxSub	LGscore	MaxSub
TLR4	4.208	0.576	5.608	0.362
MAL	2.861	0.357	2.885	0.389
MyD88	6.286	0.374	7.319	0.518

Rangos de calidad de los modelos según indicador:

LGscore 1.5 relativamente bueno; 2.5 muy bueno; 4 extremadamente bueno

MaxSub 0.1 relativamente bueno; 0.5 muy bueno; 0.8 extremadamente bueno [69]

## 5.2. Modelos de interacción en la vía mediada por TLR4 y adaptadores

### 5.2.1. Efecto del algoritmo de acoplamiento utilizado sobre la obtención de complejos proteína-proteína

El problema de acoplamiento proteína-proteína puede ser resuelto mediante el uso de herramientas *in silico*. Existe una amplia serie de algoritmos de acoplamiento que se dedican a resolver este problema, presentando diferentes aproximaciones con sus respectivas ventajas y desventajas.

Con el objeto de aprovechar las ventajas de cada algoritmo, se probó la acción del software GRAMM, HEX, ZDOCK, DOCK y pyDOCK en la determinación del complejo entre el receptor homodimérico TLR4 y los adaptadores MAL y TRAM vía acoplamiento irrestricto. El resultado esperado en cada caso es la unión de los adaptadores a los sectores cercanos a la interfase del homodímero TLR4 [20, 21, 30]. El mejor modelo obtenido a partir del uso de cada algoritmo se muestra en la Figura 5.2. Tanto en el caso del MAL como TRAM, se observa que los resultados del mejor modelo de cada acoplamiento varían de acuerdo al algoritmo utilizado. Los mejores resultados, de acuerdo a lo esperado, se obtienen con los algoritmos GRAMM y HEX. Por su parte, el algoritmo DOCK (resultados no mostrados) muestra resultados cercanos a la zona de interacción del receptor TLR4. Los algoritmos ZDOCK y pyDOCK muestran, en ambos casos, resultados en los extremos del receptor TLR4 alejados de la zona previamente reportada para la interacción.

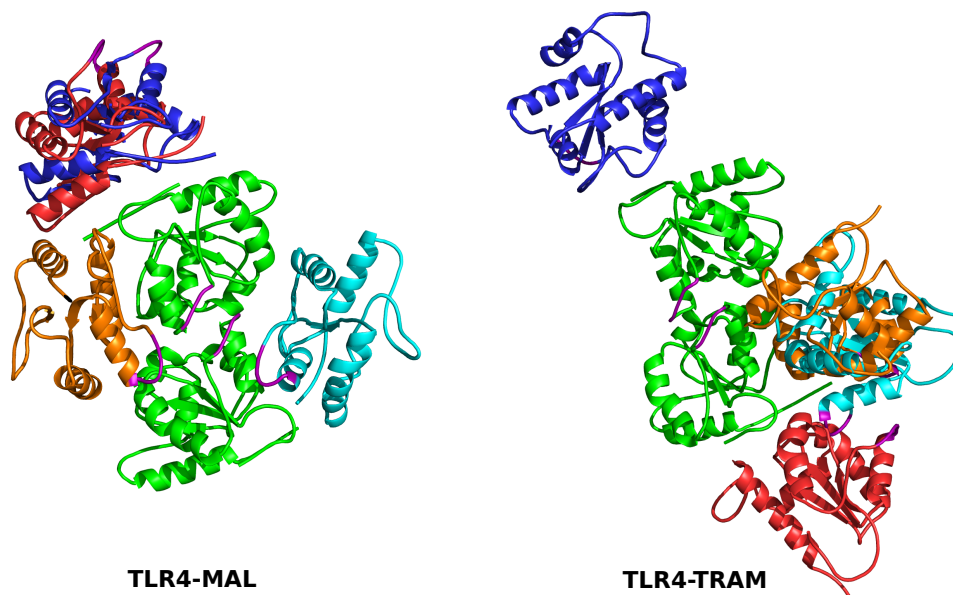


Figura 5.2: Modelo de interacción según distintos algoritmos de acoplamiento. La zona de acoplamiento del dominio TIR del receptor TLR4 y el dominio TIR del adaptador MAL varía de acuerdo al algoritmo de docking utilizado. TLR4 (verde), MAL y TRAM: GRAMM (naranja), HEX(cian), pyDOCK(rojo), ZDOCK(azul)

Otro indicador utilizado para visualizar el potencial que tiene cada algoritmo de acoplamiento, para encontrar soluciones apropiadas para el sistema, es el porcentaje de soluciones que rodea la zona de interacciones buscada. Para esto se analizaron 100 soluciones de cada algoritmo, en las que se observó si los acoplamientos se encontraron en las zonas de interacción del homodímero de TLR4. Los resultados (ver Tabla 5.3) muestran que los mejores algoritmos bajo este criterio son GRAMM y HEX con altos porcentajes de encuentro de modelos positivos. En contraste, los algoritmos ZDOCK y pyDOCK no fueron capaces de reconocer la unión de los adaptadores a la región de la interfase del receptor TLR4, y favorecieron resultados como los observados en la Figura 5.2. En este punto, es necesario destacar la posibilidad de que los resultados reportados para ZDOCK y pyDOCK estén revelando algún sitio de interacción importante distinto a la zona de interacción entre los dominios TIR del homodímero de TLR4. El estudio de los contactos vía  $\Delta$ ASA muestra que los modelos obtenidos por ZDOCK y pyDOCK generan interacción en otros sectores: lazos N-terminal y C-terminal para TLR4; lazo CD y C-terminal para MAL y TRAM (Resultados no publicados). De estos contactos, solo aquellos que se relacionan al lazo C-terminal de los dominios TIR de TLR4 encuentran sustento en evidencia experimental. Sin embargo, en vista de que los últimos modelos de interacción para la vía mediada por TLR4 han apoyado la hipótesis de la formación del homodímero de TLR4 [21, 30, 31, 44], se decide continuar con esta suposición en vista de la alta concentración de resultados observados en la zona de interacción del homodímero, obtenidos vía GRAMM y HEX.

Algoritmo de acoplamiento	% Modelos positivos MAL <sup>1</sup>
GRAMM	76.0
HEX	71.0
ZDOCK	0.0
pyDOCK	7.0

<sup>1</sup> Se define como modelo positivo a los acoplamientos que muestran la unión del adaptador a la zona de interfase en el homodímero TLR4.

Tabla 5.3: Porcentaje de modelos positivos TLR4-MAL según algoritmo de acoplamiento

### 5.2.2. Uso de acoplamiento sitio dirigido en la búsqueda de modelos de interacción para la vía mediada por TLR4 y adaptadores

Una alternativa a la estrategia de búsqueda irrestricta es el uso de una aproximación sitio dirigida. Esta permite reducir el espacio de búsqueda, simplificando el problema de optimización y búsqueda de una mejor pose. La estrategia definida considera la sectorización del receptor TLR4 mediante el uso de 26 cubos que abarcan la totalidad de la superficie del receptor TLR4, de modo que 9 cubren la cadena A de TLR4, 9 la cadena B de TLR4 y 8 cubren la interfase del homodímero TLR4 (Ver Anexo A.3). Para la implementación se utilizó los algoritmos Autodock 4 y Vina, en base a la estrategia planteada, para encontrar los mejores modelos de TLR4-MAL y TLR4-TRAM. En este caso se cambió el algoritmo con que se estaba trabajando debido a que los algoritmos anteriormente mencionados no permitían llevar a cabo una búsqueda sitio dirigida, sino que realizaban por defecto una búsqueda de conformaciones en todo el espacio definido. Aprovechando el hecho de que tanto Autodock 4 como Vina cuentan con funciones empíricas para el cálculo de energía de un complejo, se generó modelos de interacción TLR4-MAL y TLR4-TRAM en cada uno de los sectores previamente definidos. El objetivo de esto es comparar la afinidad relativa del adaptador por cada sector. Los resultados del ejercicio anterior se muestran en la Figura 5.3.

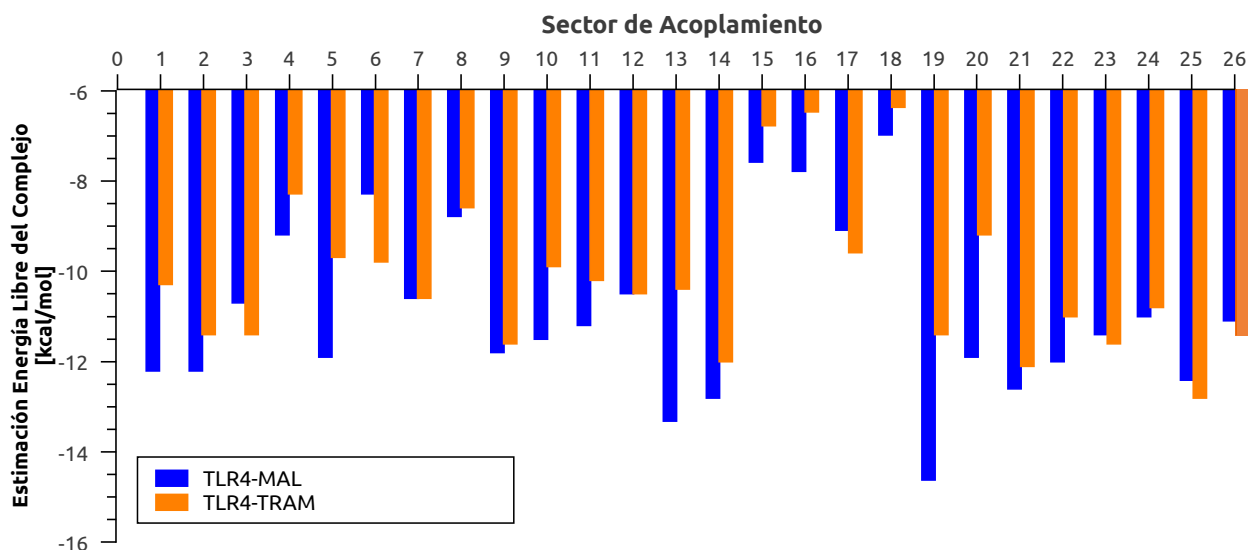


Figura 5.3: Energía para el mejor modelo obtenido por ADT4 según sector de acoplamiento. La energía libre de los complejos TLR4-MAL o TLR4-TRAM varía de acuerdo a las distintas zonas en las cuales se produzca la interacción. Cabe destacar que los sectores 19, 21, 22, 24 y 25 son aquellos que rodean al sector del BB loop. Todos los valores son el resultado de un análisis en triplicado.

A partir de lo obtenido es posible dar cuenta que los valores más bajos, y por tanto favorables, de energía

se observaron en sectores relacionados al BB loop (sector 19 y 25 para MAL y TRAM respectivamente). En el caso de TRAM los resultados no son concluyentes, ya que la función de energía que posee el algoritmo Vina tiene un error de 2.62 [kcal/mol] [103], por lo que existen otros resultados que pueden ser considerados equivalentes energéticamente. Los sectores que mostraron un menor valor de energía libre, de acuerdo a la función de *scoring* de Autodock 4, se muestran en la Figura 5.4.

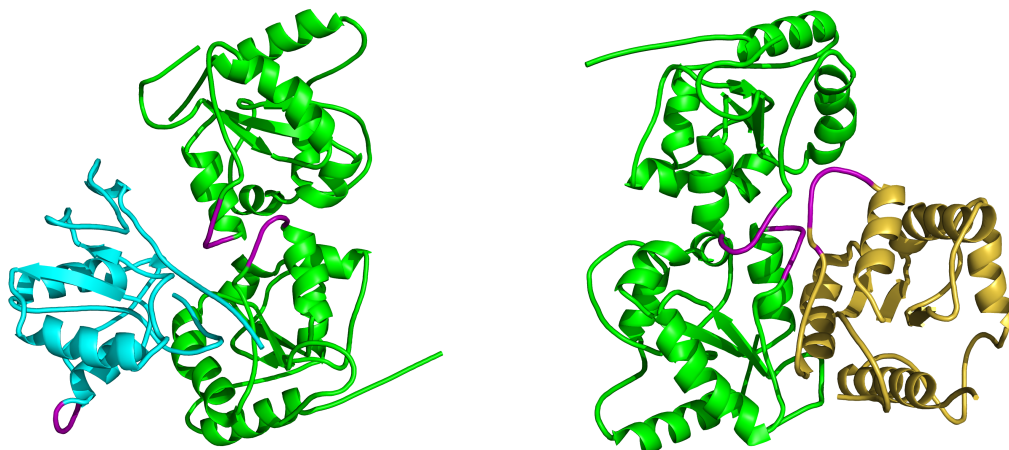


Figura 5.4: Resultados del acoplamiento sitio dirigido de TLR4-MAL y TLR4-TRAM vía Autodock 4. En la figura de la izquierda se observa la interacción de menor energía entre los dominios TIR del homodímero de TLR4 y el adaptador MAL, correspondiente al sector 19 de acuerdo al diagrama expuesto en el Anexo A.3. La figura de la derecha muestra la interacción de menor energía de unión para los dominios TIR de TLR4 y TRAM, correspondiente al sector 25 de acuerdo a la sectorización ya planteada. En ambos casos se ha destacado los BB loops en color púrpura.

Independiente de estos últimos resultados se privilegió la evidencia experimental que se tiene de las interacciones TLR4-MAL y TLR4-TRAM, por lo que se revisó los resultados de los sectores que involucran al BB Loop. Dichos resultados mostraron, al contrario de lo esperado, que la interacción de los adaptadores con el receptor TLR4 se daba de tal manera que el BB Loop del adaptador (MAL o TRAM) se encontraba en todos los casos alejado de la zona BB loop del homodímero TLR4. Con la idea de mejorar el resultado del acoplamiento se utilizó el algoritmo Autodock para realizar un acoplamiento en los sectores en que Vina entregó mejores resultados. En el caso del modelo del adaptador MAL, ninguno de los resultados logró reorientar al BB *loop* hacia la zona del receptor TLR4. Por el contrario, en el caso de TRAM, uno de los cluster de resultados obtenidos muestra una correcta orientación del BB *loop*.

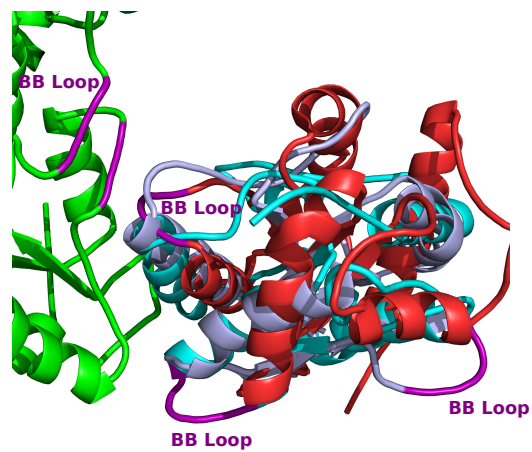


Figura 5.5: Modelos de interacción TLR4-MAL mediante acoplamiento sitio dirigido vía Autodock 4. La figura muestra parte del dominio TIR de TLR4 (verde) en la que se acoplan modelos del dominio TIR de MAL (diferentes modelos presentados en color cian, púrpura y rojo). En la figura se han destacado con color morado los residuos pertenecientes a los BB loops del receptor o adaptadores.

De lo anterior se desprende que en el caso del adaptador MAL, el uso de las herramientas Autodock 4 y Vina no permitió responder bien a la evidencia biológica, que indica la interacción entre segmentos BB *loops* tanto de TLR4 como de las proteínas adaptadoras [20–23, 34, 35]. Bajo los criterios de estas herramientas los modelos que presentaron la mejor puntuación en la función de energía libre del programa fueron aquellos en los que se observaba la orientación de los segmentos distintos al BB *loop* de los adaptadores, hacia el BB *loop* del homodímero de TLR4. Lo anterior indica que si bien Autodock 4 y Vina son buenas herramientas para la generación de modelos de acoplamiento, no entregan las soluciones esperadas, de acuerdo a los supuestos adoptados, para el sistema con el cual se está trabajando. Lo anterior puede tener relación con el hecho de que tanto Autodock 4 como Vina se especializan en la unión proteína-ligando, por lo que trabajan con proteínas de un tamaño mucho menor al de las utilizadas. En el caso de TRAM, el uso de la estrategia sitio dirigida sí permitió encontrar una conformación con la orientación esperada. Sin embargo, de acuerdo a los cálculos energéticos de Autodock 4, no es posible establecer, dadas las leves diferencias energéticas encontradas por sector, que la interacción de los BB *loops* resulta en la formación de un complejo más estable desde el punto de vista energético.

Finalmente, es necesario destacar que existe la posibilidad de que Autodock efectivamente este mostrando las conformaciones reales y de menor energía para cada interacción. No obstante, de acuerdo a las suposiciones consideradas (interacción entre BB *loops*), las conformaciones del complejo TLR4-MAL no responden la fenomenología estudiada. Si bien esta decisión responde a la evidencia revisada, no es posible asegurar que esto sea lo que sucede *in vivo* y existe un claro sesgo en la elección del modelo a estudiar vía dinámica molecular, aún cuando se trata de respaldar dicha elección de la mejor manera posible.

### 5.2.3. Descripción de los modelos de interacción obtenidos

Una vez establecido el algoritmo y la estrategia de búsqueda, se seleccionaron los modelos para las interacciones de las proteínas TLR4-MAL, TLR4-TRAM y TRAL-MAL-MyD88. Los modelos fueron definidos a partir del uso de reglas o directrices que permitieron acotar el número de modelos hasta dar con un único modelo representativo para cada interacción. Cabe destacar que todos los modelos están basados en el concepto de homodimerización del receptor TLR4 como primer paso dentro de la cascada de señalización en la vía pro-inflamatoria en la que participan los receptores estudiados [20,21,30,31,42,44,84]. Adicionalmente, todos los modelos planteados consideran la unión exclusiva de MAL o TRAM a ambos

costados del dímero de TLR4, y la posterior unión de MyD88 al complejo TLR4-MAL previamente formado.

### 5.2.3.1. Modelo TLR4-MAL

El modelo TLR4-MAL obtenido se observa en la Figura 5.7. Tanto el complejo modelado mediante GRAMM como el presentado por Nuñez señalan que el adaptador MAL puede unirse a ambos costados del receptor TLR4. La simetría que presenta la estructura del homodímero de TLR4 favorece la ubicación del adaptador MAL a ambos costados del receptor. De acuerdo al modelo planteado, en ambos casos, las moléculas de MAL se unen orientando sus BB *loops* (color morado en la figura) hacia la zona de interacción del homodímero TLR4 en la que se encuentran sus correspondientes BB *loops*, aunque el acoplamiento y orientación es distinta para cada molécula de MAL.

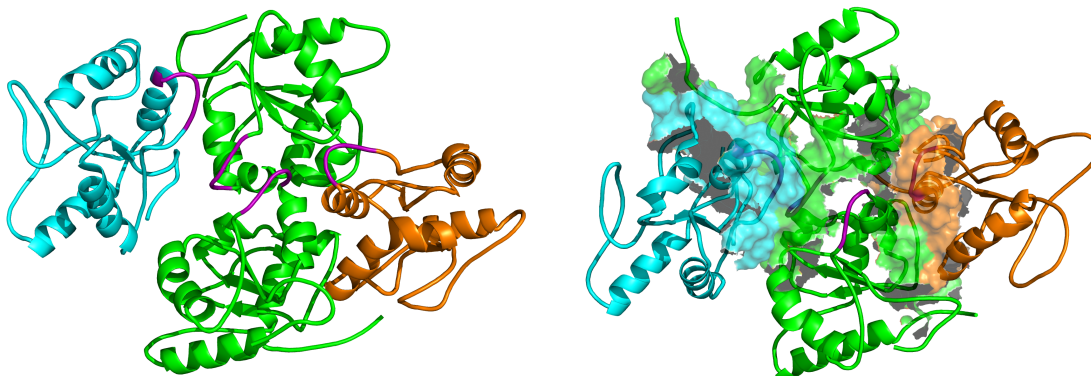


Figura 5.6: Modelo obtenido para la interacción TLR4-MAL. Se muestran las interacciones propuestas para el dominio TIR de TLR4 (verde) y los dominios TIR de MAL Nuñez (dorado) y MAL propuesto (cian). En la figura de la derecha se han resaltado los residuos que interactúan por medio de una representación tridimensional de su superficie.

La superficie escondida en la interfase por la interacción TLR4-MAL es de 3881.9 [ $\text{\AA}^2$ ] para el acoplamiento del modelo de Nuñez y de 1873.6 [ $\text{\AA}^2$ ] para la interacción del modelo propuesto. Para el caso del complejo TLR4-MAL propuesto existen 22 residuos de MAL, 10 residuos de la cadena A de TLR4 y 3 de la cadena B de TLR4 que pertenecen a la zona de interacción. Estos residuos muestran un  $\Delta\text{ASA}_{rel} > 40$ , lo que indica que son residuos no expuestos y que por tanto son definidos como posibles zonas de contacto entre el homodímero y el adaptador [20]. Los residuos mencionados corresponden, en el caso de MAL, principalmente a los sectores de las hojas  $\beta A$  y hélices  $\alpha A$  y  $\alpha C$ , además de los *loops* AB, BB y CD. En el caso de TLR4 A corresponden principalmente al BB *loop* y la hélice  $\alpha B$ , mientras que en la cadena B de TLR4 los residuos escondidos forman parte del AA *loop*.

En el caso del complejo TLR4-MAL de Nuñez, según la misma definición, se tiene que 36 residuos de TLR4 y 36 residuos de MAL pertenecen a la zona de interacción. El mayor número de contactos que muestra el modelo de Nuñez se correlaciona con el mayor tamaño de la superficie escondida en la interfase por la interacción TLR4-MAL. En el caso de MAL los residuos mencionados se encuentran principalmente en los *loops* BB, CD y DD, además de las hélices  $\alpha B$  y  $\alpha C$ . Por su parte los residuos escondidos en TLR4 pertenecen al CC *loop* en la cadena A, y principalmente el BB y CD *loop* junto a la hélice  $\alpha B$  en la cadena B de TLR4.

El tipo de aminoácidos que forman parte de la interfase de cada modelo se aprecia en la tabla ??.

Los datos muestran que, en el caso del modelo de Nuñez, la mayor parte de estos residuos son de carácter hidrofóbico o polares sin carga. Por su parte el modelo propuesto muestra que los residuos escondidos por el adaptador MAL son principalmente polares sin carga e hidrofóbicos, mientras que en TLR4 son mayoritariamente polares cargados e hidrofóbicos. Si bien no se puede hablar de la existencia núcleo hidrofóbico en las interacciones se aprecia que de manera transversal la presencia de al menos un 30% de residuos de carácter hidrofóbico formando parte de la interfase.

Tipo de aminoácidos	Modelo Propuesto		Modelo Nuñez	
	MAL [%]	TLR4 [%]	MAL [%]	TLR4 [%]
Hidrofóbicos	27.3	30.8	47.2	33.3
Polares sin carga	45.5	15.4	22.2	36.1
Polares negativos	18.2	30.8	13.9	8.3
Polares positivos	4.5	15.4	2.8	11.1
Otros <sup>1</sup>	4.5	7.7	13.9	11.1

<sup>1</sup> Corresponden a PRO, GLY y CYS

Tabla 5.4: Tipo de aminoácidos presentes en la zona de interacción TLR4-MAL según modelo de acoplamiento

### 5.2.3.2. Modelo TLR4-MAL-MyD88

A partir de los modelos de interacción TLR4-MAL se generaron los modelos de interacción con MyD88 por medio del algoritmo de acoplamiento GRAMM. De acuerdo a las características de la interacción estudiada entre TLR4 y MAL, se plantean dos modelos de interacción de MyD88, uno a cada lado del receptor TLR4. Ambos modelos se unen en sectores simétricos y equivalentes, pero con orientaciones distintas.

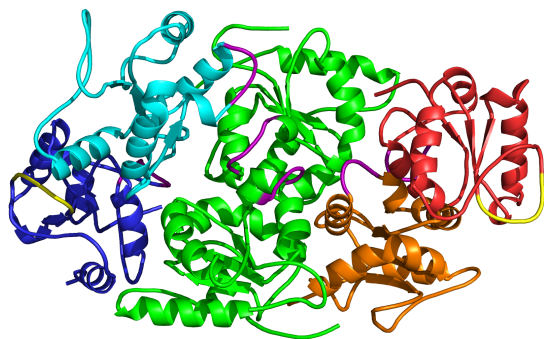


Figura 5.7: Modelo obtenido para la interacción TLR4-MAL-MyD88. De acuerdo a los resultados obtenidos el adaptador MyD88 es capaz de unirse al complejo TLR4-MAL interactuando en ambos casos sobre la nueva superficie establecida por la interacción TLR4-MAL. La figura muestra, sobre una misma representación del dominio TIR del receptor TLR4, las interacciones que produce el dominio TIR del adaptador MyD88 sobre los complejos TLR4-MAL que forman los modelos del dominio TIR de MAL de Nuñez y Salgado. Los modelos para los dominios TIR del receptor y los adaptadores se observan en distintos colores: TLR4 (verde), MAL Nuñez (dorado), MAL propuesto (cian), MyD88 modelo 1 (azul), MyD88 modelo 2 (rojo). En la figura de la derecha se han resaltado los residuos que están interactuando por medio de la representación de su superficie. Los BB loop se han destacado en morado y los CD loop de MyD88 en amarillo



En el caso del modelo 1 de MyD88 se calculó una superficie escondida de 2176.6 [Å<sup>2</sup>] en la interfase del complejo TLR4-MAL-MyD88 (sin considerar la ya escondida superficie TLR4-MAL), mientras que para el modelo 2 la misma superficie alcanzó un valor de 1363.8 [Å<sup>2</sup>]. De acuerdo a la definición planteada ( $\Delta ASA_{rel}$ ) para determinar aquellos residuos que presentan potenciales contactos, se observa que 9, 9 y 18 residuos de TLR4, MAL y MyD88 respectivamente pertenecen a la zona de interacción en el modelo 1. Estos residuos provienen de las hélices  $\alpha A$ ,  $\alpha C$  y los lazos BB y EE de MyD88; los lazos DD y EE de TLR4; y el *loop* AB junto con la hélice  $\alpha E$  de MAL. Por su parte en el segundo modelo planteado se presentan 7, 17 y 25 residuos interactúan desde TLR4, MAL y MyD88 respectivamente. Los residuos pertenecen al AA, BB, EE *loop* y la hélice  $\alpha A$  de MyD88; AA, BB *loop* y hélice  $\alpha A$  de MAL; y sectores de las hélices  $\alpha A$  y  $\alpha E$  de la cadena B de TLR4. Este último modelo presenta choques intermoleculares en la zona de interacción entre la hélice  $\alpha A$  de MyD88 y el modelo MAL (N), por lo que se modificó la estructura de modo de evitar el colapso de la estructura durante la simulación de dinámica molecular.

Tal como se aprecia en la Tabla 5.5, la interfase en la interacción TLR4-MAL-MyD88 del modelo 1 está compuesta principalmente por aminoácidos de tipo hidrofóbicos y polares sin carga en el contacto MAL-MyD88, mientras que en el contacto TLR4-MyD88 se observa una mayor proporción de residuos con carga. En el segundo modelo la zona de interacción está compuesta por aminoácidos principalmente hidrofóbicos en el caso MAL-MyD88, mientras que la interacción TLR4-MyD88 no muestra preponderancia de algún tipo particular de aminoácidos.

Tipo de aminoácidos	Modelo 1			Modelo 2		
	MAL [%]	TLR4 [%]	MyD88 [%]	MAL [%]	TLR4 [%]	MyD88 [%]
Hidrofóbicos	44.0	22.2	11.1	47.1	28.5	44.0
Polares sin carga	33.3	22.2	33.3	11.7	28.5	16.0
Polares negativos	11.1	33.3	27.7	29.4	28.5	12.0
Polares positivos	11.1	22.2	11.1	5.8	14.2	8.0
Otros <sup>1</sup>	0.0	0.0	16.6	5.8	0.0	20.0

<sup>1</sup> Corresponden a PRO, GLY y CYS

Tabla 5.5: Tipo de aminoácidos presentes en la zona de interacción TLR4-MAL según modelo de acoplamiento

### 5.2.3.3. Modelo TLR4-TRAM

Los modelos TLR4-TRAM estudiados se observan en la Figura 5.8. Nuevamente, la geometría que presenta el homodímero TLR4 permite que la unión del adaptador TRAM sea factible en ambos costados de la zona de interacción del receptor. En efecto, tanto el modelo de interacción propuesto como el presentado por Nuñez, muestran el acoplamiento de TRAM a la interfase del homodímero TLR4. Sin embargo, los adaptadores de TRAM se unen en extremos opuestos y muestran una orientación antiparalela. De manera similar a la unión del adaptador MAL, los modelos de TRAM se unen al receptor TLR4 orientando sus *loops* BB hacia la zona de interacción del homodímero lo que resalta la importancia que puede llegar a tener este segmento en las interacciones de esta familia de proteínas.

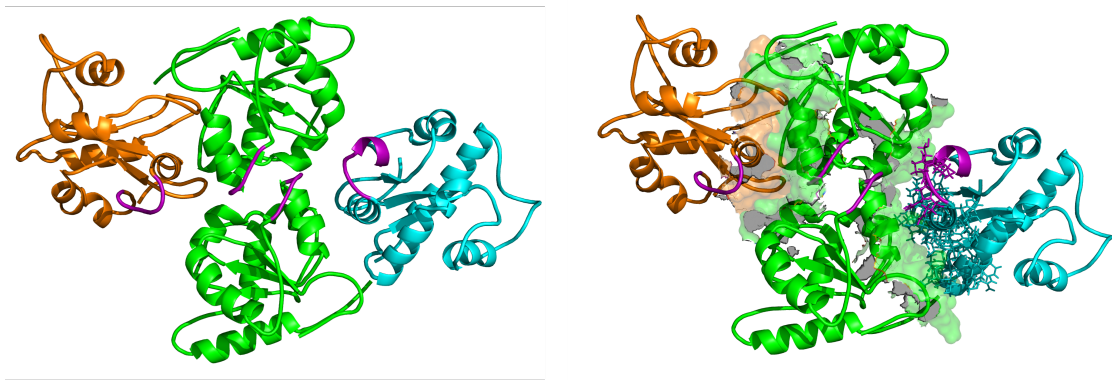


Figura 5.8: Modelo obtenido vía *docking* para la interacción TLR4-TRAM. Los modelos propuestos indican la unión del adaptador TRAM a la zona de interacción entre los dominios TIR de los receptores TLR4. Sobre una misma representación del receptor TLR4 se han ubicado los modelos propuestos para la interacción del dominio TIR de TRAM de los modelos de Nuñez y Salgado. TLR4 (verde), TRAM Nuñez (dorado), TRAM propuesto (cian). En la figura de la derecha se han resaltado las superficies de interacción.

La superficie escondida en la interfase por la interacción TLR4-TRAM es de 3853.1 [ $\text{\AA}^2$ ] para el acoplamiento del modelo de Nuñez y de 2422.4 [ $\text{\AA}^2$ ] para la interacción del modelo propuesto. En el complejo TLR4-TRAM propuesto existen 15 residuos de TRAM, 6 residuos de la cadena A de TLR4 y 3 de la cadena B de TLR4 que pertenecen a la zona de interacción, ie. que muestran un  $\Delta\text{ASA}_{rel} > 40$ . En el caso de TRAM los residuos a los que se hace referencia pertenecen principalmente a los sectores de la hélice  $\alpha\text{B}$  y  $\alpha\text{C}$ , además de los *loops* BB, CC y CD. En el caso de TLR4 A corresponden al BB y la hélice  $\alpha\text{B}$ , mientras que en la cadena B de TLR4 los residuos escondidos forman parte de los *loop* AA y CC.

En el caso del complejo TLR4-TRAM modelado por Nuñez se observa que son 20 los residuos de TRAM y 19 provenientes del receptor TLR4 los que forman parte de la zona de interacción. En este caso los residuos que participan en la interacción pertenecen, en el caso del adaptador TRAM, al CD *loop*, la hélice  $\alpha\text{B}$  y la hoja  $\beta\text{B}$ . En el caso del homodímero TLR4, los residuos corresponden a las hélices  $\alpha\text{B}$  y  $\alpha\text{C}$ , y el *loop* CD en el caso de la cadena A. En la cadena B de TLR4 los residuos señalados hacen referencia a las estructuras del CC *loop* y la hélice  $\alpha\text{D}$ . De manera similar al caso del modelo TLR4-MAL, el mayor número de contactos observados en el modelo de Nuñez está directamente relacionado con el mayor tamaño del área escondida en la interacción TLR4-TRAM.

El detalle acerca del tipo de aminoácidos que forman parte de la interfase se muestra en la Tabla 5.6. De lo anterior se desprende que a diferencia de la interfase de interacción TLR4-MAL, en este caso se observa una mayor proporción de residuos polares con carga, en general sobre 30%, lo que puede ser un indicador de la importancia de las interacciones iónicas en el acoplamiento del adaptador TRAM.

Tipo de aminoácidos	Modelo Propuesto		Modelo Nuñez	
	TRAM [%]	TLR4 [%]	TRAM [%]	TLR4 [%]
Hidrofóbicos	26.7	22.2	25.0	36.8
Polares sin carga	20.0	33.3	30.0	36.8
Polares negativos	26.7	0.0	25.7	5.3
Polares positivos	13.3	33.3	10.0	15.8
Otros <sup>1</sup>	6.7	11.1	10.0	5.3

<sup>1</sup> Corresponden a PRO, GLY y CYS

Tabla 5.6: Tipo de aminoácidos presentes en la zona de interacción TLR4-TRAM según modelo de acoplamiento

## 5.3. Aplicación de dinámica molecular sobre los modelos de interacción

### 5.3.1. Caracterización de las simulaciones de dinámica molecular

Los complejos modelados para TLR4-MAL, TLR4-TRAM y TLR4-MAL-MyD88 fueron sometidos a dinámica molecular para estudiar estabilidad y dinámica de las interacciones. Se simuló el comportamiento de los complejos en solvente explícito (TIP3P) durante una etapa de producción de 6 [ns], siguiendo el protocolo detallado e implementado en GROMACS con un paso de intergración de 2 [fs] y el campo de fuerza AMBER99SB. Los estados de protonación de los residuos fueron asignados automáticamente por GROMACS. En todos los casos se extrajo datos acerca de la estructura y variables termodinámicas de los modelos simulados en intervalos de 20 [ps].

Los perfiles de energía y temperatura obtenidos durante la etapa de producción de la simulación del complejo TLR4-MAL se observan en la Figura 5.9. En ambos casos se observa que las variables termodinámicas estudiadas se mantienen estables y dentro de los valores esperados, con excepción de la presión. En particular, esta última variable muestra grandes oscilaciones y no logra permanecer bajo control durante la simulación. Las variaciones presentadas por la presión durante la simulación han sido documentadas previamente por los desarrolladores de GROMACS [55]. Si bien no responden a la fenomenología del sistema, estas no tienen mayores implicancias en el resto de las variables del sistema ni en la calidad general de la simulación. Para el resto de las variables se observó, durante la etapa de producción, coeficientes de variación de %0.51 %, 0.15 % y 0.23 % para la energía cinética, potencial y total del sistema. Por su parte la temperatura presentó fluctuaciones promedio del orden de 0.51 %.

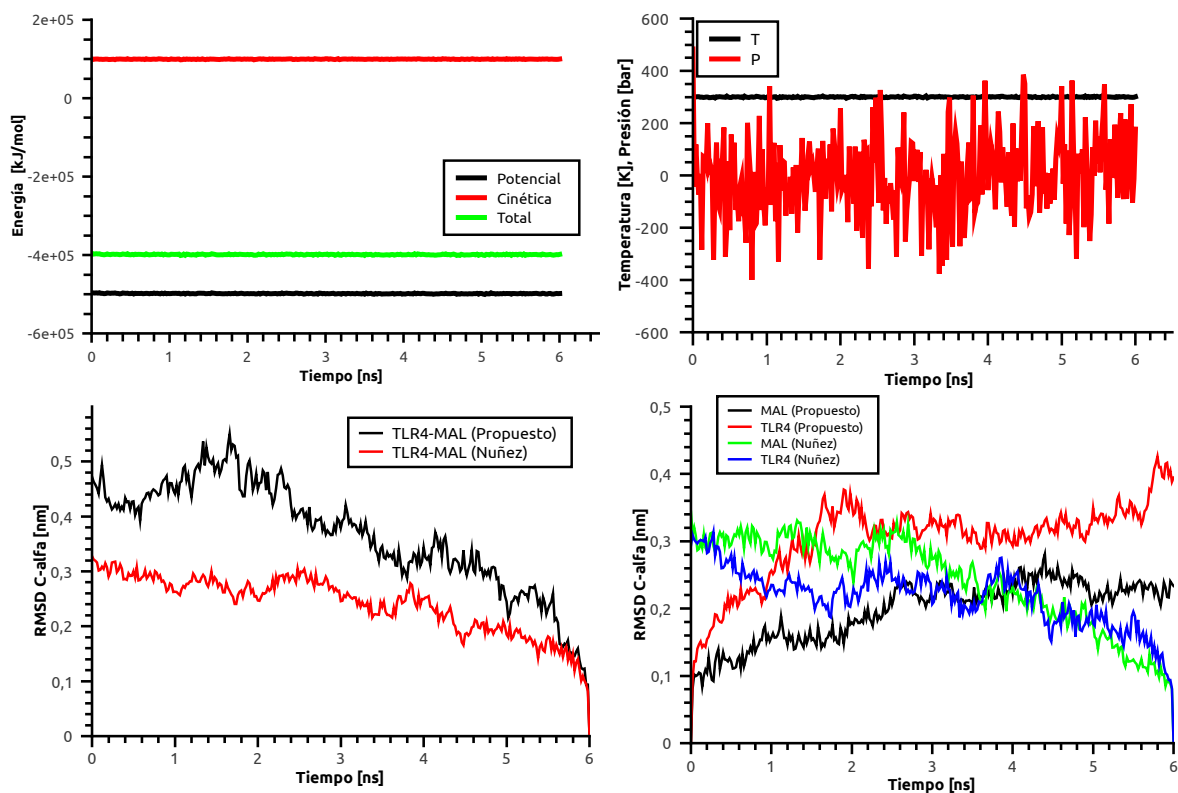


Figura 5.9: Perfiles de energía, temperatura, presión y RMSD ( $C\alpha$ ) para la dinámica molecular de TLR4-MAL

Adicionalmente se verificó la estabilidad de las estructuras a lo largo de la simulación calculando el RMSD  $C\alpha$  usando como referencia la estructura inicial de la etapa de producción. En el caso del complejo propuesto se obtuvo un RMSD  $C\alpha$  promedio de 3.6 [Å] y un coeficiente de variación de 26.3 %, mientras que en el complejo de Nuñez se observó un RMSD  $C\alpha$  promedio de 2.4 [Å] y variaciones de un 22 %. Estos valores si bien son bajos en magnitud, muestran que ninguno de los complejos alcanza completa estabilidad durante la simulación. Esto último indica la necesidad de extender la duración de la etapa de equilibrio. Se verificó también el RMSD  $C\alpha$  de las estructuras TLR4 y MAL durante la simulación usando el mismo tipo de referencia. En el caso del complejo TLR4-MAL propuesto, la estructura de TLR4 se estabiliza a partir de los 2 [ns] de la simulación del complejo mostrando un RMSD  $C\alpha$  promedio de 3.2 [Å] y un coeficiente de variación de un 6.1 %. Por su parte el adaptador MAL también se estabiliza después de los 2.5 [ns], manteniendo un RMSD  $C\alpha$  promedio de 2.2 [Å] y variaciones de un 6 %. Por otro lado, los datos recuperados de la simulación del complejo TLR4-MAL(Nuñez, en adelante modelo N) indican que tanto la estructura de TLR4 como MAL parecen estabilizarse durante los primeros 3 [ns] de la simulación con un promedio 2.4 y 2.7 [Å] y variaciones de un 10.5 y 5.6 % respectivamente. Sin embargo, una vez alcanzados los 4 [ns] de simulación ambas estructuras muestran un descenso en el RMSD  $C\alpha$ , no volviendo a alcanzar un equilibrio durante el resto de la simulación. De manera general, estos resultados indican la necesidad de extender la etapa de equilibrio previo a la extracción de datos desde la simulación.

Evidencia adicional de la estabilidad del sistema se obtiene a partir de la observación de la dinámica molecular del complejo usando la herramienta de visualización molecular VMD. A partir de la observación del complejo TLR4-MAL propuesto (en adelante modelo P) se observa que el sistema se mantiene unido durante toda la simulación. Sin embargo se aprecia que la estructura de MAL adopta dos orientaciones claras que presentan ciertas diferencias. En primera instancia se observa el BB *loop* del adaptador MAL esta

orientado hacia los lazos BB del receptor TLR4, pero posteriormente se observa, a partir de los 3 [ns] un leve desplazamiento del lazo BB que permite una mejor interacción del lazo CD con el receptor TLR4 (ver Figura 5.10)

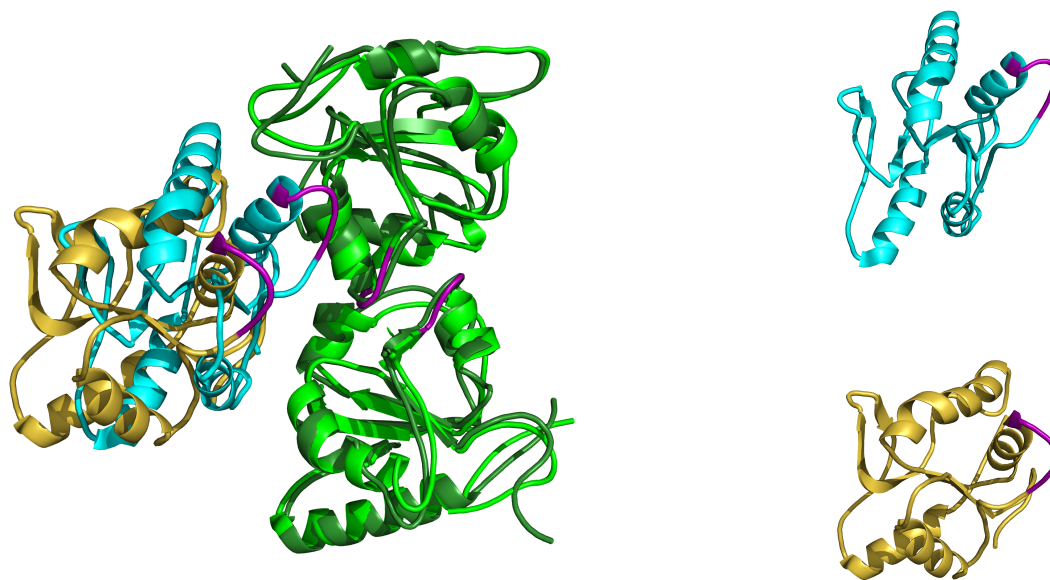


Figura 5.10: Cambios conformacionales en la estructura de MAL durante dinámica molecular de complejo TLR4-MAL (P). A la izquierda de la figura se muestra la estructura del complejo formado por los dominios TIR del homodímero receptor TLR4 y el adaptador MAL. Se observan los cambios conformacionales desde el inicio de la simulación de dinámica molecular  $t = 0$  [s] (TLR4 en verde oscuro, MAL en cian) hasta la conformación adoptada en  $t = 3$  [ns] (TLR4 en verde claro, MAL en amarillo). A la derecha de la figura se observan la misma vista frontal para el dominio TIR de MAL en  $t = 0$  [s] (cian) y  $t = 3$  [ns] (amarillo) En cada caso se han destacado las estructuras de los BB loops en color morado.

Este comportamiento está en acuerdo con los datos que entrega el cálculo del RMSD  $C\alpha$ , que indican cierta estabilización del sistema pasados los 2 [ns] de simulación. Por otro lado, la inspección de la dinámica molecular del complejo TLR4-MAL(N) muestra un sistema que permanece unido durante la totalidad de la simulación. En este caso las zonas de interacción se mantienen sin mayores variaciones durante la simulación, y además gran parte de los cambios en la estructura del complejo provienen principalmente de la vibración del sistema y el movimiento de las cadenas terminales. De lo anterior se desprende que, si bien los datos de RMSD  $C-\alpha$  no muestran una completa estabilidad del sistema, la visualización del complejo sí muestra que un complejo estable. De la misma manera, los histogramas de frecuencias A.6 muestran la estabilidad del sistema por cuanto muestran todas distribuciones unimodales. En vista de esto se considera que la simulación sí puede ser utilizada para la extracción de datos acerca de la interacción del complejo y para el cálculo de la energía libre. Sin perjuicio de lo anterior, se considera necesario extender las simulaciones de estos complejos de manera de estudiar los efectos a largo plazo (escala 10-100 [ns]) sobre la estructura y estabilidad del complejo.

Al integrar el adaptador secundario MyD88 al complejo TLR4-MAL y llevar a cabo la dinámica molecular se observaron los perfiles de energía, temperatura y RMSD  $C\alpha$  que muestra la Figura 5.11. En todos los casos se observa que las variables termodinámicas estudiadas se comportan de acuerdo a lo

esperado, mostrando todas estabilización, con excepción de la presión. Durante la simulación de producción se observó que las energías potencial, cinética y total presentaron coeficientes de variación de 0.15, 0.51 y 0.23 %. Por su parte la temperatura se estabilizó alrededor de los 300 [K] mostrando variaciones del 0.48 %. Estos valores confirman la estabilidad termodinámica del sistema simulado y validan su uso en la obtención de datos para el cálculo de la energía libre.

La medición de la estabilidad de las estructuras a partir del RMSD  $C\alpha$  indica que el complejo TLR4-MAL-MyD88 muestra un RMSD  $C\alpha$  promedio de 2.3 [Å] y una variación del 30 %, lo que indica que la estructura inicial no alcanza a equilibrarse a lo largo del tiempo transcurrido en la simulación. En cuanto a las estructuras individuales, se observa que tanto MAL y MyD88 muestran cierto grado de estabilización en el rango 2-6 [ns] con un RMSD  $C\alpha$  promedio de 1.3 Å y 1.1 [Å], con fluctuaciones de 16.0 y 13.0 % respectivamente. Por el contrario, la estructura del receptor TLR4 no muestra estabilización durante la simulación y el gráfico de su RMSD  $C\alpha$  muestra un perfil similar al del complejo TLR4-MAL-MyD88, con un RMSD  $C\alpha$  promedio de 2.5 [Å] y variaciones del 30 %. Esto indica que la inestabilidad observada del complejo puede estar relacionada al movimiento de las cadenas de TLR4 durante la simulación. Sin embargo, la inspección visual de la simulación del complejo muestra que no se producen cambios de gran magnitud en la estructura. Los cambios más notables se observan en los extremos terminales de TLR4, y en este caso son despreciables dado que naturalmente el receptor TLR4 se une mediante estos lazos a la membrana plasmática. Por otro lado, de los datos observados se desprende que la interacción del adaptador MyD88 según este modelo particular, no introduce una mayor estabilidad al sistema aunque sí se observa una reducción del RMSD calculado. La estructura de TLR4 es aquella que muestra las mayores variaciones en su conformación, lo que se observa también en los histogramas de frecuencia (Anexo A.6) que muestran una distribución multimodal. Al igual que en el caso anterior se vuelve necesario extender el tiempo de simulación del sistema para los efectos del estudio de los cambios conformacionales que se puedan producir en la estructura del complejo.

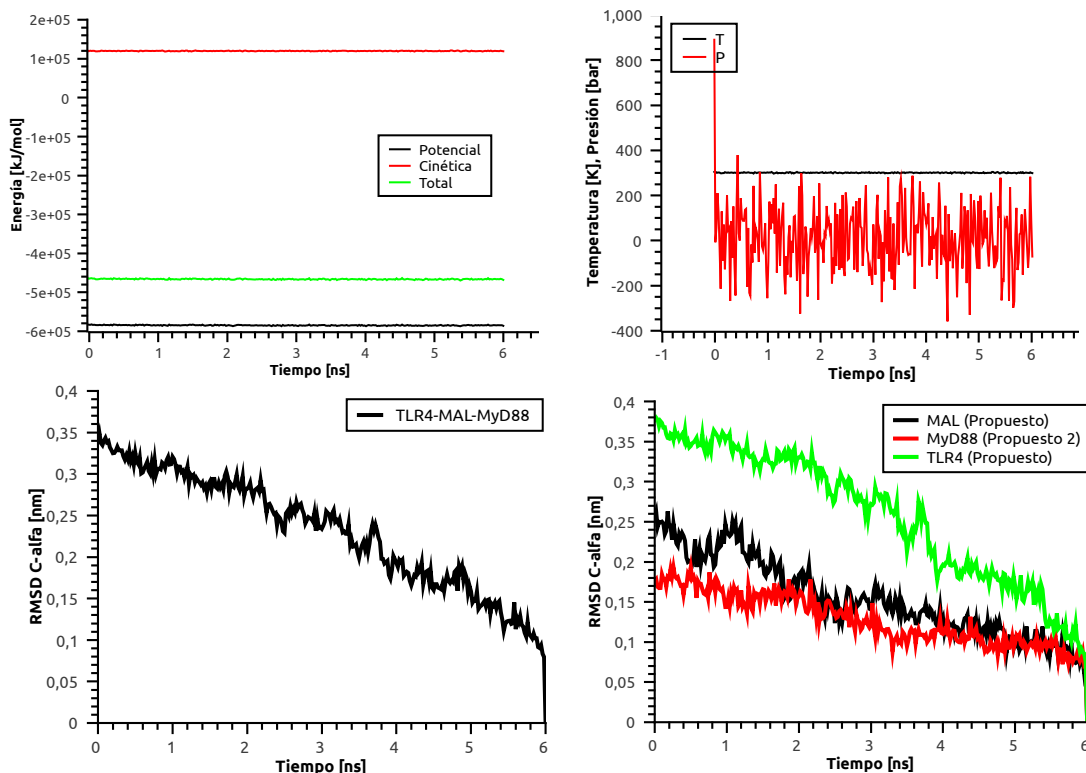


Figura 5.11: Perfiles de energía, temperatura, presión y RMSD (C $\alpha$ ) para la dinámica molecular de TLR4-MAL-MyD88

En el caso del complejo TLR4-TRAM, las variables termodinámicas del sistema, al igual que en los casos anteriores, se mantuvieron dentro de los rangos esperados. En particular los valores de las energías potencial, cinética y total mostraron solo mínimas fluctuaciones durante la simulación del orden de 0.1, 0.4 y 0.2 %. La temperatura, por su parte, muestra fluctuaciones del 0.4 %. La magnitud y el valor de los datos reportados indican la estabilidad termodinámica del sistema simulado.

La estabilidad del sistema, medida a través del RMSD C $\alpha$ , muestra que el complejo propuesto TLR4-TRAM (P) se equilibra alrededor de los valores 2.8 y 1.5 [Å] (con fluctuaciones de un 13 y 17 %) durante los intervalos 1-4 y 4-6 [ns]. A pesar de esto, al observar la simulación del complejo no se observan grandes variaciones en la estructura proteica al comparar la estructura del complejo en los intervalos señalados. Por su parte, el RMSD C-alfa del complejo TLR4-TRAM (N) se mantiene estable durante los primeros 2 [ns] de la simulación (promedio 2.9 [Å] y variaciones de un 2 %) pero luego varía su valor sin alcanzar otra zona de estabilidad y muestra variaciones de hasta un 30 %. La observación de la dinámica molecular de este complejo muestra que, tal como en los casos anteriores, no se presentan grandes cambios conformacionales y solo se observa cambios considerables en los extremos terminales.

Al analizar la estabilidad estructural de cada proteína de manera individual, es posible observar que en este caso las proteínas muestran comportamientos similares a los que muestra el complejo al que pertenecen. Los histogramas de RMSD (Anexo A.6) muestran distribuciones unimodales en todo los casos, con excepción del adaptador TRAM en el modelo TLR4-TRAM(P). En el último caso, si bien se observa una distribución multimodal, la observación de la estructura del adaptador muestra que la estructura es estable a lo largo de la simulación. A esta altura, y al igual que en el estudio de los complejos TLR4-MAL y TLR4-MAL-MyD88, se reconoce la necesidad de extender el largo de la simulación para dar cuenta de si efectivamente las estructuras presentadas son efectivamente modelos representativos del sistema.

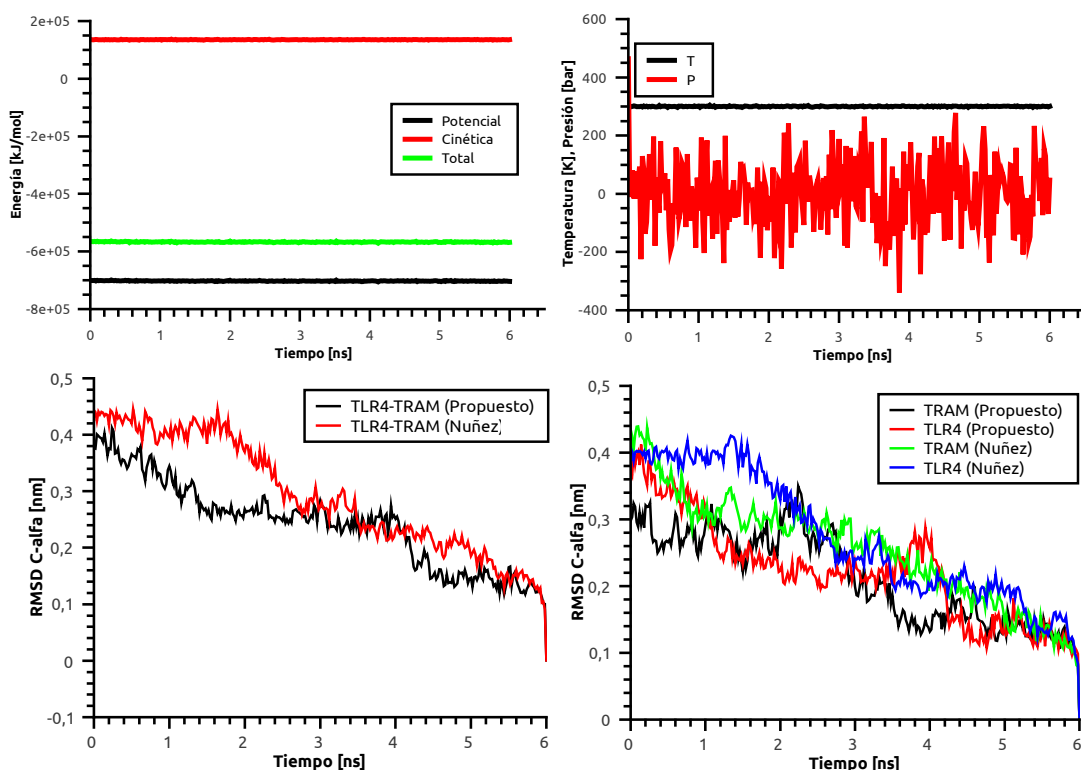


Figura 5.12: Perfiles de energía, temperatura, presión y RMSD (C $\alpha$ ) para la dinámica molecular de TLR4-TRAM

### 5.3.2. Descripción de las interacciones proteína - proteína

Mediante la aplicación de dinámica molecular se busca determinar las interacciones producidas en cada uno de los complejos y determinar cuáles son estables a lo largo del tiempo. A priori se han definido los contactos como aquellas interacciones que se mantuvieron por al menos un 50% de la simulación. De todas maneras se consideraron los contactos con una presencia mayor al 20% de la dinámica molecular, dado que estos contactos pueden ayudar a estabilizar un complejo en presencia de otro enlace más estable. Cabe destacar que el considerar la extensión de la duración de la simulación puede afectar directamente las interacciones propuestas, haciendo que interacciones que en principio se observan en menor proporción, pudiesen en realidad manifestarse durante una mayor fracción de la simulación. Todos los contactos se definieron a partir de los criterios que muestra la tabla 4.4. En la práctica los contactos se determinaron mediante el uso de las herramientas VMD y GROMACS en base a los criterios ya descritos.

#### 5.3.2.1. Descripción de los contactos en el homodímero de TLR4

La estructura del homodímero TLR4 no fue obtenida a partir del uso de algoritmos de acoplamiento, sino que fue modelada a partir de la secuencia de aminoácidos usando como modelo la estructura de TLR10. Aún así, se utilizó la dinámica molecular para observar los contactos presentes en el complejo y dar cuenta de si la estructura generada en el modelamiento muestra contactos relevantes en el tiempo. En efecto, los resultados de la dinámica molecular muestran que durante la simulación se observan en promedio por frame 8 interacciones de tipo hidrofóbico, 1 interacción por puente de hidrógeno y 1 interacción catión- $\pi$ . El listado de la totalidad de interacciones observadas se muestra en la Tabla 5.7. Las zonas de contacto pueden



ser observadas en la Figura 5.13.

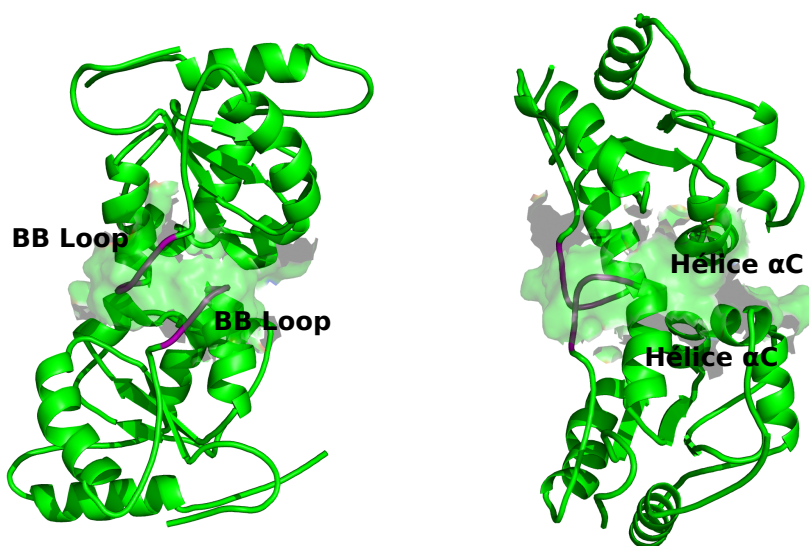


Figura 5.13: Zonas de contacto observadas vía MD para el complejo TLR4-MAL vía MD. *Mediante dinámica molecular se estudiaron las zonas de interacción entre los dominios TIR de los receptores TLR4 al formar un homodímero. Una vista frontal (izquierda) y lateral (derecha) se observan para las interacciones TLR4-TLR4. En la figura se han resaltado las superficies de los residuos que participan en la interacción. Los lazos BB han sido destacados en color morado*

El detalle de los contactos observados en el complejo muestra la aparición de hasta 12 interacciones de tipo hidrofóbico, 9 de las cuáles se presentan por sobre un 60 % de la simulación. Estas interacciones se producen entre los residuos de los lazos BB y la hélice  $\alpha$ C de ambos monómeros de TLR4 (ver Figura 5.14). Adicionalmente, se observa que dentro de los mismo sectores se producen 2 interacciones catión- $\pi$  que prevalecen un 50 % de la simulación. Por otro lado, los datos de la simulación indican la aparición de 2 puentes de hidrógeno (uno de mayor prevalencia que el otro). Esto últimos enlaces se producen entre las hélices  $\alpha$ E de los monómeros y entre la hélice  $\alpha$ C y el lazo BC. Es importante destacar que los contactos reportados muestran en su mayoría la participación de residuos que han sido objeto de estudios de mutagénesis.

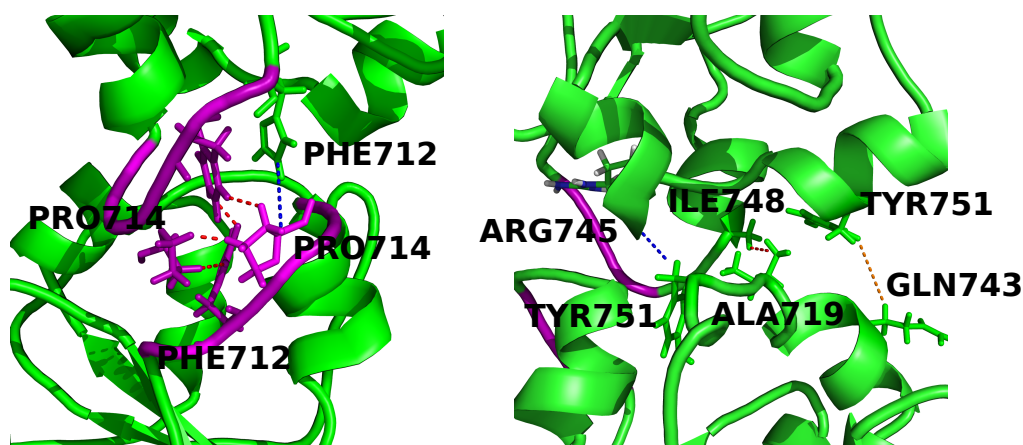


Figura 5.14: Detalle de la interacción en el complejo TLR4-TLR4. A la izquierda de la figura se observa como interactúan los lazos BB de los dominios TIR de cada monómero de TLR4, mientras que al lado derecho se observa para el mismo instante los contactos que se producen en los sectores de la hélice  $\alpha C$ . Los contactos unidos en rojo corresponden a interacciones hidrofóbicas, en azul se muestran contactos catión- $\pi$  y en naranja puentes de hidrógeno.

Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
ARG809	B	GLU824	A	63,5	RR809-810AA	II
TYR751	A	GLN743	B	27,6	YE751-752AA	I
Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
ARG745	A	TYR751	B	57,0	SR744-745AA	I
HIS708	B	PHE712	A	43,7	DFI711-713AAA	III
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
ALA719	A	ILE748	B	98,7		
ILE713	A	ILE718	B	64,1	I718A	I
ILE713	A	ILE748	B	94,7	DFI711-713AAA	I
ILE713	A	PHE749	B	72,1	DFI711-713AAA	I
ILE748	A	ILE748	B	86,7		
PHE712	A	PRO714	B	47,2	DFI711-713AAA, P714H	I, IV NFKB/IFNB
PHE712	A	ILE718	B	34,9	I718A	I
PHE712	A	PHE712	B	77,4	DFI711-713AAA	I
PHE712	A	PRO714	B	70,8	DFI711-713AAA, P714H	I, IV NFKB/IFNB
PRO714	A	PHE712	B	75,4	DFI711-713AAA, P714H	I, IV NFKB/IFNB
TYR751	A	ILE748	B	87,7	YE751-752AA	I
PRO714	A	PRO714	B	27,6	P714H	I, IV NFKB/IFNB

Efecto de las mutaciones (de acuerdo a evidencia experimental Ronnie *et al.* (2003) y Nuñez *et al.* (2007)):

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial) en la vía señalada IFN- $\beta$  o NF- $\kappa$ B

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

Tabla 5.7: Contactos reportados vía MD para la zona de interacción TLR4-TLR4. A: TLR4 cadena A, B: TLR4 cadena B

### 5.3.2.2. Descripción de los contactos en el complejo TLR4-MAL

Durante la simulación de la dinámica molecular del complejo TLR4-MAL se observaron los contactos que se generaron entre el receptor TLR4 y los modelos del adaptador propuesto tanto en este trabajo (Modelo P) como en el realizado por Nuñez (Modelo N). El número de contactos promedio observados por cada frame de la simulación se puede observar en la Tabla A.8 del Anexo A.8. El modelo generado por Nuñez presenta en general un mayor número de contactos promedio por frame que el modelo propuesto, en particular su modelo presenta promedios de 11 puentes de hidrógeno, 3 interacciones Cation- $\pi$ , 5 interacciones hidrofóbicas, 1 interacción iónica y 1 interacción aromática por frame. Por su parte el modelo propuesto presenta en promedio por frame la presencia de 4 puentes de hidrógeno, 1 interacción Cation- $\pi$  y 10 interacciones hidrofóbicas. Esta diferencia en el número de contactos es esperable de acuerdo al hecho de que los modelos presentan orientaciones y superficies de interacción diferentes en su unión al receptor.

Un resumen de los contactos persistentes (presentes sobre un 50% de la simulación) encontrados a lo largo de la dinámica molecular en el complejo TLR4-MAL se muestran en la Tabla 5.8. El listado de interacciones da cuenta que el modelo propuesto por Nuñez muestra un mayor número y tipos de contacto que el modelo generado en este trabajo a partir del algoritmo de acoplamiento GRAMM. En particular el modelo de Nuñez muestra 14 contactos de alta prevalencia (2 Hidrofóbicos, 2 Cation- $\pi$ , 1 Aromático, 2 Puentes Salinos y 7 Puentes de Hidrógeno) por su parte el modelo obtenido a partir de GRAMM solo muestra 5 contactos de alta prevalencia (3 Hidrofóbicos y 2 Puentes de Hidrógeno). Los datos presentados indican que la mayor parte de los contactos involucran a aminoácidos hidrofóbicos y polares cargados con un 40% en cada caso. Además en el caso del modelo propuesto los contactos provienen del BB *loop*, hélice  $\alpha$ B y CD *loop* del adaptador MAL y de los sectores del *loop* N-terminal, BB *loop* y hélice  $\alpha$ B de la cadena A de TLR4. Para el modelo presentado por Nuñez las áreas de interacción corresponden a *loop* N-terminal, BB y CD *loop*, lazo C-terminal y hélice  $\alpha$ B de MAL; y los lazos AA, CC y DD de la cadena A y hélice  $\alpha$ B,  $\alpha$ C y CD *loop* de la cadena B de TLR4 (ver Figura 5.15 y 5.16). Desde el detalle de los contactos se desprende que el modelo propuesto si bien se inserta en la zona de interacción del homodímero de TLR4, no presenta contactos de alta prevalencia con la cadena B del receptor TLR4, mientras que el modelo de Nuñez si presenta contactos duraderos con ambas cadenas del receptor lo que podría permitir una mayor estabilidad estructural del complejo en el tiempo. En efecto, se puede relacionar esta falta de contactos del MAL (P) con la cadena B en el complejo TLR4-MAL (P, Propuesto) al movimiento observado de MAL, en el que se el lazo BB se desplaza para dar paso una mayor participación del lazo CD en las interacciones. Por lo mismo el hecho de que el receptor MAL se una a TLR4 por medio de las dos cadenas en el complejo TLR4-MAL (N, Nuñez), muestra que esa conformación en particular favorece el anclaje del adaptador a la zona de interacción.

Durante la simulación también se observan otros contactos menos persistentes, que son detallados en el Anexo A.8. Estas interacciones son definidas como aquellas presentes durante un porcentaje del 20 - 50 % del total de la simulación. Si bien este tipo de interacciones presentan porcentajes bajos de existencia durante el total de la simulación, deben ser considerados dado que pueden ser relevantes en el caso de extenderse las simulaciones de dinámica molecular. Entre los contactos de esta categoría se aprecia la presencia de 14 y 4 nuevos puentes de hidrógeno para el modelo de Nuñez y el propuesto respectivamente. De la misma manera, en el modelo propuesto pueden observarse 3 nuevos contactos iónicos, 2 hidrofóbicos y 1 cation- $\pi$ . Estos contactos aparecen, en el caso del modelo de Nuñez, en las zonas estructurales del AA, CC, CD *loops* y hélice  $\alpha$ A para TLR4 y en la hélice  $\alpha$ A en conjunto con los lazos CD y BB de MAL. En el caso del modelo propuesto, las interacciones se centran en los AA, CC *loops*, hélices  $\alpha$ A y  $\alpha$ B de TLR4, y los *loops* CD, BB y hélices  $\alpha$ A y  $\alpha$ C de MAL.

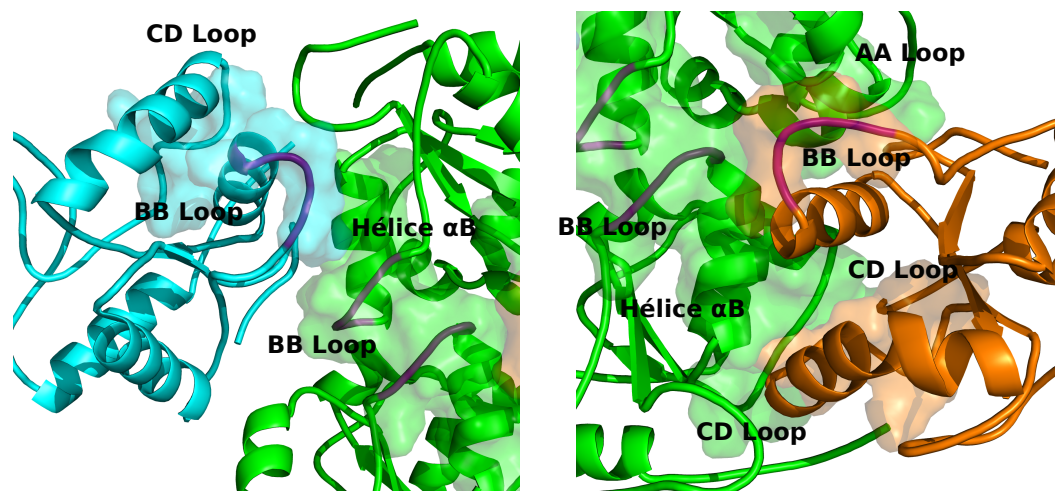


Figura 5.15: Zonas de contacto observadas vía MD para el complejo TLR4-MAL. A la izquierda de la figura se observa la interacción del dominio TIR de TLR4 (verde) con el modelo del dominio TIR de MAL propuesto por Salgado (cian). De manera similar, en la figura de la derecha se muestra el modelo generado por Nuñez (dorado) interactuando con el dominio TIR del receptor TLR4. En ambos casos se han resaltado los residuos que interactúan por medio de la representación de sus superficies. Los lazos BB han sido destacados en color morado

Parte de los residuos de TLR4 reportados en este trabajo como parte de contactos de alta permanencia, han sido objeto de estudios de mutagénesis determinándose que su mutación puede afectar la capacidad del receptor para transmitir río abajo la señal proinflamatoria o en otros casos puede producir la desestabilización del complejo. Esta evidencia habla acerca de la importancia de tales residuos en las interacciones del receptor con las proteínas adaptadoras. En la tabla se destacan los residuos que interactúan, detallándose el tipo de mutación que se estudio sobre el residuo y el efecto que tuvo la mutación sobre las capacidades de interacción y transmisión del receptor. Son 9 los contactos más relevantes según esta evidencia experimental y son en su mayoría reproducidos por el modelo de Nuñez. En particular el residuo W757 de TLR4, que participa en dos tipos de interacciones, parece ser importante dentro de esta vía, ya que su mutación provoca reducciones importantes de la señalización según evidencia experimental. Otros contactos que pueden ser relevantes son aquellos en los que participan residuos marcados con mutaciones de efecto tipo II (que afectan la unión con adaptadores) como en los contactos GLN758-ALA168, HIS724-SER131 y ARG763-GLU167. En todos estos casos, el residuo afectado es capaz de generar interacciones con carga o puentes de hidrógeno con permanencia por sobre el 80 % de la simulación, lo que indica la relevancia que pueden tener este tipo de contactos en la vía mediada por TLR4.

Dentro los contactos de menor permanencia en la simulación (Anexo A.8) se observa la presencia de alguno aminoácidos que resultan importantes de acuerdo al evidencia experimental. En particular los puentes de hidrógeno formados entre GLU167-GLN758, ARG763-GLU167, ARG787-GLU221, GLN863-ASP122 del complejo TLR4-MAL(N) pueden responder al efecto que tienen las mutaciones Q758A, R763A y Q683A en TLR4. Las mutaciones mencionadas provocan impedimentos en la unión del receptor TLR4 con adaptadores, por lo que resulta esperable su participación en las interacciones TLR4-MAL tal como se muestra en este modelo. De la misma manera, los contactos GLN683-SER113, HIS724-SER83 y ARG115-HIS740 presentes en el modelo TLR4-MAL(P) responden al efecto que tienen las mutaciones Q683A, H724A y QH739-740AA sobre la unión del adaptador MAL al receptor TLR4.

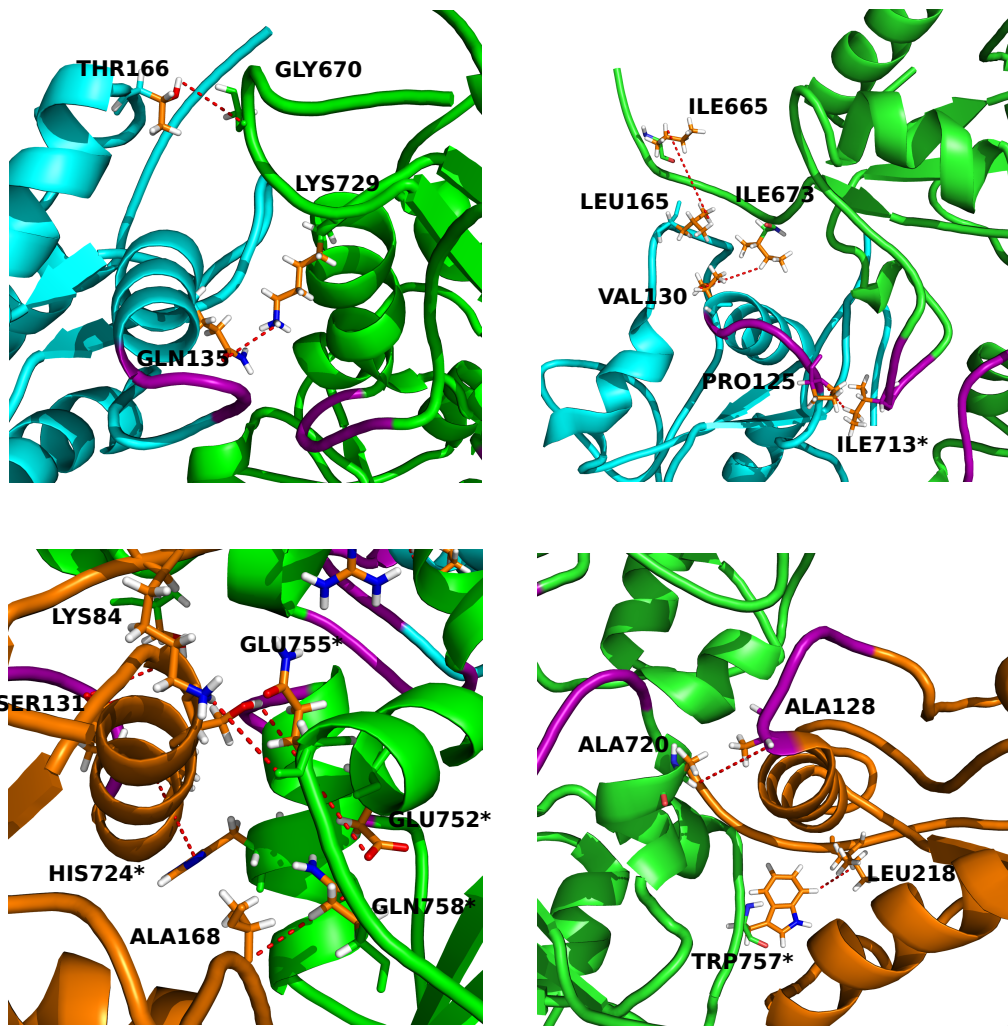


Figura 5.16: Detalle de la interacción en el complejo TLR4-MAL. Se han marcado algunas de las interacciones observadas entre los dominios TIR de MAL y TLR4. Los contactos que pueden ser validados mediante evidencia experimental han sido marcados con \*. En todos los casos el receptor TLR4 se muestra en verde, modelo de acoplamiento de MAL propuesto en este trabajo en cian y el planteado por Nuñez en naranja. Se han destacado los lazos BB en color morado.

Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
LYS84	N	GLN755	B	77.7	Q755A	I, III
SER131	N	HIS724	B	83.7	H724A	II
SER138	N	GLU752	B	57.1	YE751-752AA	I
ALA168	N	GLN758	B	58.5	Q758A	II
SER682	A	GLU132	N	54.5		
ARG780	A	GLU223	N	64.8		
ARG780	A	GLU223	N	64.1		
LYS729	A	GLN135	G	56.5	K729A	III
THR166	G	GLY670	A	57.8		
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ALA720	B	ALA128	N	73.2		
ILE665	A	LEU165	G	80.5		
ILE673	A	VAL130	G	64.9		
ILE713	A	PRO125	G	90.7	I713A	I
TRP757	B	LEU218	N	84.1	W757A	IV-RS-IFN- $\beta$
Contactos Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
HIS740	A	TYR86	N	100.0		
LYS84	N	TRP757	B	73.2	W757A	IV-RS-IFN- $\beta$
Contactos Aromáticos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
TYR86	N	HIS740	A	89.4		
Puentes Salinos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
GLU167	N	ARG763	B	52.6	R763A	II
ASP85	N	ARG780	A	56.3		

Efecto de las mutaciones (de acuerdo a evidencia experimental Ronnie *et al.* (2003) y Nuñez *et al.* (2007)):

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial) en la vía señalada IFN- $\beta$  o NF- $\kappa$ B

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

Tabla 5.8: Contactos encontrados vía MD en la zona de interacción TLR4-MAL según modelo de acoplamiento. A: TLR4 cadena A, B: TLR4 cadena B, N: MAL modelo Nuñez, G: MAL modelo propuesto

### 5.3.2.3. Descripción de los contactos en el complejo TLR4-MAL-MyD88

Para representar la interacción entre los dominio TIR del homodímero TLR4 y los adaptadores MAL y MyD88, se generaron dos modelos vía *docking* uno a partir del complejo TLR4-MAL de Nuñez y otro propuesto a partir de los resultados de GRAMM. A continuación se procede a describir solo las interacciones del complejo TLR4-MAL-MyD88 (P, Propuesto). Hasta este momento no se ha desarrollado la dinámica molecular para el complejo TLR4-MAL-MyD88 (N, Nuñez) por lo que no se describirán los contactos por medio de esta estrategia. El modelo de interacción analizado para el adaptador secundario MyD88 muestra la unión del adaptador a la zona de interacción entre MAL y la cadena B de TLR4. A partir de la dinámica molecular se encontraron en promedio por frame 5 puentes de hidrógeno y un contacto iónico entre el receptor TLR4-MyD88, mientras que en la interacción MAL-MyD88 se identificaron, en promedio por

frame, 4 puentes de hidrógeno, 1 contacto catión- $\pi$  y 1 interacción hidrofóbica (ver Anexo A.8, Tabla A.10 y A.11). Se identificaron 12 contactos con alta prevalencia durante la simulación, 6 entre TLR4(B)-MyD88 y 6 MAL-MyD88. Dentro de los contactos reportados se encuentran 9 puentes de hidrógeno, 1 interacción hidrofóbica, 1 enlace Cation- $\pi$  y 1 enlace iónico. En estos contactos se observa una alta participación de aminoácidos polares cargados, alcanzando un 58 % del total residuos participantes en interacciones de alta prevalencia (un 60 % de estos cargados positivamente). De los contactos detectados vía dinámica molecular se tiene evidencia experimental de la importancia del residuo LYS773 en la estabilidad de los complejos generados (ver Figura 5.17), por lo que se debe analizar la implicancia de este residuo u otros análogos en naturaleza (carga positiva) sobre la estabilidad del sistema.

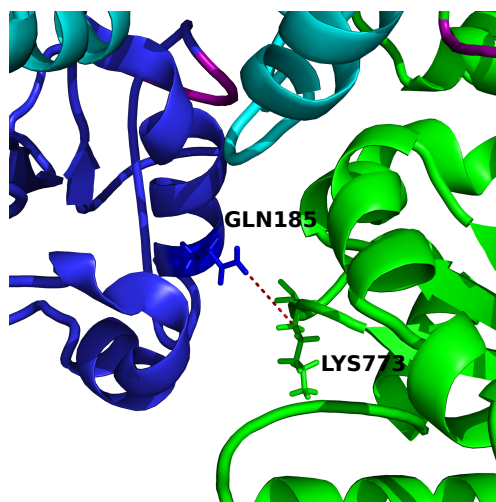


Figura 5.17: Contactos validados experimentalmente para el modelo TLR4-MAL-MyD88. Se destaca la interacción del residuo LYS773 cuya mutación por ALA causa pérdida de estabilidad en interacciones del dominio TIR. Se observa en la figura TLR4 (verde), MAL (cian) y MyD88 (azul). En la figura se han resaltado las superficies de interacción.

Los contactos entre TLR4 y MyD88 se generaron entre las zonas DD *loop* y  $\alpha$ C de TLR4, y la hélice  $\alpha$ A de MyD88. Por otro lado, las interacciones entre el adaptador MAL y MyD88 se produjeron entre las zonas hélice  $\alpha$ A y  $\alpha$ C de MyD88, y los lazos AB, N-terminal y C-terminal de MAL (Figura 5.18). Esto último puede restar validez al modelo, ya que cabe la duda de si ese sitio de unión puede estar impedido por la membrana plasmática. Además el hecho de que participen los lazos terminales en la interacción resalta otro problema que puede presentar este modelo, ya que estos segmentos presentan cierto grado de movilidad según lo observado durante la simulación. Un segundo modelo generado para MyD88, que une a la cadena A de TLR4 y el adaptador MAL del modelo de Nuñez, no muestra este último tipo de uniones con los lazos terminales, pero hasta este punto no se ha desarrollado la dinámica molecular ni se han definido los contactos para ese complejo.

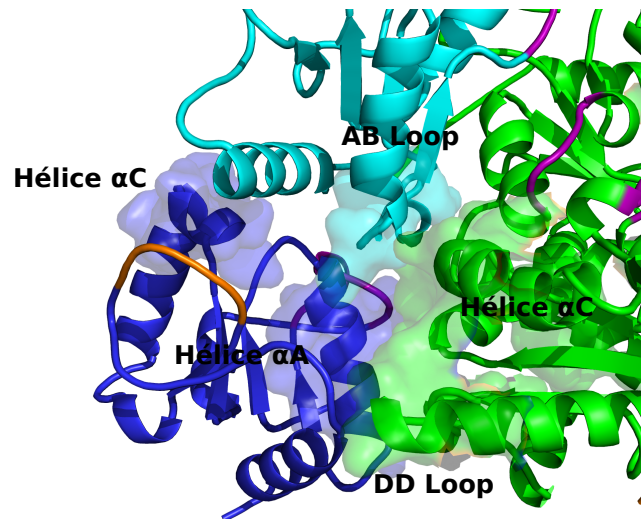


Figura 5.18: Zonas de contacto observadas mediante dinámica molecular para el complejo TLR4-MAL-MyD88. TLR4 (verde), MAL (cian), MyD88 (azul). En la figura se han resaltado las superficies de interacción. Los lazos BB han sido destacados en color morado

Otros contactos de menor permanencia en la simulación (Anexo A.8) incluyen la presencia de 3 puentes de hidrógeno en la interacción TLR4-MyD88, además de 4 puentes de hidrógeno y 1 contacto hidrofóbico que se forman entre MAL y MyD88 (Modelo 1). Estas interacciones con menor presencia involucran al CC loop y  $\alpha C$  de TLR4; hélice  $\alpha A$ , lazos AA y BB de MyD88; y lazos CC y terminales del modelo de MAL.



Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ARG181	M	HIS740	B	82.4	K773A	III
ARG181	M	HIS740	B	68.1		
GLN185	M	LYS773	B	63.8		
HIS740	B	GLN177	M	66.1		
GLN772	B	GLU178	M	74.8		
GLN174	M	SER113	G	68.1		
GLY202	M	SER79	G	65.1		
ARG215	G	ASP227	M	53.2		
TYR216	G	ASP227	M	59.5		
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
LEU217	G	ILE173	M	76.8		
Contactos Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ARG181	M	HIS740	B	100.0		
Puentes Salinos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ASP227	M	ARG215	G	59.9		

Efecto de las mutaciones (de acuerdo a evidencia experimental Ronnie *et al.* (2003) y Nuñez *et al.* (2007)):

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial) en la vía señalada IFN- $\beta$  o NF- $\kappa$ B

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

Tabla 5.9: Contactos encontrados en la zona de interacción TLR4-MAL-MyD88 vía MD según modelo de acoplamiento. A: TLR4 cadena A, B: TLR4 cadena B, N: MAL modelo Nuñez, G: MAL modelo propuesto, M: MyD88 Modelo 1

#### 5.3.2.4. Descripción de los contactos en el complejo TLR4-TRAM

Los resultados de la simulación del complejo TLR4-TRAM (N) muestran un promedio de enlaces por frame de 15 puentes de hidrógeno, 11 interacciones catión- $\pi$ , 4 hidrofóbicas un contacto aromático. En contraste de los 7 puentes de hidrógeno y 2 enlaces catión- $\pi$  promedio por frame encontrados en el complejo TLR4-TRAM propuesto (Ver Anexo A.8, Tablas A.12 y A.13). Se reportaron 7 contactos de alta prevalencia (4 puentes de hidrógeno, 1 contacto hidrofóbico, 2 interacción cation- $\pi$ ) para el modelo propuesto, mientras que el modelo de Nuñez mostró 20 contactos de estas características (11 puentes de hidrógeno, 4 contactos hidrofóbicos, 4 interacciones cation- $\pi$  y 1 iónica). De esta manera, al igual que en el caso de TLR4-MAL, el modelo de Nuñez reproduce muchos más contactos que el otro modelo propuesto en este trabajo. Además este último modelo responde en mayor grado a la evidencia experimental generada mediante la mutación sitio dirigida de residuos. Esta información se muestra en la Tabla 5.10 y corresponden a estudios realizados por Ronni *et al.* (2003) y Nuñez *et al.* (2007) en lo que se demuestra el efecto sobre la estabilidad de los complejos que provocan ciertas mutaciones del receptor TLR4. Parte de los contactos relevantes puede observarse en la Figura 5.19.

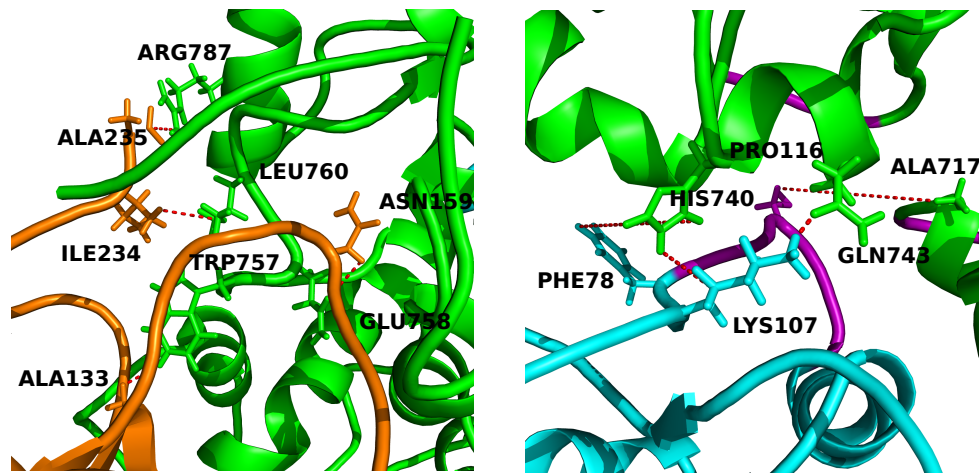


Figura 5.19: Detalle de la interacción en el complejo TLR4-TRAM. Se han destacado, mediante la representación de sus residuos, algunas de las interacciones observadas entre los dominios TIR de TRAM y TLR4. Los contactos corresponden a aquellos que involucran a residuos que han sido validados experimentalmente. En todos los casos el receptor TLR4 se muestra en verde, modelo de acoplamiento de TRAM propuesto en este trabajo en cian y el planteado por Nuñez en naranja. Se han destacado los lazos BB en color morado.

Al observar la naturaleza de los residuos participantes en los distintos modelos se observa que cerca de un 30 % son de carácter hidrofóbico y un 40 % son residuos polares con carga. Estructuralmente, los sectores involucrados en las interacciones para el modelo de Nuñez son: BC, DD *loop*, lazos terminales y hélice  $\alpha$ D en TRAM; hélice  $\alpha$ B, el lazo CD y los lazos terminales de la cadena A de TLR4, y el DD *loop* de la cadena B de TLR4. En el caso del modelo propuesto, interactúan los lazos AB, BB y terminales de TRAM, con el BB *loop* (cadena B) y la hélice  $\alpha$ A (cadena A) de TLR4 (Figura 5.20).

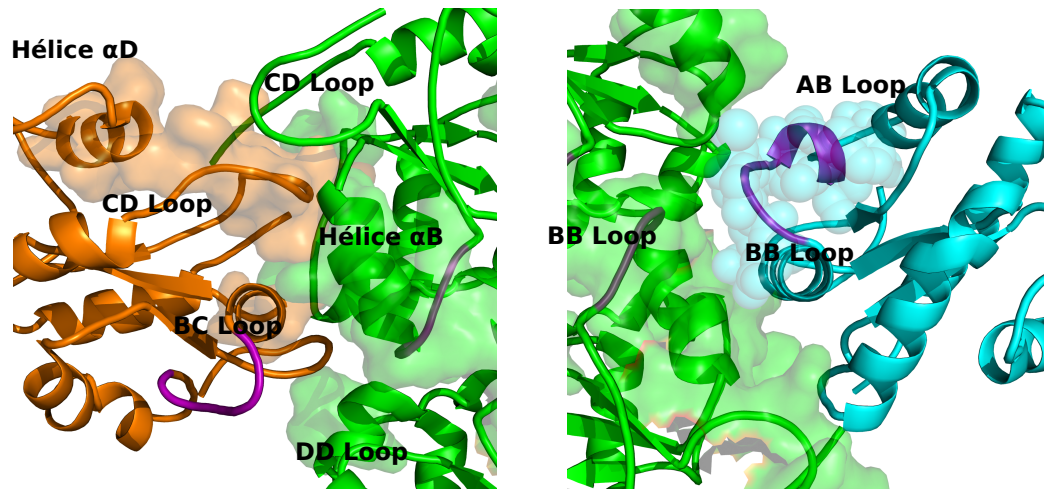


Figura 5.20: Zonas de contacto observadas mediante dinámica molecular para el complejo TLR4-TRAM. A la izquierda de la figura se observa la interacción del dominio TIR de TLR4 (verde) con el modelo del dominio TIR de TRAM propuesto por Nuñez (dorado). De manera similar, en la figura de la derecha se muestra el modelo de TRAM generado por Salgado (cian), interactuando con el dominio TIR del receptor TLR4. En ambos casos se han resaltado los residuos que interactúan por medio de la representación de sus superficies tridimensional. Los lazos BB han sido destacados en color morado

Entre los contactos con menor prevalencia (Anexo A.8, Tabla A.13) se observan hasta 10 nuevos puentes de hidrógeno, 3 interacciones hidrofóbicas y 2 catión- $\pi$  en el modelo de Nuñez. Estas interacciones involucran principalmente sectores del CD *loop* de TLR4 en contacto con los lazos terminales y CD de TRAM. Para el modelo propuesto, se detectan 8 nuevos puentes de hidrógeno y 2 interacciones del tipo catión- $\pi$ . Estructuralmente, en este último caso, los contactos involucran a los *loops* AA, CC y terminales de TLR4 interactuando con residuos pertenecientes a los lazos BB y CD del adaptador TRAM. De la totalidad de contactos de menor prevalencia que participan en las interacciones, existen algunos que logran responder a evidencia experimental. En particular los enlaces GLN683-CYS117, TRP821-ARG163, GLN755-GLU74, ARG763-PHE233, LEU760-ILE234 del complejo TLR4-TRAM(N) muestran la participación de residuos de TLR4 cuya mutación interrumpe la señalización en esta vía. Lo mismo ocurre con el residuo GLN683 que participa en uno de los puentes de hidrógeno GLN683-GLU114 observados en el modelo TLR4-TRAM(P)

Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
LYS107	G	GLN743	A	60.4	IQ742-743AA  H728D Q758A	I  II, IV-RS-NF $\kappa$ B-IFN- $\beta$ II
ARG163	G	GLU824	B	50.2		
ARG717	A	GLY118	G	58.8		
SER156	N	HIS728	A	80.1		
ASN159	N	GLN758	A	97.0		
ARG231	N	GLU824	A	63.8		
ARG780	B	GLU75	N	90.7		
GLN782	B	GLU75	N	88.7		
GLU671	A	TYR154	N	92.4		
ASN672	A	SER156	N	66.1		
ARG787	A	ALA235	N	61.8	RL787-788AA	II
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ALA717	B	PRO116	G	73.2	FL759-760AA	II
LEU760	A	ILE234	N	73.5		
TRP757	A	ILE235	N	68.2		
TRP757	A	ALA133	N	58.3		
TRP757	A	MET158	N	57.6		
Contactos Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ARG669	A	PHE187	N	60.9	W757A	IV-RS-IFN $\beta$  IV-RS-IFN $\beta$
ARG780	B	PHE69	N	71.5		
HIS165	N	TRP757	A	99.7		
HIS740	A	PHE78	G	63.2		
LYS166	N	TRP757	A	85.8		
LYS107	G	HIS740	A	70.2		
Interacciones Aromáticos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
TRP757	A	HIS165	N	97.0	W757A	IV-RS-IFN $\beta$

Efecto de las mutaciones (de acuerdo a evidencia experimental Ronnie *et al.* (2003) y Nuñez *et al.* (2007)):

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial) en la vía señalada IFN- $\beta$  o NF- $\kappa$ B

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

Tabla 5.10: Contactos encontrados en la zona de interacción TLR4-TRAM vía MD según modelo de acoplamiento. A: TLR4 cadena A, B: TLR4 cadena B, N: TRAM modelo Nuñez, G: TRAM modelo propuesto

### 5.3.2.5. Interacciones conservadas en la vía mediada por TLR4 y adaptadores

A partir del alineamiento de las secuencias de TLR4 y adaptadores se busca determinar el nivel de conservación de los contactos y la existencia de segmentos estructurales con preferencia por algún tipo de interacción dentro de la vía proinflamatoria mediada por TLR4. El alineamiento se realizó con la herramienta PRALINE, la que utiliza una estrategia de alineamiento denominada *Homology Extended Multiple Sequence Alignment* en la que la información de las secuencias del set a estudiar, es complementada con la recolección de secuencias homólogas [104]. La ventaja de esta estrategia es que se cuenta con una mayor cantidad de

información específica acerca de las posiciones de los residuos. Tanto el algoritmo como el desempeño de la estrategia utilizada por PRALINE puede revisarse en Simossis *et al.* (2005) [104]. Los resultados del alineamiento generado por PRALINE se observan en las Figuras 5.21 y 5.22. Sobre los alineamientos producidos por PRALINE se destacaron los contactos obtenidos a partir de las simulaciones de dinámica molecular para los complejos TLR4-MAL (N y P), TLR4-MAL-MyD88 (P) y TLR4-TRAM (N y P). Adicionalmente se resaltan aquellos residuos que han sido objeto de estudios de mutagénesis.

De manera general, a partir de los resultados del alineamiento se desprende la existencia de sectores en los que se acumulan contactos de manera transversal dentro de esta familia de proteínas. Estos sectores corresponden al segmento previo al BOX I que comprende los residuos 11-15 (según alineamiento), BB *loop*, hélice  $\alpha$ C, CD *loop* y el segmento C-terminal que comprende los residuos 170-174 (según alineamiento). Si bien el BOX I es un fragmento altamente conservado dentro de esta familia de proteínas, los residuos 11-15 previos al BOX no lo son y de hecho presentan una puntuación media baja de conservación según el algoritmo utilizado. El BB *loop* (BOX II) que muestra una alta concentración de contactos encontrados (para los modelos propuestos), presenta un nivel de conservación alto y ya ha sido indicado como relevante dentro de las interacciones de esta familia de proteína [8, 20, 21, 33, 42]. Lo mismo ocurre con el CD *loop* que también muestra una gran acumulación de contactos (principalmente modelo Nuñez) y que presenta un nivel de conservación medio o alto según el sector que se observe. Este segmento también ha sido señalado como participante dentro de las interacciones de esta vía [21]. El caso del sector 170-174 es distinto, por cuanto muestra un nivel de conservación muy bajo, lo que puede indicar que los contactos acumulados en este sector pueden ser circunstanciales y por lo mismo resultar de poca relevancia dentro de la vía mediada por TLR4.

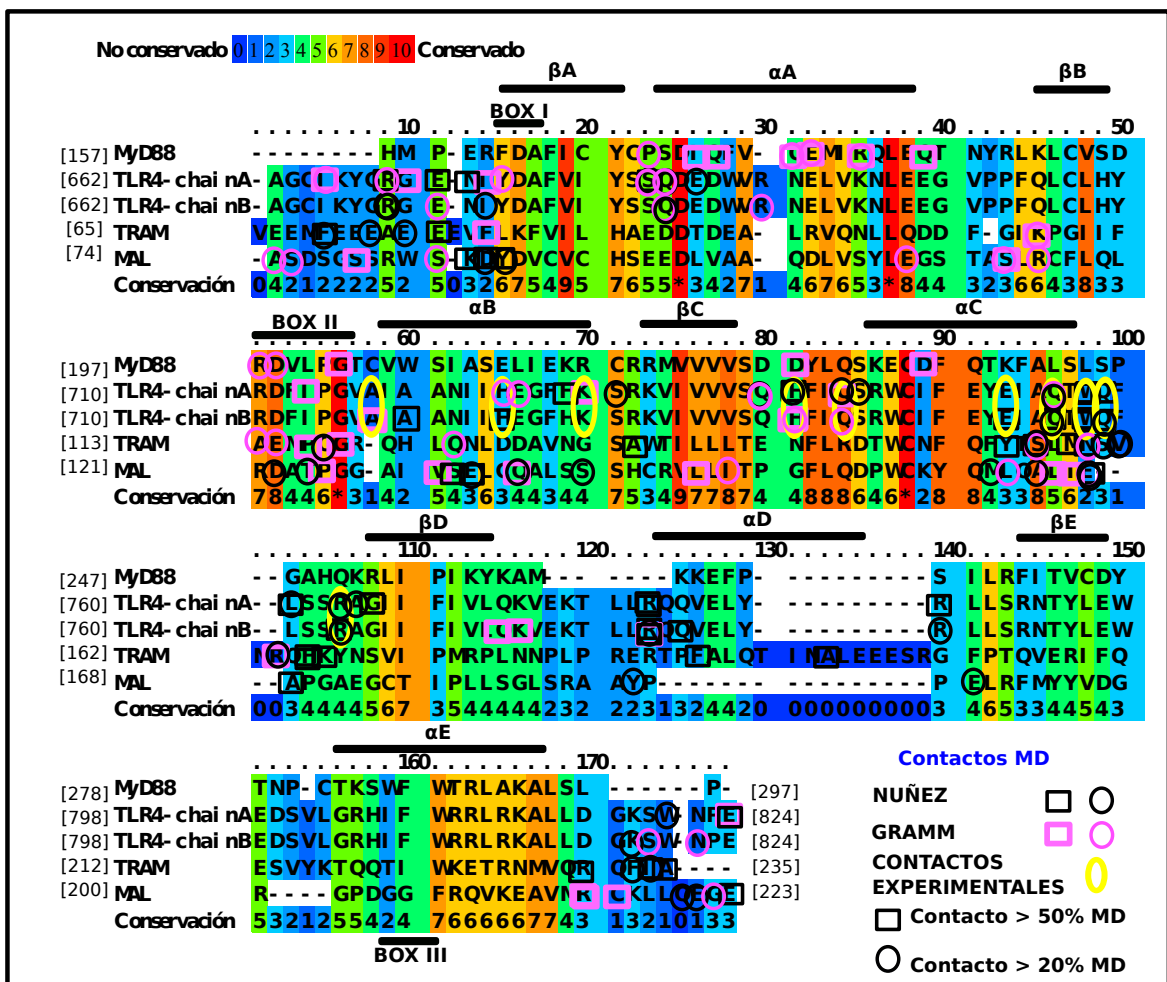


Figura 5.21: Alineamiento de las secuencias para el estudio de conservación de residuos (PRALINE). □: Contactos de alta prevalencia, ○: Contactos prevalencia media

Por otro lado, mediante la herramienta PRALINE, se estudio la hidrofobicidad de las distintas proteínas que participan en esta vía. De la figura 5.22 es posible dar cuenta que los sectores que muestran una acumulación de contactos, en general no presentan un carácter de hidrofobicidad marcado, si no más bien muestran características intermedias de hidrofobicidad. Excepcionalmente el CD loop muestra un caracter principalmente hidrofílico.

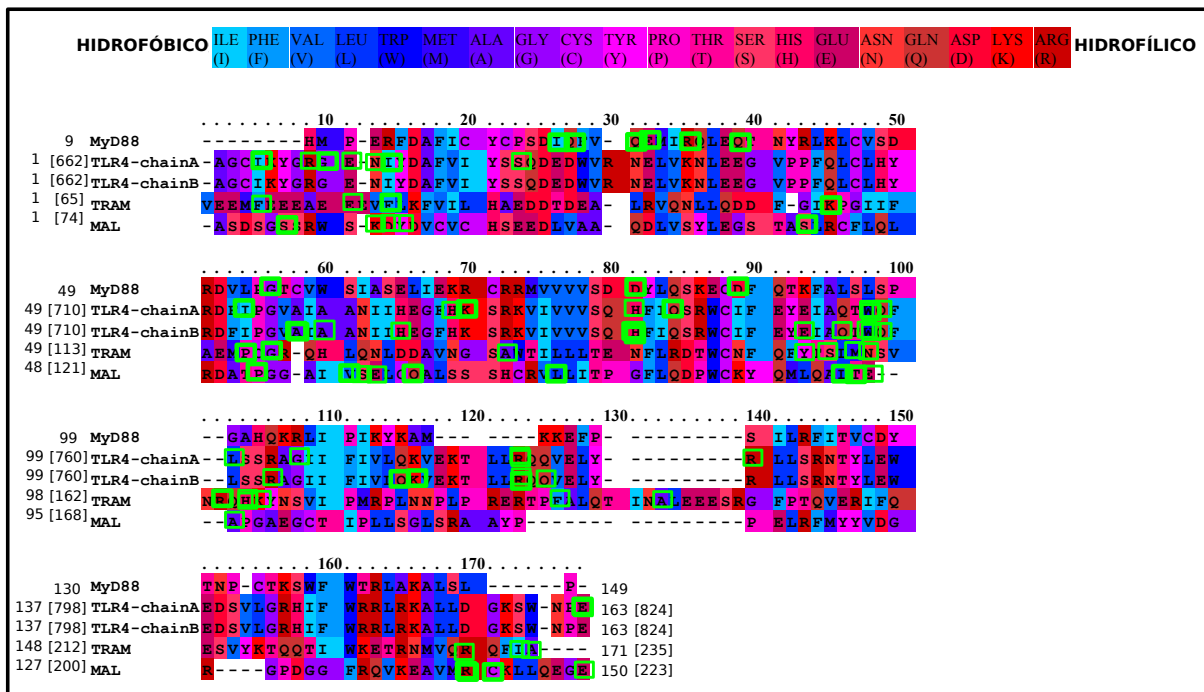


Figura 5.22: Alineamiento de las secuencias y análisis de la hidrofobicidad de los residuos (PRALINE)

En particular, dentro del sector de los residuos 11-15 (según alineamiento), se aprecia que el aminoácido en la posición 11 (correspondiente a PRO en MyD88, GLU en TLR4 y TRAM, y SER en MAL) tiene un nivel de conservación medio y un nivel de hidrofobicidad similar según la escala que muestra la Figura 5.22. El otro extremo conservado dentro del sector mencionado es el correspondiente al residuo 15 (según alineamiento) que forma parte del BOX I y que corresponde en todos los casos a un residuo hidrofóbico (PHE en MyD88, TYR en TLR4 y MAL, LEU en TRAM). Al cotejar los contactos observados en esta zona es posible observar que el residuo 11 (según alineamiento) en cada caso tiene la característica de formar puentes de hidrógeno que se mantienen por gran parte de la simulación (GLU75 en TRAM, GLU671 en TLR4, SER83 en MAL), mientras que los residuos 14-15 muestran interacciones de tipo Hidrofóbico o Cation- $\pi$ . Cabe destacar que en el caso de los puentes de hidrógeno se tiene que todas las interacciones se encuentran al menos un 75% de la simulación, esto puede indicar la importancia de la presencia de este puente de hidrógeno para la unión y estabilidad de este sector y podría permitir además la formación de otras interacciones como las que presenta el residuo 15, como por ejemplo contactos hidrofóbicos TYR674-ALA128 y MET161-ILE673 del complejo TLR4-MAL, entre otros. La relevancia particular de un puente hidrógeno, al punto de resultar crucial para la unión estable receptor-ligando, es un fenómeno ha sido observado en otro tipo de interacciones proteína-proteína [105, 106].

Una segunda zona en la que se acumulan contactos es la zona del BB *loop* (residuos 54-58 según alineamiento), que ha sido ampliamente señalada como clave en las interacciones de esta familia de proteínas [8, 20, 21, 33, 42]. La mayor parte de los contactos concentrados en este sector, corresponden a contactos generados por el modelo propuesto y no el planteado por Nuñez. En general este segmento posee una conservación media según el alineamiento donde destacan los residuos PRO y GLY, además no posee un carácter hidrofóbico/hidrofílico definido (ver Figura 5.22). Los contactos capturados a partir del modelo propuesto para esta zona son principalmente de carácter hidrofóbico. Estos contactos tienen la particularidad de generarse entre los BB *loops* de TLR4 y las proteínas adaptadoras MAL y TRAM respectivamente. En particular, el residuo de PRO52 del BB *loop* tanto de MAL como TRAM aparece involucrado en las interacciones hidrofóbicas descritas, en ambos casos interactuando con un residuo en TLR4 para el cual se

ha estudiado que su mutación genera receptores no funcionales (Ver tabla MUTANTES). Los puentes de Hidrógeno aparecen como un segundo tipo de contacto relevante dentro de este segmento. Los residuos GLY y ALA pertenecientes a MAL, TRAM y MyD88, presentes en las posiciones 56-58 del alineamiento, tienen la capacidad de generar puentes de hidrógeno con otros sectores. En el caso de los adaptadores MyD88 y MAL se forman puentes de hidrógeno con regiones cercanas al BOX I de TLR4, mientras que para el caso del adaptador TRAM se forman puentes de hidrógeno con el BB *loop* de TLR4. A partir de lo anterior es posible que, al igual que en el caso del BOX I, la presencia de un enlace de hidrógeno establezca el complejo y permita la generación de un segundo contacto, en este caso hidrofóbico. Cabe destacar que, en este caso, son los contactos hidrofóbicos los que tienen una mayor prevalencia durante la simulación con porcentajes por sobre el 70 %, mientras que la permanencia de los puentes de hidrógeno no supera el 65 %.

Otra de las estructuras consideradas relevantes a partir del último análisis son la zonas del lazo CD y parte de la hélice  $\alpha$ C. El lazo CD ha sido destacado, en publicaciones anteriores, como una estructura relevante en las interacciones de los adaptadores MAL/TRAM con el receptor TLR4 [20, 21]. Efectivamente, los resultados del alineamiento y la dinámica molecular muestran que estos sectores generan interacciones tanto en los complejos TLR4-TRAM como TLR4-MAL. Desde la dinámica molecular de los complejos se observan importantes interacciones que involucran directamente los lazos CD. En el caso del complejo TLR4-TRAM(N) la interacción entre lazos CD se observa en la formación de los enlaces TRP757-HIS165, TRP757-LYS166 (enlaces cation- $\pi$ ), GLN758-ASN159 (puente de hidrógeno), ALA754-VAL161 (interacción hidrofóbica). Interacciones similares se observan en el complejo TLR4-MAL(N) por medio de los enlaces TRP757-LYS84 (enlace cation- $\pi$ ), GLN758-ALA168, GLN758-GLU167 (puentes de hidrógeno). En el caso de las interacciones del receptor MyD88, también se observa la interacción del lazo CD de TLR4 mediante el enlace GLN755-ARG197, aunque en este caso el contacto se produce con el BB *loop* de MyD88. Los contactos anteriormente destacados muestran que en ambos casos, se observa la presencia de un enlace cation- $\pi$  y un puente de hidrogeno de alta prevalencia entre los lazos CD del adaptador MAL o TRAM con el receptor TLR4. En ambos casos los enlaces involucran específicamente los residuos TRP757 y GLN758 del lazo CD de TLR4. Esta evidencia muestra la relevancia que tienen el lazo CD en las interacciones de esta vía, pero más aún hablan acerca del comportamiento de hot spot que podrían presentar estos residuos. Más aún, la evidencia experimental indica que la mutación de estos residuos afecta las interacciones entre el receptor TLR4 y los adaptadores MAL/TRAM.

#### **5.3.2.6. Estimación de la Energía Libre de los complejos proteína-proteína observados en la vía mediada por TLR4.**

Para analizar el componente energético de la interacción de los complejos TLR4-MAL, TLR4-TRAM y TLR4-MAL-MyD88, se evaluaron los distintos componentes de la energía libre durante los últimos dos nanosegundos de cada simulación. Esta ventana de tiempo corresponde a aquella en que las estructuras de los complejos mostraron mayor estabilidad de acuerdo al RMSD  $C\alpha$ . El análisis energético se basa en la metodología MM-PBSA, donde la energía libre de unión ( $\Delta G_{union}$ ) de cada complejo, es calculada a partir de 3 componentes energéticos: energía mecánica molecular ( $\Delta E_{MM}$ ), energía de solvatación ( $\Delta G_{solv}$ ) y la contribución entrópica del sistema ( $-T\Delta S$ ). La energía mecánica molecular es el término que comprende la contribución energética interna del sistema  $\Delta E_{int}$  (enlaces, ángulos, torsiones), electrostática  $\Delta E_{elec}$  y de no enlace o Van der Waals  $\Delta E_{vdW}$ . Por su parte el término de solvatación considera las contribuciones electrostáticas  $\Delta G_{solv,polar}$  y no polares  $\Delta G_{solv,apolar}$ . La estimación de la energía libre de los complejos se observa en la Tabla 5.11.

Los datos muestran que para el caso del complejo TLR4-MAL, el valor estimado de la energía libre es de -2423.6 [kJ/mol] para el modelo propuesto y -1176.2 [kJ/mol] para el modelo de Nuñez. En ambos casos el signo de la energía libre es consistente con la estabilidad observada para el sistema. Se observa además que la formación de los complejos es un proceso entálpicamente favorable pero entrópicamente desfavorable, en



Tipo de Contribución	TLR4-MAL(P)	TLR4-MAL(N)	TLR4-TRAM(P)	TLR4-TRAM(N)
$\langle E_{int} \rangle$	-1958.9	-1463.2	699	937.1
$\langle E_{elec} \rangle$	-1308.2	-992.3	-2694.7	868.2
$\langle E_{vdw} \rangle$	-665.5	-709.5	-902.1	-509.2
$\langle E_{mm} \rangle$	-3932.5	-3165	-2897.8	1296.2
$\langle \Delta G_{solv,apolar} \rangle$	-64.5	-75.8	-74.1	-62.2
$\langle \Delta G_{solv,polar} \rangle$	1358	1852.3	2903.3	1652.2
$\langle \Delta G_{solv} \rangle$	1293.5	1776.5	2829.1	1590
-T $\Delta S$	215.4	212.3	209.9	209.8
$\Delta G_{union}$	-2423.6	-1176.2	141.3	3096.2

P: Modelos propuestos. N: Modelos Nuñez. Cálculo de energía libre considera:  $\Delta G_{union} = \langle E_{MM} \rangle + \langle G_{solv} \rangle - T^* \Delta S$ .  $\langle \rangle$ : Valor promedio a lo largo de la simulación. Valores en [kJ/mol]

Tabla 5.11: Estimación de la energía libre de los complejos

el que finalmente la magnitud de ambos términos indica que la unión de los dominios TIR de TLR4 y MAL es un proceso espontáneo ( $\Delta G < 0$ ). Adicionalmente, los datos muestran que el componente mecánico de la energía ( $E_{int}$ ,  $E_{elec}$  y  $E_{vdw}$ ) y el componente apolar de la solvatación ( $\Delta G_{solv,apolar}$ ) favorecen la interacción entre el receptor TLR4 y el adaptador MAL. Por otro lado, en ambos casos se observa que la componente polar de la solvatación ( $\Delta G_{solv}$ ) resulta desfavorable para la interacción TLR4-MAL. Este término se refiere a la energía asociada con la transferencia del soluto desde un medio con una constante dieléctrica baja ( $\epsilon = 2$ ) a un medio continuo con la constante dieléctrica de agua ( $\epsilon = 78.4$ ). Lo anterior implica que las moléculas de agua pueden tener un efecto desestabilizador sobre el complejo, no favoreciendo su formación. Finalmente, en el caso de ambos complejos se observa que la entropía no favorece la formación de los complejos. En ambas situaciones se aprecia que tanto el receptor TLR4 como el adaptador MAL presentan una mayor movilidad en sus estados libres en comparación con sus estructuras dentro del complejo. Esta pérdida de movimiento, por la fijación relativa de ciertos residuos y el ordenamiento de la estructura del complejo, hace que la entropía vibracional del complejo sea menor que la de sus componentes por separado. El hecho de que la entropía presente un valor positivo no corresponde a un evento inesperado. En efecto, se pueden observar tendencias similares en otros estudios que utilizan la metodología MM-PBSA para la estimación de energía libre [51,56]. El detalle de los componentes de la estimación de energía libre se aprecia en la siguiente tabla.

Tipo Contribución	Ligando (MAL)		Receptor (TLR4)		Complejo (TLR4-MAL) [P]		Complejo (TLR4-MAL) [N]	
	Promedio	$\sigma$	Promedio	$\sigma$	Promedio	$\sigma$	Promedio	$\sigma$
$\langle E_{int} \rangle$	21823,8	197,0	44608,7	257,2	64470,9	216,4	64908,7	423,3
$\langle E_{elec} \rangle$	-17537,3	497,2	-45624,0	805,4	-64316,6	851,8	-63781,9	786,1
$\langle E_{vdw} \rangle$	-1208,4	91,7	-3137,7	126,1	-5023,1	168,2	-4996,1	170,4
$\langle E_{MM} \rangle$	-3077,9	557,8	-4153,0	953,6	-4868,8	1019,4	-3869,4	1020,8
$\langle G_{solv,apolar} \rangle$	228,7	6,7	493,2	7,4	653,8	11,3	651,0	8,4
$\langle G_{solv,polar} \rangle$	-5326,5	254,4	-11871,9	407,4	-16028,5	338,4	-15692,9	347,2
$\langle G_{solv} \rangle$	-5097,8	250,0	-11378,7	400,0	-15374,7	330,6	-15041,9	342,4
-TS	-552	—	-268,2	—	-489,0	—	-321,3	—
$\langle G_{union} \rangle$	-2571,7	—	-15799,9	—	-20733,2	—	-19232,6	—

P: Modelos propuestos. N: Modelos Nuñez. Cálculo de energía libre considera:  $\Delta G_{union} = \langle E_{MM} \rangle + \langle G_{solv} \rangle - T^* \Delta S$ .  $\langle \rangle$ : Valor promedio a lo largo de la simulación. Valores en [kJ/mol]

Tabla 5.12: Detalle de componente energética por estructura protéica para la estimación de energía libre del complejo TLR4-MAL

En el caso del complejo TLR4-TRAM, la estimación de la energía libre para el complejo propuesto es de 141.3 [kJ/mol] mientras que para el de Nuñez es de 3096.2 [kJ/mol]. Al contrario de lo observado con MAL, el signo de la estimación de energía libre muestra que los complejos no tienen conformaciones termodinámicamente favorables. Esto se observa con mayor propiedad en el modelo de Nuñez cuyo valor positivo es incluso un orden de magnitud mayor que en el modelo propuesto. El hecho de que la energía libre de Gibbs sea positiva implica que los dominios TIR del homodímero de TLR4 y el monómero de TRAM no se unen de manera espontánea, por lo que de acuerdo a los valores reportados en este trabajo no es posible la formación del complejo TLR4-TRAM en las conformaciones estudiadas. Si bien es posible que la estimación realizada para la energía libre de unión no sea del todo precisa, también es necesario considerar la posibilidad de que quizás el complejo TLR4-TRAM no puede formarse de manera independiente y requiera de la participación de otro adaptador.

En particular, las diferencias más importantes en la estimación de la energía de estos modelos se encuentra en el cálculo de la energía mecánica. El modelo propuesto para TLR4-TRAM muestra que la componente de energía mecánica resulta favorable para la interacción, mientras que este término aparece como desfavorable en la interacción del complejo de Nuñez. Otra diferencia considerable se puede apreciar en la contribución energética provocada por la solvatación polar que presenta un aumento considerable en el modelo propuesto con un valor de 2903.3 [kJ/mol], mientras que en el segundo modelo se alcanza un valor de 1652.2 [kJ/mol]. Finalmente, y de manera similar al caso del adaptador MAL, se observa que la entropía no favorece la formación de los complejos, ya que dentro del complejo se produce la fijación y reordenamiento de algunas cadenas de residuos, por lo cual la entropía vibracional resulta menor que en el caso del receptor y ligando. Adicionalmente se muestra el detalle del cálculo de energía libre para el complejo TLR4-TRAM. La Tabla 5.13 muestra el valor energético de cada componente considerada para el adaptador TRAM, el receptor TLR4 y el complejo formado.

Tipo Contribución	Ligando (MAL)		Receptor (TLR4)		Complejo (TLR4-TRAM) [P]		Complejo (TLR4-TRAM) [N]	
	Promedio	$\sigma$	Promedio	$\sigma$	Promedio	$\sigma$	Promedio	$\sigma$
$\langle E_{int} \rangle$	22956,6	360,0	44608,7	257,2	68341,9	228,5	68444,8	417,7
$\langle E_{elec} \rangle$	-22763,5	550,0	-45624,0	805,4	-70906,5	826,9	-67835,5	1045,6
$\langle E_{vdW} \rangle$	-1389,6	90,5	-3137,7	126,1	-5271,3	188,4	-4986,1	159,0
$\langle E_{MM} \rangle$	-1196,5	696,9	-4153,0	953,6	-7835,9	1004,2	-4376,8	1254,7
$\langle G_{solv,polar} \rangle$	292,9	5,7	493,2	7,4	706,9	12,3	722,6	8,5
$\langle G_{solv,polar} \rangle$	-9574,7	285,1	-11871,9	407,4	-18648,1	349,4	-19701,7	479,0
$\langle G_{solv} \rangle$	-9281,8	282,3	-11378,7	400,0	-17941,2	341,6	-18979,1	474,8
-TS	-351,0	—	-268,2	—	-315,8	—	-315,9	—
$\langle G_{union} \rangle$	-10829,6	—	-15799,9	—	-26092,9	—	-23671,9	—

Tabla 5.13: Detalle de componente energética por estructura protéica para la estimación de energía libre del complejo TLR4-TRAM

# Capítulo 6

## Discusión

### 6.1. Validación de los modelos estructurales utilizados

A la fecha, se han determinado, vía cristalografía de rayos X, las estructuras del dominio TIR de TLR1, TLR2, TLR10 y MyD88 propias de los humanos [33, 107, 108]. Adicionalmente se han descrito estructuras tridimensionales para los dominios TIR humanos de TLR4, MAL y TRAM [20, 21]. En el último caso los modelos han sido generados a partir de la secuencia de aminoácidos de las proteínas por medio de modelamiento por homología utilizando un software como MODELLER o similares. De igual manera las estructuras consideradas en este trabajo fueron modeladas independientemente por Salgado (Resultados no publicados) siguiendo una metodología similar.

Los trabajos realizados por Nuñez *et al.* (2007) y Basith *et al.* (2011) entregaron descripciones importantes del sistema estudiado que permiten tener una base comparativa para los resultados aquí expuestos. En vista de esto, se vuelve necesario considerar la calidad de los modelos utilizados en cada caso, y utilizar esta información al analizar las discrepancias que puedan presentar los distintos modelos de interacción. De los resultados del estudio de calidad estructural de los modelos, se determinó que el modelo del homodímero TLR4 producido por Nuñez, usando los indicadores estudiados, resulta de mayor calidad que el presentado por Salgado (Resultados no publicados). Adicionalmente, los resultados demuestran que la calidad de los modelos estructurales de los adaptadores MAL y TRAM es superior en los modelos planteados por Salgado. En vista de esto, resultan de gran interés los contactos que puedan ser descritos a partir de este trabajo por cuanto provienen de un modelo estructural de alta fidelidad.

Es necesario destacar que los modelos propuestos por Nuñez no fueron refinados por medio de dinámica molecular. Si se relaciona este hecho con la calidad de los modelos, se revela el impacto que puede tener la dinámica molecular en la calidad de un modelo estructural generado por homología. Adicionalmente, el hecho de que la calidad estructural del modelo de TLR4 refinado por Salgado no sea altamente superior a la del receptor TLR4 propuesto por Nuñez, podría indicar que el efecto de la dinámica molecular sobre la calidad del modelo es mayor en los sistemas monoméricos y con mayores grados de libertad en proporción al número total de átomos del sistema. Esto no implica que se deba descartar el uso de dinámica molecular para refinar complejos de gran tamaño, sino que puede ser necesario complementar el proceso de refinamiento de estructuras por medio del estudio de otras conformaciones del sistema, o afinar los parámetros de la dinámica molecular para obtener una estructura de mejor calidad.

Por su parte, los modelos utilizados por Basith fueron refinados durante una simulación de dinámica molecular de 5.5 [ns], mientras que los modelos generados por Salgado se refinaron por 6.7 [ns]. Dado que no se contaba con las estructuras propuestas por Basith, se probó la calidad estructural de los modelos de Salgado usando la herramienta ProQ tal como se plantea en el trabajo de Basith [69]. Los resultados muestran que si bien en ambos casos los modelos son catalogados como excelentes, el valor de los indicadores es superior en los modelos de Basith. Lo anterior muestra que una mayor duración de la simulación de dinámica molecular durante la etapa de refinamiento no implica necesariamente que la estructura final sea de mayor

calidad. Esto se debe a que la dinámica molecular es un proceso dependiente de la estructura inicial a partir de la cuál se desarrolla, por lo que si se desea mejorar los modelos ya planteados es necesario explorar de manera más exhaustiva el universo de conformaciones libres para cada proteína. La necesidad aquí planteada ha sido tratada por una serie de estrategias que se enmarcan dentro del concepto de Muestreo Conformacional (*Conformational Sampling* en inglés). Bajo este concepto, se reconoce la necesidad de explorar la mayor cantidad de conformaciones posibles, de modo de poder observar la prevalencia de alguna de estas en el tiempo. Por lo mismo, se diseñan estrategias que permiten la exploración intensiva de las conformaciones de un sistema permitiendo el estudio de escalas de tiempo extensas, de un modo mucho más práctico que en el caso de la dinámica molecular tradicional. Existen diferentes aproximaciones para la implementación de la dinámica molecular acelerada y se les puede distinguir de acuerdo a la estrategia utilizada, observándose el uso de funciones de potenciales de energía modificados, estrategias alteradas de sampleo y uso de dinámicas acentuadas [49].

Debe destacarse que mientras este trabajo era desarrollado, se resolvió la estructura cristalográfica del dominio TIR del adaptador MAL [47]. Los resultados de ese trabajo muestran que la estructura de MAL presenta diferencias significativas con la de los dominios TIR conocidos y modelados por éste y otros trabajos. La principal diferencia yace en la presencia de un lazo largo que comprende las estructuras antes consideradas como hélice  $\beta$ B y lazo BB. De esta manera el motivo conservado y funcional de los residuos correspondientes al lazo BB adopta una conformación mucho más expuesta. Estos resultados tienen un efecto inmediato en la validez de los complejos generados, sobre todo considerando que los modelos obtenidos para el complejo TLR4-MAL muestran en ambos casos que el adaptador MAL orienta su lazo BB hacia la misma zona de homodímero de TLR4. Más aún, el modelo TLR4-MAL(P) muestra la interacción directa de ambos sectores vía la interacción hidrofóbica PRO125-ILE713 entre MAL-TLR4, y además muestra otros residuos del lazo BB de MAL, identificados como relevantes experimentalmente, interactuando directamente (R115 de MAL) o localizados en zonas cercanas al lazo BB de TLR4 (F117A, L118A, R121A, D122A e I129A de MAL), donde estos últimos no alcanzan a cumplir los criterios para calificar como interacciones.

Adicionalmente, se debe considerar que el principal algoritmo utilizado está basado en un criterio exclusivamente geométrico. De modo que cualquier modificación en las estructuras afecta directamente los complejos obtenidos. Sin embargo, debe considerarse que el fin de este trabajo es encontrar características que sean generalizables en la interacción, por sobre la reproducción exacta de las interacciones de la vía. En vista de esto, se plantea que tanto la metodología como las estructuras aún permiten el reconocimiento de patrones e interacciones en miras del diseño racional de estrategias de bloqueo. Más aún, los residuos del sector hélice  $\alpha$ B y lazo BB quedan mucho más expuestos en esta nueva conformación por lo que son potenciales candidatos a formar contactos de acuerdo tanto a los antecedentes previos y los resultados obtenidos.

## **6.2. Uso de algoritmos de acoplamiento en el modelamiento de interacciones en la vía mediada por TLR4.**

Una vez validados los modelos de las estructuras del receptor TLR4 y adaptadores MAL, TRAM y MyD88, se produjo los modelos para los complejos mediante el uso del algoritmo GRAMM. La selección de este algoritmo en desmedro de otras herramientas, se tomó de acuerdo a los resultados de pruebas preliminares, en las que se buscaba responder a la evidencia relacionada a que la unión del adaptador MAL o TRAM debía realizarse en el sector de la interfase del homodímero de TLR4. El uso de este algoritmo de acoplamiento tiene asociadas ciertas desventajas relacionadas principalmente al funcionamiento mismo del algoritmo. GRAMM trabaja a partir de la descomposición del espacio de búsqueda en una grilla tridimensional sobre la cuál se trazan las superficies del receptor y ligando que están siendo estudiados. Posteriormente el algoritmo busca la conformación que maximice el área de contacto receptor-ligando, al

mismo que penaliza la superposición de las estructuras del receptor y ligando.

El algoritmo GRAMM trabaja en base a un criterio estrictamente geométrico para determinar la mejor pose en la interacción receptor-ligando. Si bien esto favorece el desempeño computacional del algoritmo, sobretodo mediante el uso de FFT en los cálculos, se presenta finalmente como una desventaja por cuanto el algoritmo no considera, entre otras variables, la naturaleza de los aminoácidos en contacto, la efectiva interacción receptor-ligando, componentes electrostáticas, componentes de desolvatación, ni la energía de los complejos generados.

Este algoritmo es ampliamente utilizado por su capacidad de predecir el acoplamiento de todo tipo de sistemas proteína-proteína y proteína-ligando [109], ya que las interacciones proteína-proteína efectivamente están ligadas a la complementariedad de las superficies en contacto [110]. Para el caso estudiado es esta característica del algoritmo la que permite encontrar con mayor facilidad conformaciones en que el adaptador se ubique en la zona de la interfase del homodímero de TLR4. En particular, la naturaleza cóncava de la zona de interacción TLR4-TLR4 favorece el acoplamiento de los adaptadores MAL y TRAM en este sector, ya que al insertarse en ella el adaptador maximiza la superficie de contacto con el receptor lo que permite la alta recuperación de conformaciones que consideren esa área de interacción. Lo anterior también es aplicable a la unión del adaptador MyD88 al complejo TLR4-MAL, en la que el receptor mostró acoplamientos con una clara preferencia por las nuevas superficies formadas por la interacción TLR4-MAL.

Con respecto a otros de los algoritmos usados, los resultados mostraron que el algoritmo HEX tiene la capacidad de entregar resultados similares a GRAMM. Este algoritmo de acoplamiento funciona en base a la superposición de las superficies de las proteínas y complementariedad electrostática de éstas [75]. El hecho de que los resultados de este algoritmo sean favorables y similares a los que presenta GRAMM, indica que la complementariedad electrostática de las estructuras de MAL, TRAM y TLR4 también favorece la interacción en la interfase del homodímero. En efecto, la complementariedad electrostática entre MAL y TLR4 ha sido tratado en O'Neill *et al.* (2007), donde se plantea que posiblemente son las interacciones electrostáticas las que explican las interacciones específicas entre los adaptadores y TLR4. En particular, la superficies electrostáticas de TLR4 y MyD88 son negativas, mientras que la de MAL es positiva. Esta información favorece las estructuras propuestas, ya que se predice la repulsión en la interacción directa TLR4-MyD88, pero se favorece la interacción TLR4-MAL y MAL-MyD88, lo que resulta consistente con los modelos estudiados y los propuestos en este trabajo.

Los algoritmos Zdock y pyDOCK también fueron utilizados en el estudio preliminar sobre los sistemas TLR4-MAL y TLR4-TRAM. Los resultados mostraron que estos algoritmos no lograron modelar el acoplamiento buscado entre TLR4 y adaptadores. Por el contrario, mostraron resultados en sectores alejados a los BB loops de TLR4. Ambos algoritmos seleccionan los modelos predichos en base a la complementariedad de superficies y utilizando adicionalmente funciones de scoring que contemplan componentes electrostáticos y de desolvatación [77, 78].

La complementariedad de superficies que presentan estos algoritmos es similar a la de GRAMM, ya que consiste en la generación de una grilla de puntos que rodea tanto a receptor como ligando, y una posterior búsqueda por acoplamientos que abarquen la mayor superficie posible, penalizándose el choque de las estructuras. Adicionalmente, la componente electrostática de los algoritmos está basada en la fórmula de Coulomb al igual que en el caso de HEX. En vista de esto, se cree que el componente crucial que impide a Zdock y pyDOCK encontrar las conformaciones buscadas es el componente de desolvatación de la función de puntuación. Este componente se incorpora por la penalización de la desolvatación al esconder grupos polares previamente expuestos al solvente [78, 111]. En el caso de las proteínas estudiadas la hidrofobicidad superficial promedio calculada en función de la escala de Kyte-Doolittle [87] (Ver Anexo A.9), indica que las proteínas estudiadas poseen en la superficie un carácter polar (valores en rango -1.7 en promedio para TLR4 y adaptadores). Dada esta característica del receptor y adaptadores, es posible que la penalización por la desolvatación de estos grupos polares al esconderse provoque que las interacciones buscadas sean

caracterizadas por estos algoritmos como no favorables.

Otra de las implicancias que trae el uso del algoritmo GRAMM se relaciona con el hecho de que para efectos de este algoritmo, tanto el receptor como el ligando son tratados como entidades estrictamente rígidas. Esto indica que solo pueden ser reorientadas por el algoritmo y que se ignora totalmente la movilidad interna que pudieran presentar las estructuras estudiadas. Esta simplificación del problema no es menor, ya que la unión de un ligando a su receptor introduce generalmente cambios conformacionales y cambia la función de energía favoreciendo la interacción. Usualmente estos cambios pueden abarcar desde el movimiento de las cadenas laterales de los aminoácidos a cambios estructurales de mayor envergadura, que involucren el re-acomodamiento de las cadenas principales de las proteínas involucradas en la interacción [110]. Dado que el estudio del sistema mediado por TLR4 es aún limitado, no es posible determinar a este punto el efecto que puede tener la simplificación del movimiento del receptor TLR4 y sus adaptadores. Con el objetivo de remediar la rigidez del receptor y adaptadores impuesta por el uso del algoritmo GRAMM, se puede utilizar la dinámica molecular para estudiar los cambios conformacionales que se producen por la interacción entre las partes del complejo. En efecto, la dinámica molecular permite dar cuenta de que existen sectores que se re-acomodan producto de la interacción, sobre todo los lazos. Además, la dinámica molecular realizada permite afirmar que para tiempos del orden de los nanosegundos, no se observan cambios conformacionales importantes en la estructura del complejo. Se debe destacar que la dinámica molecular solo muestra aquellos movimientos que se producen una vez formado el complejo y no permiten observar el re-acomodamiento de la estructura del complejo, mientras se produce la aproximación del adaptador a la superficie del receptor TLR4. Adicionalmente, el movimiento está limitado a las condiciones iniciales impuestas por el modelo seleccionado para la dinámica molecular. En este punto nuevamente podría implementarse como alternativa al *docking* flexible, la estrategia de muestreo conformacional ya descrita.

La introducción de movilidad en un sistema proteína-proteína trae consigo una serie de interrogantes y complicaciones. Por un lado, debe determinarse si la movilidad se integrará a la estructura del receptor, ligando o a ambas. Debe considerarse que el permitir la movilidad del sistema, de cualquier manera, implica un aumento de los grados de libertad del sistema y de la complejidad del problema de optimización a resolver, lo que tiene un efecto considerable sobre el costo computacional de la implementación. Esto obliga a seleccionar aquellos sectores a los que se les brindará movilidad, lo que crea el problema de la selección de los segmentos que serán considerados como móviles. La respuesta a este problema es por ningún motivo trivial, ya que no existe información que permita privilegiar el movimiento de ciertos grupos de átomos por sobre otro. Si bien existen sectores reportados como de mayor relevancia en las interacciones, esto a priori no significa que esos átomos muestren una mayor movilidad o modificación estructural. En efecto, residuos catalogados como hot spots en la interacción proteína-proteína muestran valores de RMSD bajos a lo largo de simulaciones de dinámica molecular, lo que indica que son residuos que presentan baja movilidad [112].

Cabe destacar que la aplicación de los algoritmos ya comentados implica el uso de una estrategia de acoplamiento en la que el espacio de búsqueda en ningún momento es limitado. Se exploran todas las conformaciones posibles del ligando sobre la totalidad de la superficie del receptor TLR4. Como alternativa se probó la aplicación de los algoritmos Autodock y Vina sobre el sistema mediado por TLR4, ya que estas herramientas permiten definir un espacio de búsqueda acotado sobre la superficie del receptor. La estrategia utilizada dividió el espacio de búsqueda sobre la superficie de TLR4 en 26 áreas, con el objetivo de simplificar el problema de optimización planteado y mejorar las posibilidades de encontrar mínimos locales que pudieran ser biológicamente importantes. Los resultados de esto mostraron sectores con una menor energía libre estimada en las secciones asociadas a la interfase de interacción del homodímero TLR4 y sus lazos BB, lo que confirma la importancia de estos sectores en la vía de señalización mediada por el receptor TLR4. Aún así, de acuerdo a los resultados ya expuestos no se encontraron, en primera instancia, las interacciones buscadas entre los lazos BB de los adaptadores y de TLR4.

Al analizar los resultados de Autodock y Vina deben considerarse varios aspectos. En primer lugar,

estos algoritmos fueron diseñados para sistemas proteína-ligando, es decir para sistemas con un número menor de átomos totales. Autodock permite trabajar con un máximo de 2048 átomos en el ligando, mientras que las estructuras presentan números mayores a los 2300 átomos en el caso de los adaptadores utilizados, es decir sobre un 10% más de átomos. El límite máximo de átomos por ligando fue modificado por medio de la recompilación del código fuente de Autodock utilizado para estudiar el sistema. Sin embargo, debe considerarse que los parámetros que utiliza Autodock fueron calibrados bajo las condiciones originales, por lo que esta modificación podría disminuir la precisión de los resultados que entrega el algoritmo. Por otro lado Autodock utiliza algoritmos estocásticos (*Montercarlo Simulated Annealing*, Algoritmo Genético y Genético Lamarckiano) lo que implica que para asegurarse de encontrar estructuras correspondientes a mínimos globales, es necesario realizar repeticiones en cada caso de modo de validar los resultados encontrados. Aun así, la probabilidad de encontrar mínimos locales, en lugar de un mínimo global, es alta dada la complejidad de la función de energía libre de un complejo proteico del tamaño estudiado. Al trabajar con estos algoritmos solo se trabajó con 3 repeticiones mientras que en publicaciones anteriores se comenta el uso de un mínimo de 10 iteraciones para la obtención de resultados confiables [113]. En un principio se limitó el número de repeticiones en función del costo computacional ligado a desarrollar más repeticiones considerando que se contaba con 26 sectores distintos, en donde cada uno requería la misma cantidad de repeticiones. Sin embargo, como análisis *ex post* es importante recalcar el hecho de que el mejor resultado, en 2 de las 3 repeticiones, siempre alcanzaba la misma estructura. Esto está relacionado con la limitada cantidad de conformaciones que pueden ser exploradas al considerar como rígidas las estructuras del receptor TLR4 y los adaptadores. Más aún, los resultados de Patiño (2012, Resultados no publicados) al utilizar esta misma estrategia con péptidos de bloqueo (15 aminoácidos aproximadamente), en vez del dominio TIR de los adaptadores, mostraron una mayor variabilidad en los mejores resultados obtenidos en cada repetición cuando se incluía la flexibilidad de algunos residuos, en contraste con la situación en la que se consideraba rígida cada una de las estructuras.

La función de energía utilizada por Autodock y Vina es una función semi-empírica. Estima la energía del proceso de unión de dos o más moléculas en un ambiente acuoso evaluando por pares de átomos las contribuciones a la interacción intramolecular, electrostática y de desolvatación. Adicionalmente utiliza un método empírico para estimar la contribución del de las moléculas de agua en los alrededores. Si bien el uso de Autodock y Vina entrega resultados que pueden considerarse favorables al analizar las energías de unión según sectores de acoplamiento (valores energéticos más bajos en sectores del lazo BB de TLR4), la desviación estándar del valor energético (2.63 [kcal/mol]) muestra que habrían otros sectores (similares a los encontrados por Zdock y pydock) que son energéticamente equivalentes [113]. Esto podría indicar que el término energético relacionado al efecto de la adición del solvente puede provocar que sectores no pertenecientes a la interfase de interacción del homodímero de TLR4 alcancen puntajes que los caracterice como favorables.

Finalmente, a partir del uso de diferentes estrategias y algoritmos fue posible observar que el desempeño del algoritmo GRAMM superó al de otros en el estudio del sistema mediado por el receptor TLR4. Las principales limitaciones que trae el uso de este algoritmo están relacionadas al criterio estrictamente geométrico con que trabaja GRAMM y a la rigidez con que se trata tanto a receptor como ligando. El uso de la dinámica molecular permite subsanar en parte las limitaciones presentadas por GRAMM. A partir de simulaciones de la dinámica de los complejos es posible brindar movilidad y estudiar efectivamente las interacciones de los complejos estudiados y los cambios conformacionales provocados por el acoplamiento del ligando. Si bien la conformación inicial se sigue realizando a partir de un modelo rígido, la dinámica permite que se produzca el refinamiento de los modelos de acoplamiento. En el caso estudiado se utiliza la dinámica molecular para refinar los complejos predichos por GRAMM para el acoplamiento TLR4-MAL. Posteriormente se utiliza esa estructura para estudiar el acoplamiento de MyD88 sobre este sistema refinado.

### 6.3. Definición de las interacciones proteína-proteína por medio de una estrategia de docking y dinámica molecular

A partir del trabajo realizado se modeló las interacciones proteína-proteína características de las primeras etapas de la transmisión de la señal pro-inflamatoria en la vía mediada por TLR4. Concretamente, se modelaron las interacciones del receptor homodimérico TLR4 con los adaptadores MAL y TRAM. Posteriormente se estudió la interacción del adaptador secundario MyD88 con el ya formado complejo TLR4-MAL. El modelamiento de las interacciones en esta vía ha sido tratado en publicaciones anteriores, destacando las descripciones realizadas por Nuñez *et al.* (2007) y Basith *et al.* (2011). Sin embargo, ninguno de estos trabajos utilizó la dinámica molecular como herramienta para el estudio y descripción de las interacciones de esta vía. Por el contrario, el uso de la dinámica molecular se limitó al refinamiento de las estructuras que luego fueron acopladas mediante algoritmos de docking rígido.

Cabe destacar que las interacciones, tanto en el caso de los modelos TLR4-MAL como TLR4-TRAM, fueron definidas a partir de dos conformaciones distintas. Esto es, que la unión del adaptador MAL, o TRAM, con el homodímero del receptor TLR4 presenta una orientación distinta en los modelos propuestos en este trabajo y en los planteados por Nuñez. En ambos casos los modelos corresponden a la mejor solución encontrada vía *docking rígido*, validada por medio del uso de evidencia experimental. Las diferencias que presentan las conformaciones en la orientación del adaptador están relacionadas, primer lugar, a las diferencias existentes en la representación geométrica de los dominios TIR del receptor TLR4 y los adaptadores de los modelos propuestos por Nuñez y los utilizados en este trabajo. Este hecho afecta notablemente los resultados del acoplamiento, ya que en ambos trabajos se utilizó el algoritmo GRAMM, basado en un criterio estrictamente geométrico. En vista de esto, cualquier diferencia estructural presente entre los modelos limita las conformaciones que es posible encontrar como resultado del uso del algoritmo. Adicionalmente, la selección de los modelos esta claramente sesgada por la interpretación de la evidencia disponible hasta este momento. Más aún, ninguno de los modelos ha podido ser corroborado completamente por medio de evidencia experimental, por lo mismo ambos modelos pueden potencialmente entregar información relevante acerca de las interacciones dentro de la vía mediada por el receptor TLR4.

En este trabajo se utiliza un enfoque distinto a la hora de definir las interacciones proteína-proteína. En primera instancia se define un modelo de acoplamiento entre el receptor TLR4 y los adaptadores MAL/TRAM utilizando el algoritmo de acoplamiento GRAMM. Este modelo puede ser analizado para encontrar información general acerca de la interacción: área de interacción, aminoácidos escondidos por el acoplamiento, segmentos de las proteínas que muestran mayor contacto y participación en la interacción. Posteriormente se somete la estructura de los complejos a una simulación de dinámica molecular utilizando solvente explícito. En base a la trayectoria utilizada, y mediante el uso de herramientas disponibles en VMD, GROMACS y otras desarrolladas por Patiño (2012, Resultados no publicados), se describe los contactos en base a criterios de corte geométricos (distancias y ángulos atómicos). De esta manera es posible describir de manera detallada cuales son los residuos que interactúan, que tipo de interacción forman y durante que porcentaje de la simulación se observan tales contactos. Esta forma de modelar los contactos es sin duda más compleja que la presentada en los trabajos de Nuñez y Basith, pero aproxima de mejor manera el cómo se producen realmente las interacciones proteína-proteína.

En la práctica, el uso de la dinámica molecular demuestra que a lo largo del tiempo para una mismo complejo, es posible encontrar una serie de interacciones distintas que aparecen y desaparecen en el tiempo. Además muchas de las interacciones observadas mediante el enfoque de *docking* desaparecen al usar la dinámica molecular y los criterios geométricos establecidos para describir las interacciones. En efecto, en el caso del modelo de interacción TLR4-MAL, mediante las deficiones de Nuñez o la variación establecida en este trabajo, se detectaron 36 contactos para el modelo TLR4-MAL(N) y 22 para el TLR4-MAL(P), en cambio por medio de la dinámica molecular se estableció en promedio por *frame* 20 y 15 interacciones



respectivamente. Lo mismo se observó en el caso de los modelos del complejo TLR4-TRAM. Mediante la definición de contactos vía ASA se observaron 20 y 15 contactos para los modelos de Nuñez y el propuesto en este trabajo, mientras que mediante dinámica molecular se observó en promedio 31 y 9 interacciones por *frame*. Además, existen contactos que fueron observados exclusivamente a partir de la dinámica molecular y los criterios geométricos, como también contactos que solo se observó a partir de la definición vía ASA para los contactos. A modo de ejemplo, se tiene que las interacciones PRO125-ILE713 y R115-H740, validadas mediante evidencia experimental, se observan solo mediante los criterios propuestos en este trabajo y no por medio de la definición vía ASA. De manera similar los residuos L118, Q119, D122, T124 pertenecientes al BB *loop* de MAL se muestran como residuos en contacto de acuerdo a la definición vía ASA, pero mediante los criterios propuestos en este trabajo no muestran participación en ningún tipo de interacción.

Debe recordarse que tanto en el trabajo de Nuñez como en el de Basith, los contactos en cada complejo fueron identificados por medio del cálculo de la variación del ASA (*Accesible Surface Area*) de los distintos aminoácidos, esto es el cambio en el área expuesta al solvente de cada aminoácido. La definición utilizada por Nuñez establece que los residuos que muestran decrecimientos de más de 1 [Å<sup>2</sup>] en su ASA son considerados como residuos en contacto. Así mismo, los residuos que muestran decrecimientos de ASA mayores a 40 son considerados como contactos importantes [20]. Esta manera de modelar los contactos posee varias limitaciones. En primer lugar, no considera la naturaleza de los distintos contactos. Esto implica que, *a priori*, no es posible distinguir entre contactos de tipo hidrofóbico, puentes de hidrógeno, catión- $\pi$ , puentes salinos, entre otros. Además, definir los contactos en función del ASA entrega muchas veces falsos positivos. Esto se debe a que se considera como parte de la interacción a residuos que son arrastrados de manera fortuita en la zona de interacción, al estar cerca de un residuo que está interactuando, encontrándose posteriormente que estos no necesariamente forman contactos. Adicionalmente, la definición utilizada por Nuñez considera exclusivamente la variación absoluta en el ASA. Esto puede ser un problema en residuos que poseen inicialmente valores altos de ASA (típicamente expuestos al solvente), por cuanto considerar un decrecimiento en 1 [Å<sup>2</sup>] (o incluso 40 [Å<sup>2</sup>]) puede no representar un cambio suficiente en el área expuesta si su valor inicial es muy alto. En vista de esto se planteó, en las etapas iniciales de este estudio, la modificación de la definición de Nuñez para los contactos, de manera de considerar la variación relativa del ASA de un aminoácido con respecto a su estado inicial. Aún así, la principal desventaja de esta definición es que no permite identificar la naturaleza de los contactos detectados. El objetivo de utilizar los cambios en el ASA de los residuos para identificar zonas de interacción, aun conociendo las limitaciones de esta metodología, es poder contrastar, bajos los mismos o similares criterios, los resultados expuestos en los trabajos de Nuñez y Basith con los propuestos en este trabajo.

De manera complementaria al criterio basado en ASA, los trabajos de Nuñez y Basith utilizan herramientas como HBPLUS y ELECINT para definir los puentes de hidrógeno y las interacciones electrostáticas dentro de cada complejo. Este tipo de software sí consideran la distancia y orientación de los átomos en la estructura y presentan una mejora importante en la descripción del sistema. Sin embargo, los contactos son definidos en función de una única conformación, la obtenida producto del acoplamiento, por lo que no es posible *a priori* definir si los contactos observados representan efectivamente la interacción. Esta es una clara ventaja que muestra la metodología utilizada, que al considerar el movimiento de la estructura a lo largo del tiempo permite identificar aquellos contactos predominantes dentro de un set de conformaciones que representan el movimiento del complejo. Además, las herramientas utilizadas consideran una definición en base a criterios geométricos, no solo de puentes de hidrógeno, sino también de contactos hidrofóbicos e interacciones electrostáticas como puentes salinos e interacciones catión- $\pi$ .

### 6.3.1. Limitaciones de la estrategia de dinámica molecular utilizada

Aun cuando la metodología planteada en este trabajo permite definir con mayor precisión y detalle las interacciones de los complejos estudiados, es necesario tener en consideración las limitaciones que

presenta. Estas restricciones están ligadas tanto al diseño de la metodología utilizada, como también a la implementación de la dinámica molecular. En lo referido al diseño experimental, la metodología planteada considera la ejecución de una única simulación de dinámica molecular utilizando la mejor conformación obtenida para cada complejo. Al ejecutar una simulación, la trayectoria obtenida se genera a partir de la resolución de las ecuaciones clásicas de Newton considerando un set de condiciones iniciales. En particular, las velocidades iniciales son generadas aleatoriamente por lo que, en estricto rigor, al realizar una nueva repetición de la dinámica molecular los contactos observados pueden ser diferentes. Esto implica que para realizar una descripción más precisa de los contactos presentes en la interacción puede ser necesario realizar repeticiones de la dinámica de un sistema, de modo de explorar los posibles movimientos producidos a partir de una misma conformación inicial. Aún así, las conformaciones exploradas por la dinámica están limitadas por la posición inicial, lo que puede determinar la exploración de la superficie de la función energética del complejo.

Para estudiar efectivamente otras conformaciones que pueda presentar el sistema, se debe considerar que los cambios conformacionales, en general, se observan en escalas de tiempo mucho mayores a la utilizada. Muchos procesos biológicos pueden considerar escalas de tiempo de micro hasta milisegundos o más, sobre todo si se requiere superar barreras energéticas altas entre un estado y otro [49]. Si bien este tipo de escalas puede alcanzarse al utilizar dinámica molecular convencional, se presentan restricciones dado que el costo computacional de la simulación crece considerablemente al aumentar la duración de la simulación. El tiempo de ejecución de la dinámica molecular puede ser reducido al realizar cálculos en paralelo. En efecto, la implementación en paralelo de GROMACS permitió reducir de manera importante el tiempo de ejecución de cada simulación. Sin embargo, de acuerdo al trabajo realizado, la duración promedio de una simulación para los complejos TLR4-MAL o TLR4-TRAM fue de 3-4 semanas para 6 [ns] de simulación utilizando un nodo de cálculo (8 cpu). Estos datos muestran que obtener simulaciones de un largo mayor resulta poco práctico ya que requiere de un extenso período de ejecución. Adicionalmente, de acuerdo a lo desarrollado, se observó que el aumento del número de nodos de cálculo no permite una reducción considerable del tiempo de ejecución de la simulación, ya que la transferencia de datos entre un nodo y otro se vuelve la etapa limitante en el proceso.

Ya se ha comentado en este trabajo la existencia de una serie de alternativas a la dinámica molecular convencional, que podrían permitir mejorar y completar la descripción realizada. Estas estrategias favorecen la observación de eventos de menor prevalencia en la dinámica molecular que usualmente permiten convertir un estado en otro. Existen diferentes aproximaciones para la implementación de la dinámica molecular acelerada, ejemplos concretos de la dinámica molecular acelerada, que podrían ser utilizados en la descripción de este u otro sistema son: *Umbrella Sampling* [114], *Replica-Exchange* [115, 116], *Leap Dynamics* [117].

*Umbrella Sampling* corresponde a los métodos que usan potenciales modificados; en él se suma una función potencial a la energía del sistema para sesgar el muestreo de conformaciones. La desventaja de este método está en que la construcción de la función potencial (*Umbrella potential*) requiere del conocimiento *a priori* de las conformaciones de interés [114]. Esto se puede presentar como una limitación en el caso del sistema estudiado, ya que se desconoce aún el espectro de conformaciones que pueden adoptar tanto el receptor TLR4, las proteínas adaptadores MAL, TRAM, MyD88 y los complejos que pueden formar. Por su parte, la técnica *Replica-Exchange* utiliza una serie de simulaciones de MD o Monte Carlo, llamadas réplicas, las que se ejecutan a distintas temperaturas dentro de un rango previamente definido. Luego de ciertos intervalos de tiempo, las temperaturas de las distintas réplicas son intercambiadas [115, 116]. Estos cambios permiten explorar distintas conformaciones y superar más fácilmente las barreras energéticas que limitan la transición entre las conformaciones. La última de las estrategias mencionadas corresponde a *Leap Dynamics*. Esta aproximación implica la imposición de perturbaciones en la estructura de la molécula de modo de forzar la superación de las barreras energéticas, pudiéndose alcanzar otras zonas de la superficie energética. Posteriormente las conformaciones se refinan por medio de la aplicación de dinámica molecular

en solvente implícito [117]. Los últimos dos métodos no requieren conocimientos previos del sistema a estudiar, por lo que podrían ser implementados de manera de complementar la descripción realizada del sistema, precisando y permitiendo refutar o confirmar los contactos y las conformaciones propuestas.

### 6.3.2. Consideraciones acerca de la implementación de la dinámica molecular en la definición de interacciones

La definición de los contactos a partir de la dinámica molecular con solvente explícito incluye una serie de consideraciones y simplificaciones que permiten modelar el sistema estudiado. Estas van desde la definición de las condiciones termodinámicas iniciales del sistema, a la selección del algoritmo utilizado para la integración, paso de integración, largo de la simulación, campo de fuerza y modelos de agua explícita apropiados para la descripción del sistema, además de la definición de los estados de protonación de los distintos residuos.

El estado de protonación de los residuos es un factor importante a considerar dentro de la dinámica molecular, no solo para generar simulaciones realistas, sino también para estimar de manera correcta la energía libre y la representación correcta de algunas interacciones del sistema [91]. La principal dificultad que se encuentra en la descripción de los estados de protonación de una proteína, es que no se cuenta, en dinámica molecular, con algoritmos que consideren un pH constante. En general, los algoritmos consideran un número constante de protones. Esto implica que en una dinámica molecular tradicional no se observarán cambios de protonación/deprotonación del sistema, sino que los residuos conservarán los estados de protonación que se les asigne en un principio. Esta es una suposición fuerte, ya que ignora el hecho de que *in vivo* puede existir una transferencia de protones por efecto de interacciones electrostáticas, entre los residuos del adaptador MAL o TRAM y el receptor TLR4, cuando el adaptador se acerca a la zona de interacción. Este tipo de problemas en los cuales se considera la dinámica de la protonación/desprotonación de los residuos está siendo estudiado actualmente dentro del Laboratorio de Modelamiento de Procesos y Computación Distribuida del Departamento de Ingeniería Química y Biotecnología de la Universidad de Chile. En particular, el tema está siendo tratado por Pilar Fernández, quien se encuentra desarrollando el trabajo "Identificación de sitios posibles de desprotonación en la estructura de la enzima monoaminoxidasa A".

En el caso de GROMACS, los estados de protonación son asignados por el paquete *pdb2gmx*. Un inconveniente que presenta esta asignación es que GROMACS considera el estado de protonación más común para cada aminoácido en un solvente a pH 7. Esta es una simplificación importante, ya que además no considera el efecto que puede tener la estructura proteica en el pKa de los diferentes residuos. Para comprobar los estados de protonación establecidos por GROMACS se utilizó el servidor dedicado H++ (sitio web: <http://biophysics.cs.vt.edu/H++>) [92]. Esta herramienta permite obtener una estimación de los pKs de cada residuo en las estructuras estudiadas, así como otras características como puntos isoeléctricos, curvas de titulación y energías de protonación. No obstante, ambas herramientas asumen que los estados de protonación se mantienen fijos a lo largo de la simulación, ignorando cualquier intercambio de protones que pudiese existir por efecto de la proximidad entre los adaptadores y el receptor.

La comparación de los estados de protonación asignados por GROMACS con los propuestos por el servidor H++, para los aminoácidos ASP, GLU, HIS, LYS, ARG, CYS y TYR, muestra que la simplificación utilizada por GROMACS parece ser adecuada para las estructuras de MAL, TRAM y TLR4. En todos los casos se asignaron los mismos estados de protonación (Ver Anexo A.14), exceptuando los siguientes casos: LYS210 en MAL, LYS80 en TLR4 aparecen protonados en GROMACS pero no en H++; HIS728 en TLR4 (cadenas A y B) y GLU70, GLU72, GLU92, GLU196, GLU207 en TRAM se muestran desprotonados en GROMACS mientras que debiesen estar protonados según los datos de H++. De todos estos residuos el único que participa en interacciones importantes es el residuo HIS728. De acuerdo a las simulaciones este

residuo forma un puente de hidrógeno de alta prevalencia en la formación del complejo TLR4-TRAM (SER, participando como aceptor). Esto es posible gracias al estado deprotonado asignado por GROMACS. En caso de estar protonado como propone el servidor H++, el residuo no podría haber participado como aceptor en la interacción. El caso anterior muestra que efectivamente el estado de protonación de los residuos puede afectar la descripción de las interacciones de los complejos estudiados. Sin embargo, no parecen existir otras discrepancias producidas por la asignación en los estados de protonación realizado por GROMACS. En vista de esto, se considera que los estados de protonación asignados permiten describir el sistema de manera adecuada, tomando en cuenta las limitaciones que plantea el hecho de que los estados de protonación se mantengan fijos a lo largo de la simulación.

Otro factor importante en la implementación de la dinámica molecular es la selección del largo de la simulación. Esta variable resulta relevante ya que tiene un efecto directo sobre la definición de los contactos, la estabilidad observada en los complejos y las conformaciones finalmente exploradas. En el caso del estudio realizado la extensión de las simulaciones alcanzó los 6 [ns]. Este es un valor bajo y fue seleccionado en vista del tiempo disponible para la realización de este estudio, ya que el desarrollo de una simulación de este largo toma alrededor de 3-4 semanas usando 8 CPUs para el cálculo. El largo de la simulación tiene un efecto directo en la clasificación de los contactos, como de alta o baja prevalencia. Evidentemente, si la simulación fuera extendida por un lapso de tiempo mayor, podría observarse cambios considerables en los porcentajes de prevalencia reportados en este trabajo. Estas variaciones pueden ser aún más importantes al tomar en cuenta que las estructuras de los complejos, de acuerdo a las fluctuaciones de RMSD  $C\alpha$  observadas, parecen no lograr alcanzar conformaciones estables en el tiempo. Esto podría implicar que, *a priori*, no se pueda considerar ningún contacto como relevante, bajo la sola observación de su prevalencia en la simulación. No obstante al visualizar la dinámica molecular de los complejos se observa que estos si mantienen una o más conformaciones estables en el tiempo.

Por otro lado, la extensión de una simulación de dinámica molecular permite observar en mayor o menor grado las distintas conformaciones de los complejos y sus componentes a lo largo del tiempo. De los resultados se desprende la necesidad de extender las simulaciones, ya que, de acuerdo a las fluctuaciones del RMSD  $C\alpha$ , los complejos estudiados no alcanzan conformaciones estables por más de un par de nanosegundos. Si bien el orden de magnitud de las fluctuaciones parece no afectar la estabilidad macro del sistema, lo que se aprecia al visualizar la dinámica molecular, sí puede tener un efecto significativo en los contactos que están siendo definidos a partir de las simulaciones. Los cambios en el coeficiente de variación del RMSD se traducen en un leve movimiento de la cadena principal (que varía de acuerdo al complejo estudiado), lo que también puede provocar que el valor reportado para la prevalencia de los contactos sea distinto al presentado y por tanto se esté sobrevalorando (o subvalorando) algunas interacciones. Para tratar de resolver este problema se debe extender la etapa de equilibrio hasta reducir considerablemente las fluctuaciones y comenzar con una estructura estable la etapa de producción. De manera adicional, el considerar la ampliación de la etapa de producción a valores en el rango de 10-100 [ns] permitiría el estudio detallado del movimiento de los lazos y hélices de la estructura de las proteínas consideradas [49].

Finalmente debe mencionarse que no solo los parámetros con que se implementa la dinámica molecular, pueden tener un efecto sobre los resultados observados en la interacción. El diseño de las etapas con que se desarrollará la dinámica molecular también puede tener un efecto importante en los resultados obtenidos y puede entregar evidencia del como se generan los contactos. En el caso de este trabajo, originalmente se había diseñado un protocolo que incluía una primera etapa de simulación al vacío. El objetivo de esta etapa era permitir el movimiento de la molécula de modo de reparar posibles choques estéricos que presente la estructura inicial de los complejos estudiados. Sin embargo, durante las pruebas realizadas se observó que durante esta etapa el complejo receptor TLR4 se separaba, es decir la homodimerización desaparecía. Además se observó que eliminar esta etapa del protocolo de dinámica molecular y comenzar la dinámica molecular con el sistema solvatado, permitían mantener el estado de homodimerización del receptor TLR4. En este caso, el posterior análisis de los datos de la simulación dió cuenta de que es el solvente en el

sistema el que permite que se produzca la interacción entre los BB *loops*, caracterizados por ser segmentos altamente hidrofóbicos en comparación al promedio de la estructura. Por lo mismo, su ausencia en el caso de la simulación al vacío, no permitió mantener la estructura homodimérica ya que no existe una fuerza que mantenga la interacción hidrofóbica del sistema y probablemente existe una repulsión electrostática que impide la mantención del homodímero en ausencia de agua.

Se estudió la hidrofobicidad promedio de los monómeros de manera de determinar el origen de los diferentes resultados estudiados. Los monómeros de TLR4 poseen una hidrofobicidad superficial de -1.7 (escala Kyle-Dolittle) lo que refleja el carácter polar de gran parte de su superficie. Sin embargo, los monómeros poseen sectores marcadamente hidrofóbicos, como el lazo BB y las hélices  $\alpha C$  y  $\alpha E$ . El estudio de los contactos del complejo TLR4-TLR4 muestra que son estos sectores los que mantienen la unión del complejo. Estos resultados muestran que la interacción de estos sectores parece ser dirigida por el efecto del solvente, el que favorece la interacción entre estos sectores de los monómeros de TLR4 provocando la homodimerización. Por lo mismo, en este caso el protocolo de dinámica molecular no puede contener una etapa de ajuste en condiciones de vacío ya que el homodímero de TLR4 se desarma.

Hasta el momento, la única estructura cristalizada que muestra la dimerización de dominios TIR es la propuesta para el complejo TLR10-TLR10 en el trabajo de Nyman, en ella se muestra la interacción entre los lazos BB y la hélice  $\alpha C$  de los monómeros. Posteriormente, Nuñez *et al.* (2007) supuso que la estructura de TLR4 se dimeriza de la misma manera que TLR10, sin tener evidencia directa que soportara tal hipótesis. Los resultados de este trabajo apoyan el trabajo de Nuñez, así como los trabajos de Gong *et al.* (2009) y Basith *et al.* (2011) que muestran interacciones similares para el homodímero TLR4-TLR4. Finalmente, es importante destacar que este es el primer trabajo en que se utiliza la dinámica molecular para describir las interacciones de la vía mediada por TLR4. Las descripciones realizadas por otros trabajos provienen del estudio de conformaciones obtenidas desde algoritmos de acoplamiento, y como tal en la mayoría de los casos no son capaces de modelar correctamente el efecto del solvente sobre la formación del complejo.

## **6.4. Caracterización y validez de los modelos propuestos para la vía mediada por TLR4 y adaptadores**

Los modelos desarrollados para la interacción entre TLR4 y los adaptadores MAL y TRAM, consideran que el evento que gatilla la señalización dentro de esta vía es la dimerización del receptor TLR4. Esta hipótesis ha sido ampliamente considerada y apoyada por una serie de trabajos [20,21,42,84]. Sin embargo, de acuerdo a la bibliografía revisada, no existe consenso acerca de cómo se produce la interacción en el homodímero TLR4-TLR4, ya que no se cuenta con una estructura cristalizada del homodímero. Si bien se ha descrito ampliamente la importancia del lazo BB en la dimerización del receptor TLR4 [34,35,83], se han propuesto modelos contradictorios. Por un lado, los modelos de Nuñez (2007), Gong (2009) y Basith(2011) proponen la interacción directa de los lazos BB de ambos monómeros de TLR4 y la formación de un complejo con claras características simétricas, tal como el observado en la dimerización del receptor TLR10. Por otro lado, trabajos como el de Toshchakov (2011) proponen una interacción asimétrica, en la que el lazo BB de uno de los monómeros de TLR4 forma enlaces con la hélice  $\alpha E$  del segundo monómero.

### **6.4.1. Modelo TLR4-TLR4**

De acuerdo a los resultados de la simulación del complejo TLR4-TLR4, la interacción se produce principalmente entre los lazos BB, y otros sectores cercanos como la hélice  $\alpha C$ , de ambos monómeros. Esta importante interacción se produce mayoritariamente a partir de contactos de tipo hidrofóbico, lo que se cree responde al efecto que tiene el solvente en la interacción. De acuerdo a lo observado, la presencia del

agua dirige la interacción provocando que sectores de mayor hidrofobicidad (en relación al promedio de la superficie del receptor) como el lazo BB y la hélice  $\alpha$ C interactúen escondiéndose del agua. La importancia de la interacción de los sectores señalados es apoyada por estudios de mutagénesis como el de Poltorak (1998), quien observó que la mutación P714H en TLR4 termina con la respuesta del receptor a estímulos de LPS. El residuo de P714 al que se hace referencia se encuentra en el lazo BB del dominio TIR del receptor TLR4. En este caso la mutación afecta el carácter hidrofóbico de esa zona y la estructura que brinda el residuo de PRO al lazo BB, al ser mutado por un residuo de HIS el que posee un carácter polar y carga positiva. De acuerdo a los resultados planteados este residuo participa en al menos 3 interacciones, 2 de las cuáles se observan por sobre un 70 % de la simulación (PHE712-PRO714 y PRO714-PHE712 entre los lazos BB de los monómeros A y B de TLR4). La tercera de las interacciones se observa durante una menor proporción de la simulación, pero es importante destacarla ya que se genera entre los residuos de PRO714 de cada monómero. A partir del modelo descrito, se plantea que la importancia de este residuo en la cadena de señalización yace en que permite la correcta orientación del resto de los residuos del lazo BB, de modo que puedan enfrentarse para generar los contactos descritos (ver Anexo A.10). La información disponible a partir del trabajo de Ronni *et al.* (2003), también apoya la importancia de las regiones descritas. En particular, se ha reportado que las siguientes mutaciones HYR708–710AAA (BB), DFI711–713AAA (BB), PGV714–716AAA (BB), I718A (BB), YE751-752AA ( $\alpha$ C) y R809STOP ( $\alpha$ E) afectan severamente la capacidad de señalización del receptor humano de TLR4. El efecto que tiene cada una de estas mutaciones puede ser explicado a través de la dinámica molecular. Los residuos de ARG809 e HIS708 forman respectivamente un puente de hidrógeno y un enlace catión- $\pi$ , capacidad que pierden al ser mutados por residuos de ALA, perjudicando la interacción entre los monómeros. Así mismo, la mutación de los residuos PHE712, ILE713, PRO714 y TYR751 por residuos de ALA afecta la formación de interacciones hidrofóbicas y también catión- $\pi$ , en el caso de TYR751. El efecto de esta mutación se puede deber a que si bien ALA es un residuo de tipo hidrofóbico, cuenta con una cadena lateral (-CH<sub>3</sub>) no lo suficientemente larga, o que no puede orientarse correctamente, como para participar en la generación de contactos. Además, de acuerdo a la escala de hidrofobicidad de Kyte-Doolittle las mutaciones F712A e I713A afectan notoriamente la hidrofobicidad del sector, ya que el valor de hidrofobicidad baja de 2.8 y 4.5, PHE e ILE respectivamente, a un valor de 1.8 correspondiente a ALA. Para los casos de los residuos P714 y Y751, el valor de hidrofobicidad aumenta de -1.6 y -1.3, para PRO y TYR respectivamente, a 1.8. Si bien en este caso aumenta la hidrofobicidad, de acuerdo a la escala de Kyte-Doolittle, se está perdiendo cadenas laterales de gran tamaño y que además en el caso de TYR no tan solo generan contactos hidrofóbicos, sino también interacciones catión- $\pi$ .

Si bien la descripción realizada mediante dinámica molecular para el sector de interacción TLR4-TLR4 es similar a la presentada en los trabajos de Nuñez (2007), Gong (2009) y Basith (2011). Es importante notar que la metodología utilizada por los trabajos mencionados usualmente sobrestima la cantidad de residuos participantes en la interacción. La definición basada en el cambio de ASA utilizada para determinar los contactos termina identificando como formadores de enlaces a residuos que simplemente son arrastrados a la zona de interacción. Esto se aprecia en la cantidad de residuos reportados como parte de la interacción. En el caso de Basith se reportan 29 residuos, en Nuñez 20, mientras que de acuerdo a los datos generados solo participan 15 residuos en la formación de enlaces. Si bien esta diferencia puede provenir de la estructura del modelo, esto se descarta ya que la superposición de los modelos muestra su alta similitud. Por lo mismo, se destaca la precisión con la que es posible describir las interacciones usando los datos de las trayectorias de dinámica molecular y un criterio geométrico para la determinación de los contactos.

Es importante destacar que debido a la dinámica molecular se puede notar que varios contactos determinados vía ASA desaparecen. A modo de ejemplo este criterio destaca la interacción de los residuos D711, F712, A717 del lazo BB y C747, E758 de la hélice  $\alpha$ C, los que de acuerdo a los criterios impuestos en la dinámica molecular no debiesen ser considerados como contactos. De manera similar existen contactos que fueron detectados exclusivamente por la dinámica molecular, las interacciones más relevantes son ARG718-GLU824, ILE713-ILE718, ARG745-TYR751, entre otras que pueden ser validadas mediante evidencia experimental. Es relevante notar que la definición vía ASA destaca como residuos en contacto a ciertos

residuos que claramente no pueden formar interacciones por encontrarse en zonas lejanas al sector de interacción. Esto último indica que la detección de los contactos vía ASA puede ser poco precisa.

#### 6.4.2. Modelo TLR4-MAL y TLR4-MAL-MyD88

Una vez formado el homodímero de TLR4, el escalón siguiente en la cadena de señalizaciones, y que termina con la producción de NF- $\kappa$ B, es la unión del adaptador MAL al recién establecido homodímero. Los modelos estudiados en este caso fueron el propuesto y el ya planteado por Nuñez. Ambos modelos muestran acoplarse a sectores equivalentes, uniéndose a ambos costados del receptor TLR4 y orientando el lazo BB hacia la misma zona del receptor TLR4. Esto apoya la hipótesis acerca de la importancia de la formación de una nueva superficie de interacción a partir de la homodimerización del receptor TLR4 [20, 21, 42, 84]. Sin embargo, los modelos presentados muestran diferencias claras en cuanto a los sectores que muestran interacciones relevantes. Las diferencias observadas están directamente relacionados con el hecho de que las conformaciones obtenidas por Nuñez y por este trabajo presentan orientaciones distintas. Adicionalmente, la observación de los modelos muestra que para que el modelo propuesto en este trabajo genere los mismo contactos que el de Nuñez se requiere más que el desplazamiento del adaptador y su reorientación en torno al receptor TLR4. De ser correcto alguno de los modelos, la manera más probable para dar cuenta de otros contactos experimentales que el modelo no reconoce es considerando la participación de un nuevo adaptador, pudiendo ser MyD88, TRIF o un adaptador distinto.

El modelo propuesto muestra la clara interacción, ausente en el modelo de Nuñez, del lazo BB del adaptador MAL con el mismo segmento del receptor TLR4. Más aún, la interacción involucra al residuo PRO125 de MAL cuya mutación P125A ha sido señalada por Lin *et al.* (2012) como causante de la atenuación de la activación de NF- $\kappa$ B [118]. Además, se ha observado que la mutación del tipo P125H impide la unión del adaptador MAL al receptor TLR4 [37]. Esta evidencia es apoyada por el modelo propuesto, que muestra una interacción PRO125-ILE713 entre MAL-TLR4, la que se puede observar durante un alto porcentaje de la simulación. Este resultado indica la importancia que pueden tener los contactos hidrofóbicos en la señalización. Además explica la diferencia que existe en el efecto de las mutaciones ALA e HIS sobre la señalización, ya que en la mutación por HIS provoca la pérdida de la capacidad para generar el contacto hidrofóbico. Por otro lado, la mutación por ALA si bien permite la formación del contacto probablemente no logra generar un contacto fuerte si se considera el tamaño de su cadena lateral.

Adicionalmente, se han reportado otro tipo de mutaciones asociadas al lazo BB que afectan la señalización producida por la cadena. Se ha demostrado que las mutaciones R115A, F117A, L118A, R121A, D122A e I129A producen una disminución en la respuesta inflamatoria NF- $\kappa$ B [47, 118]. Con excepción del residuo R115 de MAL que muestra la formación de un enlace catión- $\pi$ , el resto de los residuos señalados no muestra la formación de contactos, de acuerdo a las definiciones planteadas, en el complejo propuesto en este trabajo, pero si muestra una proximidad con los residuos del lazo BB del receptor TLR4. Lo anterior puede indicar la necesidad de estudiar el ordenamiento del adaptador a lo largo del tiempo, ya que cabe la posibilidad de que el complejo muestre la capacidad de reordenarse encontrando una conformación de menor energía durante una simulación más extensa. Adicionalmente se debe considerar la evidencia de Valkov *et al.* (2011) cuya estructura cristalográfica de MAL, distinta a las utilizadas en este trabajo, muestra un segmento AB en desmedro del lazo BB y hélice  $\alpha$ B que están mostrando interacciones en el modelo propuesto. Bajo esta nueva conformación, se podría tener a los residuos mencionados en una posición de mayor exposición pudiendo interactuar directamente con el lazo BB del homodímero.

Si bien el modelo de Nuñez no es capaz de reproducir los contactos que se observan en el lazo BB, sí muestra otras zonas con potencial para generar interacciones. En particular, las interacciones que se producen entre el lazo CD de TLR4 con el mismo sector de MAL, o en su defecto el lazo EE, también encuentran asidero en mutaciones reportadas y que son capaces de interrumpir o impedir parcialmente la

unión con el adaptador MAL. Cabe destacar que, en general, este tipo de contacto son principalmente puentes de hidrógeno (ALA168-GLN758, SER131-HIS724 entre MAL y TLR4) e interacciones catión- $\pi$  (LYS84-TRP757, GLU167-ARG763 entre MAL y TLR4). La aparición de los contactos responde a las características que tienen los residuos pertenecientes al sector CD, ya que corresponden principalmente a residuos polares o residuos con cadenas aromáticas, como TRP. Cabe preguntarse como es posible que se observen contactos experimentales distintos en cada uno de los modelos estudiados. Si bien es probable que ninguno de los modelos esta capturando totalmente la fenomenología estudiada, también es posible que sea otra molécula la que tenga que acoplarse al sistema para que se reproduzcan tales interacciones, aunque esto no se observe en el caso de MyD88. De todas maneras puede existir otro adaptador aunque por ahora no existe evidencia que sustente esta hipótesis. Una alternativa a esta idea es que aquellos residuos estudiados experimentalmente no participen en la interacción directa entre el adaptador MAL y el receptor TLR4. En ese caso, los residuos podrían tener otro tipo de influencia en la formación del complejo, ya sea estabilizando el homodímero o manteniendo una apropiada complementariedad electrostática.

Otra variable importante al analizar las interacciones proteína-proteína, es la magnitud de las zonas de interacción. De acuerdo a lo calculado, el área escondida por la interacción TLR4-MAL es de 1850 [Å<sup>2</sup>] y 3800 [Å<sup>2</sup>] para el modelo propuesto y el de Nuñez respectivamente. Los dominios TIR tienen estructuras de tipo globular y poseen en promedio un diámetro de 3.5 [nm] [84]. Adicionalmente estos dominios forman en general complejos transientes e inestables [84]. Este tipo de interacciones no obligada está asociada a una alta polaridad en la superficie de los dominios y a tamaños de área escondida por interacción del orden de 2000 [Å<sup>2</sup>] [119, 120]. En efecto, el área escondida en un complejo formado por dos receptores TIR y un adaptador TIR debiese alcanzar el orden de 1750 - 1850 [Å<sup>2</sup>], lo que concuerda con los resultados del modelo propuesto [84]. Por el contrario el modelo de Nuñez no es capaz de responder a estas características de la fenomenología estudiada, por lo que los contactos que muestra ese modelo también pierden representatividad.

Una vez formado el complejo TLR4-MAL, se crea nuevamente una nueva superficie que permite la unión del dominio TIR del adaptador MyD88. El modelo estudiado vía dinámica molecular muestra la unión del adaptador MyD88 en sectores similares a los ya reportados por Basith (2011). Tal como en el trabajo de Basith, se observa que las interacciones muestran participación del lazo DD y hélice  $\alpha$ C de TLR4 con la hélice  $\alpha$ A de MyD88. Destacan dentro del modelo las interacciones ARG197(MyD88)-GLN755(TLR4A) y ASP198(MyD88)-TRP222(MAL), ya que por un lado forman parte del BB loop de MyD88, y además se ha mostrado que las mutaciones R197A y D198A producen la disminución de la respuesta inflamatoria. Por otro lado, el área de interacción calculada para la interacción de MyD88 muestra un valor mayor al promedio observado en interacciones TLR4, pero que aún se encuentra dentro del rango observado para interacciones transiente y no obligadas [119, 120]. Los resultados del acoplamiento del adaptador secundario MyD88, muestran que el acoplamiento no involucra la participación ni del lazo BB de MAL ni del receptor TLR4. Por lo mismo, si se quisiera evaluar la efectividad de un péptido derivado de MyD88 para bloquear la cadena de señalizaciones los esfuerzos tendrían que estar orientados a estudiar las interacciones del péptido con los sectores del lazo DD y la hélice  $\alpha$ C de TLR4 que son aquellos sectores en lo que el adaptador mostró contactos. De todas maneras debiese estudiarse el efecto del peptido, y otros diseñados en base al receptor y el resto de los adaptadores, sobre la región del BB *loop* de TLR4 y el adaptador MAL. Esto último, dadas las interacciones encontradas para el complejo TLR4-MAL y el alto nivel de conservación que muestran los lazos BB del receptor y los distintos adaptadores de la vía.

### 6.4.3. Modelo TLR4-TRAM

De acuerdo a lo estudiado, el adaptador TRAM tiene la capacidad de unirse al receptor TLR4 de manera de estimular una cascada de señalización de respuesta tardía que lleva a la activación de la vía IRF-3 [36]. Los complejos obtenidos y descritos para la interacción TLR4-TRAM muestran que la unión del adaptador



se lleva a cabo de la misma manera que para el adaptador MAL. Esto implica que aprovechan la formación de una nueva superficie de interacción para su acoplamiento. Más aún, la interacción de TRAM se produce en los mismos sectores en que se une el adaptador MAL. Lo que implica que eventualmente es posible la unión simultánea de los adaptadores. En ambos casos el adaptador se une orientando su lazo BB a la zona de los lazos BB del homodímero de TLR4, apoyando la hipótesis que plantea la importancia de estos sectores. Sin embargo, el análisis de los contactos muestra grandes diferencias entre los contactos observados por el modelo de Nuñez y el Propuesto. Lo anterior también es reflejo de la diferencia que existe en el área de interacción que muestra cada modelo. Al igual que en el caso del complejo TLR4-MAL, se observa que el área escondida por la interacción propuesta ( $2422 \text{ [Å}^2\text{]}$ ) representa de mejor manera la naturaleza de los contactos entre dominios TIR [84].

De los contactos observados en cada modelo, es interesante destacar que el modelo propuesto en este trabajo muestra la interacción directa de los lazos BB de TLR4 y TRAM a partir de la formación de contactos hidrofóbicos y puentes de hidrógeno. Esta característica es exclusiva del modelo aquí propuesto y no es una propiedad que se observe en el modelo de Nuñez. En particular el complejo propuesto en este trabajo muestra una interacción hidrofóbica equivalente a la ya observada en el complejo TLR4-MAL. La interacción ALA717-PRO116 entre los residuos de TLR4 y MAL se observa durante gran parte de la simulación, y responde de la misma manera a la mutación de PRO por HIS que el adaptador MAL. Esto, sumado a la alta conservación de este residuo (Figura 5.21) y su bajo valor de RMSF (ver Anexo A.7), da muestras claras de la importancia del residuo de PRO en la generación de la cascada de señalización de esta vía, lo que esta es concordancia con lo sugerido por Poltorak (1998), Horng (2001), Ohnishi (2009), Basith (2011), Toshchakov (2005 y 2011) y en contraposición con lo propuesto por Dunne (2003)

El estudio de la estructura propuesta por Nuñez muestra una ausencia de interacciones directas entre los lazos BB de TLR4 y TRAM. En cambio, se observan interacciones entre los lazos CD del receptor y ligando, y entre el lazo CD de TLR4 y los sectores terminales del adaptador TRAM. En este caso las interacciones observadas son de tipo puentes de hidrógeno o catión- $\pi$ , lo que está en acuerdo con la mayor proporción de residuos polares observados en esta zona (ver Tabla 5.6). Nuevamente los contactos observados entre los lazos CD se ven afectados por una serie de mutaciones como Q758A, L760A o W757A las que afectan al receptor TLR4. En todos los casos la mutación produce la pérdida de una interacción por puente de hidrógeno o catión- $\pi$  observada durante la simulación, al reemplazar un aminoácido de carácter polar cargado o aromático por uno hidrofóbico. Cabe destacar la aparición de las interacciones catión- $\pi$  en las que participa el residuo TRP757. Este tipo de interacción, que se observa tanto en el modelo TLR4-MAL como TLR4-TRAM de Nuñez, muestra una alta prevalencia durante la dinámica molecular estudiada. Estos resultados indican que es posible que este residuo tenga un rol importante en la señalización o en la alta afinidad en la interacción. Casos similares se han reportado en el caso de los receptores nicotínicos [121, 122]. Otros contactos que resultan importantes en vista de la evidencia experimental son: HIS728-SER156, GLN758-ASN159 y ARG787-ALA235. En todos estos casos la mutación de HIS, GLN y ARG en TLR4 por ALA produce problemas en la unión con los adaptadores, probablemente por la pérdida en la capacidad formadora de puentes de hidrógeno o interacciones catión- $\pi$ .

Finalmente es importante destacar que las discrepancias entre los modelos estudiados no permiten determinar con certeza un mecanismo de interacción entre el receptor TLR4 y los adaptadores MAL y TRAM. Aún así los contactos descritos permiten dar cuenta de la existencia de dos blancos importantes para el diseño de péptidos: los lazos BB y CD. La dificultad que se encuentra en la descripción del sistema tiene varias aristas. Por un lado se tiene que el sistema de señalización es altamente complejo, lo que se observa en la estequiometría de las interacciones y su carácter transitorio que involucra la formación tanto de homo como heterocomplejos. Así mismo, la información experimental que se tiene del sistema es limitada, ya que el hecho de que las interacciones funcionen en base a intrincados complejos dificulta su estudio en el laboratorio. En lo que respecta al modelamiento, las dificultades parten al momento de determinar una estructura desde un experimento de acoplamiento. El hecho de que el modelo sea guiado por las decisiones

del investigador constituye tanto una fuente de aciertos como de fallas. La gran cantidad de resultados preliminares requiere de un trabajo intensivo de observación, que debiese ser acompañado por el uso de herramientas de minería de datos de modo de poder analizar los potencialmente miles de resultados que arrojan los diferentes algoritmos de acoplamiento. Herramientas de clusterización como árboles de decisión, máquinas de soporte vectorial, etc, pueden permitir analizar los datos de modo que no sean exclusivamente dependientes de la observación del investigador. La utilización de estas herramientas fue planteada en un principio pero en vista del avance del proyecto a lo largo del tiempo se desistió de su implementación.

## 6.5. Cálculo de la energía libre de los complejos TLR4-MAL, TLR4-TRAM

En vista de las discrepancias observadas entre el tipo de contactos que muestran los modelos propuestos y los de Nuñez, la estimación de la energía libre toma gran importancia para definir un modelo de interacción. De manera general, los resultados de la estimación de energía libre de unión dan cuenta de que el acoplamiento y los contactos observados en el modelo propuesto generan un complejo más estable tanto para el caso de MAL y TRAM (desde un punto de vista termodinámico) que en el caso de Nuñez. Llama la atención que, aún cuando el modelo de Nuñez presenta una mayor cantidad de interacciones, es el modelo propuesto en este trabajo el que presenta una menor energía libre (ver Tabla 5.11).

Con el objetivo de determinar el impacto, o relación, del número y tipo de enlaces que muestra cada complejo en la estimación de la energía libre, se evaluó adicionalmente esta componente por medio de la suma de las contribuciones individuales de cada enlace de acuerdo a su naturaleza (ver Anexo A.11). La contribución individual de cada enlace es un dato que se obtuvo a través de la bibliografía y corresponde a la diferencia de la energía libre de unión entre el estado nativo de un complejo y uno mutado en el cuál se ha impedido una interacción. La diferencia en el  $\Delta G$  de unión calculado como  $\Delta\Delta G = \Delta G_{Mutante} - \Delta G_{Nativo}$  entrega el aporte energético individual de un enlace particular. De esta manera, la contribución por interacciones receptor-ligando se estimó mediante la suma del aporte energético de cada tipo de enlace multiplicado por el número de enlaces promedio observados por tipo durante el largo de la simulación. Esta es una estimación diferente a la establecida vía MM-PBSA, ya que considerada solo las contribuciones energéticas por tipo y número de enlaces, ignorando contribuciones energéticas por efectos de solvatación e interacciones electrostáticas. En particular, para el caso de los complejos TLR4-MAL esta estimación entregó un valor de -121.6 [kcal/mol] para el modelo de Nuñez y -110.4 [kcal/mol] para el modelo propuesto. Estos valores indican que si bien el modelo de Nuñez presenta un mayor número de contactos, esto no se traduce inmediatamente en una mayor estabilidad energética. En este caso se aprecia que el tipo de contactos, y su energía asociada, es de suma importancia a la hora de estimar la energía libre de unión de cada complejo. En este caso, el alto número de contactos hidrofóbicos que presenta el modelo propuesto brinda una estabilidad equivalente en la formación del complejo con un menor número de enlaces en promedio para el complejo TLR4-MAL. Esto valida aún más el análisis planteado que reconoce la importancia de las interacciones hidrofóbicas como un factor importante para el desarrollo de la señalización en esta vía.

Adicionalmente, se observa para el caso recién descrito que el complejo TLR4-MAL (N) tiene una mayor estabilidad energética por concepto de los enlaces que se forman en el complejo. Sin embargo, esto no se refleja en la estimación final de la energía libre que muestra que el complejo TLR4-MAL (P) es termodinámicamente más estable que el propuesto por Nuñez. En efecto, la diferencia que muestran las estimaciones de energía libre de unión para ambos modelos es mucho más marcada que lo que muestran las estimaciones obtenidas vía contribución energéticas por enlace. Lo anterior implica que la solvatación y complementaridad electrostática pueden efectivamente tener un papel relevante dentro de las interacciones que se producen en esta vía. Esto se sustenta en las diferencias que presentan las estimaciones de los términos mencionados, cuya magnitud además favorece al complejo propuesto.

En el caso de los modelos para el acoplamiento TLR4-TRAM, la estimación basada en el aporte energético de cada enlace indica que el modelo TLR4-TRAM (N) alcanza una mayor estabilidad que el propuesto ( $\Delta\Delta G$  con valores -186,2 y -41,2 respectivamente). A diferencia del caso de los complejos TLR4-MAL, en esta ocasión es notable lo dispar que resulta el número total de interacciones promedio que se observa en cada caso. Sin embargo, aún existiendo esta diferencia, el resultado de la estimación de energía libre indica que el complejo propuesto tiene una conformación de mayor estabilidad termodinámica (aunque en ambos casos se presente un valor positivo para la energía libre). Esto da cuenta de que, al igual que en el caso del adaptador MAL, el efecto de las interacciones electrostáticas y de solvatación juega un papel importante en la estabilidad del sistema mediado por TLR4.

De la observación de las estructuras de los adaptadores MAL, TRAM y los complejos formados con el homodímero del receptor TLR4, se puede dar cuenta de que las conformaciones muestran interacciones que a grandes rasgos no muestran diferencias estructurales notorias. En vista de esto, era esperable que la magnitud de los distintos componentes energéticas estimadas mostraran rangos de magnitud y valores similares. No obstante, en el caso de la energía libre del complejo TLR4-TRAM se aprecia que los valores estimados de  $\Delta G$  son positivos, dando cuenta de la inestabilidad y no espontaneidad de la unión entre el adaptador y el receptor. El hecho de que la energía libre de unión sea positiva puede indicar la necesidad de que participe un adaptador adicional en la interacción, de manera que logre estabilizarla. En principio esta adaptador podría ser el adaptador TRIF, homólogo en sus funciones al adaptador MyD88, u otro de los adaptadores observados río abajo en la vía. Adicionalmente, se ha comentado con anterioridad la existencia de evidencia que indica que las interacciones entre TLR4 y MAL ocurren principalmente en la superficie de la membrana plasmática, mientras que la interacción entre TLR4 y TRAM se produciría luego de la internalización en endosomas o fagosomas [45]. Esta última información podría sugerir que el adaptador TRAM se une al ya formado complejo TLR4-MAL, esto podría permitir que el adaptador TRAM forme un complejo estable, además de explicar la respuesta tardía observada en la vía mediada por el adaptador TRAM.

Al observar el detalle de los valores de la estimación de energía libre se aprecia que existen algunas componentes cuya estimación resulta, en magnitud, bastante disímil. Uno de estos casos corresponde a la gran diferencia que existe entre los valores calculados para la energía interna de los complejos TLR4-MAL y TLR4-TRAM. De manera similar los resultados de la Tabla 5.11 muestran casos como la componente electrostática de energía y  $\Delta G$  solvatación que también muestran grandes diferencias, sobretudo al considerar los valores estimados para los modelos de TLR4-TRAM.

En el caso de los valores de  $E_{int}$ , no es clara la fuente de las discrepancias o errores. Al observar los valores de esta componente para ligando (MAL o TRAM), receptor (TLR4) o complejos, estos muestran en todos los casos valores positivos de magnitud proporcional al número de átomos que compone cada sistema lo que esta en acuerdo con las ecuaciones a partir de las cuales se estima el término [123]. En función de esas mismas ecuaciones es que se puede dar cuenta que una de las razones por las cuales se pueden producir las variaciones en la energía interna pueden estar asociadas a la vibración exagerada de enlaces y ángulos o a la aparición de torsiones no permitidas. La diferencia en el valor  $E_{elec}$  se puede asociar a la mejor complementaridad de cargas que se aprecia, en el acoplamiento propuesto en este trabajo, al observar las superficies electrostáticas generadas vía Delphi (ver Anexo A.12). Por su parte, la diferencia energética observada en el término  $\Delta G_{solv,polar}$  puede ser resultado del acoplamiento del lazo BB de MAL en el modelo propuesto. El hecho de que la interacción se produzca por el sector hidrofóbico del adaptador, permite que los residuos del segmento no queden accesibles al solvente y por tanto no se vean afectados por efectos de solvatación contrarios a la interacción sobre el lazo BB (de carácter claramente hidrofóbico). Estos últimos resultados en conjunto con los ya expuestos anteriormente, dan cuenta de la importancia que tienen los contactos de tipo hidrofóbico dentro de esta vía. Más aún, los resultados reportados por Patiño (2012) dan cuenta que péptidos de bloqueo diseñados a partir del lazo BB de MAL muestran interacción con residuos hidrofóbicos del lazo BB de TLR4. Esto, sumado a la evidencia presentada por Toshchakov [34, 84], permite mostrar que los lazos BB son claves en la transmisión de la señal pro-inflamatoria, y por tanto constituyen

un blanco atractivo para el diseño de péptidos de bloqueo que impidan, a nivel de receptor, la expresión de citoquinas pro-inflamatorias.

Sin perjuicio de lo anterior, es necesario recalcar ciertos aspectos de la metodología utilizada para la estimación de la energía libre. En lo que respecta a la metodología MM-PBSA es menester mencionar algunas de las limitaciones que han sido reportadas en publicaciones anteriores [124–127], sobretodo aquellas que pueden tener un efecto importante sobre los resultados planteados. Uno de los problemas más importantes que ha sido señalado en torno a la metodología MM-PBSA está relacionado a la incapacidad del método para calcular de manera precisa el valor del cambio de energía libre. Esta imprecisión que presenta el cálculo esta relacionada a varios factores, siendo los más importante los problemas asociados al cálculo de entropía del sistema y el tratamiento de la componente electrostática de la energía.

En el caso de la entropía, se ha reportado extensivamente las dificultad que presenta la determinación de esta característica del sistema [124–127]. Usualmente se asocia a la estimación de energía libre vía MM-PBSA, el cálculo de esta variable por medio del análisis de modos normales (NMA) o al análisis quasi-harmónico del sistema (QH). De acuerdo a lo descrito en otros trabajos ambas metodologías brindan estimaciones erróneas de la energía libre de unión. Por un lado el análisis NMA provee una cota inferior para la entropía conformacional, mientras que la aproximación QH representa una cota superior para la entropía conformacional real del sistema. Otro inconveniente de la metodología utilizada para estimar la entropía tiene relacion con que la entropía conformacional muestra grandes fluctuaciones a lo largo de una trayectoria de dinámica molecular, por lo que se requiere un gran número de capturas de la dinámica para establecer una estimación correcta [126]. Como alternativas al cálculo de entropía vía NMA o QH se han propuesto otras aproximaciones como una basada en la liberación de restricciones (RR), similar al método FEP, de posición cartesianas fuertes que son impuestas sobre los átomos del ligando. Este método ha sido señalado como uno de los pocos capaces de reproducir las tendencias experimentales de las entropías de unión [128].

Más allá de los problemas encontrados para la estimación de la entropía, es necesario mencionar los inconvenientes y limitaciones asociadas al tratamiento de la solvatación polar considerada por la metodología MM-PBSA. Las críticas hechas al cálculo de la solvatación propuesto, apuntan a que a esta altura no se ha verificado el método por medio de la verificación de valores electrostáticos conocidos [124]. Adicionalmente, uno de los parámetros más importantes utilizados en el cálculo de la componente polar de solvatación corresponde a la constante dieléctrica de la proteína que en este caso fue definida con un valor de 2. Este valor es usualmente utilizado cuando la interfase de interacción está medianamente cargada y se forman interacciones iónicas entre ligando-receptor. En vista de que en este caso la interfase de interacción tiene un carácter principalmente hidrofóbico en los mejores modelos, es probable que un valor de 1 fuera más apropiado. A priori no es fácil predecir el impacto de este cambio sobre los resultados del cálculo de energía libre de solvatación, pero si debe tenerse en consideración la posibilidad de que el valor de la energía efectivamente varíe producto de esta modificación. [126]

Como alternativa al cálculo de la energía de solvatación por medio de la resolución de la ecuación de Poisson-Boltzman (PB), suele utilizarse la ecuación del modelo generalizado de Born (GB) [125, 129]. Si bien parece existir un acuerdo en que el método PB presenta mayor rigurosidad teórica, en la práctica el modelo generalizado de Born tiene la capacidad de responder, de acuerdo a la naturaleza del sistema, brindando buenas correlaciones y predicciones para la energía de solvatación. Lo anterior se produce por el menor efecto que tienen los cambios conformacionales en el cálculo que se realiza en el modelo generalizado de Born. En efecto, el trabajo realizado por Li *et al.* (2010) plantea que en el caso de utilizar PB para la estimación de la energía de solvatación, se requiere en primera instancia dar con una estructura representativa y estable en el tiempo, que usualmente se obtienen como resultado de simulaciones de por sobre los 10 [ns] de simulación. [125, 127]

Además de las técnicas utilizadas para la estimación de cada una de las componentes de la energía libre, existen otra serie de factores que pueden afectar la precisión de la estimación realizada. Se destaca como un

parámetro importante el tipo de muestreo y el largo de la dinámica molecular. En el caso de la duración de la dinámica molecular, es claro el impacto que tiene la simulación sobre la correcta descripción del sistema. Sin embargo, dependiendo del sistema una simulación MD más larga no necesariamente implica la obtención de mejores resultados, por lo que si bien se ha propuesto la extensión de la simulación para detallar sus cambios conformacionales esto no garantiza que el cálculo de energía libre alcance mejores resultados [124]. Esto último se aprecia en los resultados de las Tablas 5.12 y 5.13 para los que en el caso de la energía mecánica, aún utilizando un promedio de la magnitudes, se obtienen coeficientes de variación de hasta un 25%. Para el caso del tipo de muestreo, el procedimiento estándar de la metodología MM-PBSA plantea el cálculo de sus componentes como el promedio de los valores obtenidos para un conjunto de conformaciones observadas durante la dinámica molecular [127]. Sin embargo trabajos como el de Li *et al.* (2010) han demostrado que se puede obtener mejores resultados y correlaciones si se utiliza una estructura representativa en vez de un promedio de estructuras. El problema de aplicar este enfoque al trabajo realizado yace en la dificultad de encontrar una estructura representativa a partir de la dinámica molecular ya que no se cuenta con estudios de cristalografía de rayos X para el sistema. Como alternativa podría utilizarse un método de *clustering* de las conformaciones, de manera de poder encontrar una estructura representativa

Finalmente, si bien la metodología MM-PBSA presenta claros problemas para determinar con exactitud el cambio de energía libre, permite identificar aquellos ligandos que presentan mayor afinidad en su unión al complejo. De esta manera, es posible analizar una serie de conformaciones distintas, como en el caso presentado en este trabajo, generando un ranking de los complejos obtenidos que refleje la mayor estabilidad de una estructura por sobre las otras. Sin embargo, se ha reportado que la sensibilidad del método no es suficiente para determinar diferencias entre complejos que poseen niveles energéticos similares [124]. En vista de esto la metodología podría servir principalmente como una herramienta que permita diferenciar entre lo que se determina como ligandos fuertes o débiles (de alta o baja afinidad respectivamente). Aún en presencia de estas limitaciones, la ventaja principal que presenta esta metodología es su bajo costo computacional en comparación a otras técnicas de cálculo de energía libre mucho más precisas como FEP o TI.

## Capítulo 7

# Conclusiones

El estudio de las interacciones proteína-proteína de la vía mediada por TLR4 muestra un gran potencial para el diseño de nuevas estrategias para el tratamiento de enfermedades como la sepsis, aterosclerosis, tendinitis y artritis reumatoide. La descripción de las vías mediadas por el receptor TLR4 es una tarea altamente compleja a la cuál se han dedicado una amplia gama de publicaciones con la idea de dilucidar los mecanismos que controlan esta parte del sistema inmune innato. Algunas de las características más importantes de la vía, y con mayor respaldo de acuerdo a las publicaciones revisadas, son la homodimerización del receptor TLR4 como punto de partida de la vía y el establecimiento de la importancia de los residuos del lazo BB del dominio TIR del receptor y adaptadores en la transmisión de la señal pro-inflamatoria. Sin embargo la descripción del sistema es incompleta por cuanto se han utilizado representaciones limitadas del sistema para su descripción y este no ha sido caracterizado energéticamente.

En vista de lo anterior este trabajo planteó el uso de algoritmos de acoplamiento, dinámica molecular y herramientas de visualización y análisis energético para reconstruir *in silico* las interacciones generadas entre el receptor TLR4 y los adaptadores MAL, TRAM y MyD88.

De los resultados se desprende que el uso de algoritmos de acoplamiento permite generar una primera aproximación para el modelamiento de las interacciones proteína-proteína. Sin embargo, la definición del algoritmo de acoplamiento debe realizarse de acuerdo a la naturaleza del sistema, ya que no todos los algoritmos son capaces de capturar la fenomenología estudiada. En efecto, la selección de un algoritmo de acoplamiento adecuado permite la producción de una conformación estable a un bajo costo computacional. En el caso de este trabajo el estudio de diferentes herramientas de acoplamiento se estableció que los algoritmos GRAMM y HEX fueron los que mejor respondieron a la fenomenología del sistema de acuerdo a los supuestos considerados. Otro factor determinante a la hora de predecir el acoplamiento receptor-ligando es la calidad y geometría de los modelos del receptor y ligando estudiados. En efecto, la geometría de los modelos afecta directamente los resultados obtenidos. En vista de esto, se establece la necesidad de estudiar una serie de conformaciones distintas de modo de describir exhaustivamente las interacciones proteína-proteína de este u otros sistemas.

En el estudio de las interacciones proteína-proteína, la dinámica molecular es una herramienta clave para el estudio de las interacciones observadas en los complejos. El uso de dinámica molecular permite dar cuenta de contactos que no se observan a partir de una conformación obtenida por algoritmos de acoplamiento, logrando enriquecer el estudio de las interacciones proteína-proteína. Así mismo, los criterios geométricos determinados para la definición de los contactos proteína-proteína establecen una metodología rigurosa capaz de integrar la dinámica de las interacciones proteína-proteína y de los cambios estructurales que en ella se presentan.

De la aplicación de la dinámica molecular al sistema estudiado, se concluye que la estructura inicial a partir de la que se desarrolla la dinámica determina profundamente la trayectoria, los contactos y los cambios conformacionales observados. En vista de esto se concluye que la dinámica molecular, por si sola, no permite establecer un modelo o mecanimo de interacción entre el receptor TLR4 y los adaptadores MAL, TRAM o MyD88. Sin embargo, se convierte en una herramienta importante para el descubrimiento y descripción de sectores que puedan presentar una alta capacidad de formación de contactos.

La introducción de la metodología MM-PBSA para realizar una estimación de la energía libre de unión de los complejos estudiados permitió caracterizar termodinámicamente la formación de los complejos. Sin embargo, de los resultados se desprende que el real valor de la metodología no está en la determinación de un valor absoluto que caracterice el  $\Delta G_{union}$ , sino en su aplicación para la comparación de distintos modelos, de modo de predecir cual presenta un acoplamiento de mayor estabilidad termodinámica.

En lo que respecta a la descripción de las interacciones de la vía mediada por TLR4, se establece que tanto el adaptador MAL como TRAM se unen de manera preferencial a la nueva superficie formada por el homodímero de los dominios TIR de TLR4. Así mismo el adaptador secundario MyD88 aprovecha la superficie que se forma por el acoplamiento de MAL y TLR4. Más aún a partir de los resultados del acoplamiento, dinámica molecular y alineamiento de las secuencias se determinó sitios que tienen la capacidad de establecer contactos entre los componentes de la vía mediada por TLR4. En particular sectores como los lazos BB, CD y hélices  $\alpha A$  y  $\alpha C$  muestran ser sectores de alta interacción. Más aún, se concluye que las características hidrofóbicas del lazo BB lo convierten en una pieza fundamental de la cadena de señalización asociada a los adaptadores TLR4 y TRAM. Adicionalmente, de los resultados se desprende que existen residuos que muestran una gran capacidad generadora de contactos. De los resultados se desprende que los residuos TRP757, PRO712, GLN758, HIS740 de TLR4 corresponden a *hot spots* dentro de la interacción mediada por TLR4. La descripción de estos sectores establece una serie de blancos de estudio para el diseño de péptidos de bloqueo u otras estrategias que permitan interrumpir físicamente la cascada de

señalización característica de esta vía.

Finalmente, es necesario destacar que la metodología planteada en este trabajo se transforma en un procedimiento novedoso y no descrito en trabajos anteriores relacionados a la descripción de interacciones vía TLR. El presente trabajo muestra que el trabajo conjunto de herramientas de acoplamiento, dinámica molecular tradicional, alineamiento de secuencias y MM-PBSA para la estimación de energía libre, permiten modelar las interacciones, probar hipótesis de interacción, brindar explicación a la acción de mutantes y encontrar sectores potenciales de interacción en cualquier tipo de sistema proteína-proteína. Por lo mismo, la aplicación de estos procedimientos se transforma en una herramienta de alto valor, ya que permite explorar, de manera costo efectiva, las interacciones proteína-proteína de cualquier vía. Esto permite la aplicación de esta herramienta en el diseño de otro tipo de tratamientos en cualquier tipo enfermedad cuyo origen sea producto de la interacción de un sistema proteína-proteína.

Un ejemplo de esto, también relacionado al receptor TLR4, se encuentra en el estudio de las relaciones entre el receptor TLR4 y algunas drogas opioides como la heroína o morfina. Un reciente estudio de la Universidad de Adelaida y la Universidad de Colorado en Estados Unidos (*Journal of Neuroscience*, Agosto 2012), muestra que uno de los mecanismos claves de la amplificación de la adicción a drogas opioides es su afinidad con el receptor TLR4. Las drogas opioides unen al receptor TLR4 de una manera similar a como el sistema inmune responde a la presencia de bacterias, solo que en este caso el receptor TLR4 actúa como un amplificador de la adicción. El trabajo de Hutchinson *et al.* (2012) apuntó a describir de manera experimental y con herramientas *in silico* el bloqueo farmacológico de las interacciones morfina-TLR4/MD2. Este trabajo utiliza docking rígido para proveer un posible mecanismo de interacción [130]. En este contexto, las herramientas descritas en este trabajo podrían ser utilizadas de modo de poder describir en detalle las interacciones morfina-TLR4/MD2 y también el como actúan las drogas diseñadas para bloquear la interacción de opioides. El ejemplo anterior es solo una de las posibilidades que provee la metodología planteada en este trabajo, mostrando que efectivamente las herramientas aquí descritas brindan la posibilidad de ampliar la descripción de sistemas de interacción proteína-proteína, fármaco-proteínas o cualquier otro tipo de interacción que pueda ser descrita mediante dinámica molecular tradicional.

## 7.1. Recomendaciones

En vista del trabajo realizado, se plantean las siguientes recomendaciones como puntos interesantes para el desarrollo de nuevas investigaciones y/o la aplicación de lo desarrollado:

- A partir de los resultados del RMSD (C- $\alpha$ ) obtenidos a partir de las simulaciones de dinámica molecular de los complejos, se estableció la necesidad de extender el largo de las simulaciones desarrolladas. Idealmente se debiese considerar una extensión de al menos 100 [ns], ya que mediante esta escala de tiempo es posible observar cambios conformacionales grandes dentro de la estructura del complejo [49].
- De manera de establecer la relevancia de la conformación estudiada o de otras conformaciones, también se recomienda utilizar otras estrategias de exploración intensiva. El uso de herramientas alternativas como *Replica exchange*, *Umbrella Sampling* o *Leap Dynamics* permite favorecer la observación de eventos de menor prevalencia en la dinámica molecular, de modo que se complementa el estudio realizado.
- En vista de los resultados obtenidos mediante MM-PBSA para el cálculo de energía libre, se recomienda utilizar una estrategia distinta para estimar este valor. Se propone utilizar herramientas como integración termodinámica o FEP, ya que son herramientas si bien tienen un mayor costo computacional, permiten calcular de manera precisa el valor de la energía libre de unión asociado a

la formación de un complejo. De todas maneras, si se desea continuar aplicando la metodología MM-PBSA debe reconsiderarse el procedimiento para el cálculo de la energía libre. Adicionalmente debe estudiarse el uso de un software alternativo a APBS para el cálculo de la energía libre de solvatación, de modo de confirmar o corregir los valores obtenidos. El motivo de esto último yace en que fue este último valor el que presentaba grandes variaciones en su valor desde un complejo a otro, sin que se observaran grandes diferencias en el acoplamiento de los complejos.

- De proseguir con el estudio de las interacciones de la vía mediada por el receptor TLR4, sería interesante desarrollar el modelo estructural del adaptador secundario TRIF que forma parte de la cadena de señalización independiente de MyD88 y que según antecedentes se observa que interactúa con el ya formado complejo TLR4-TRAM. Hasta este momento solo se cuenta con la secuencia de aminoácidos extender el estudio utilizando las interacciones TRIF
- En vista de los resultados planteados por Valkov [47] en los que se describe la estructura cristalográfica del adaptador MAL, resulta imperativo estudiar nuevamente las interacciones entre el receptor TLR4 y el adaptador MAL de manera de dar cuenta de si las observaciones propuestas siguen siendo valiosas o si se generan nuevas superficies de interacción.
- Sin perjuicio de los resultados ya presentados, se establece la necesidad de implementar, o en su defecto diseñar, alguna herramienta de minería de datos que permita manejar la gran cantidad de datos que se generan a partir de los algoritmos de acoplamiento. De esta manera se evita que el investigador sesgue, en base a observaciones arbitrarias o poco representativas, los resultados de estudios como los de *docking* y dinámica molecular, caracterizados por la gran cantidad de datos que producen.
- En vista de los resultados obtenidos, resulta interesante el complementar el modelamiento realizado con un estudio experimental. Parte del diseño experimental puede estar enfocado en calcular las constantes de afinidad de modo de que a partir de esta información sea posible calcular estimar la energía libre de unión. Una vez conocido este valor, puede ser posible encontrar aquellas conformaciones que reflejen tal magnitud energética.
- Finalmente, se recomienda estudiar a partir de las interacciones propuestas la efectividad teórica de los péptidos de bloqueo con que se cuenta y estudiar la posibilidad de rediseñar los péptidos de bloqueo estudiados por Patiño (2012).

## Bibliografía

- [1] Janeway, C. *et al.* Innate immunity: The virtues of a nonclonal system of recognition. *Cell*, 91(3):295 – 298, 1997.
- [2] Janeway, C. *et al.* Approaching the asymptote? evolution and revolution in immunology. *Cold Spring Harb Symp. Quant. Biol.*, 54 Pt 1:1–13, 1989.
- [3] West, A. *et al.* Recognition and signaling by Toll-like receptors. *Annu. Rev. Cell Dev. Biol.*, 22: 409–437, 2006.
- [4] Xu, Y. *et al.* Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature*, 408(6808):111–115, 2000.



- [5] Beutler, B. *et al.* Inferences, questions and possibilities in Toll-like receptor signalling. *Nature*, 430 (6996):257–263, 2004.
- [6] Gay, N. y Gangloff, M. Structure and function of Toll receptors and their ligands. *Annu. Rev. Biochem.*, 76:141–165, 2007.
- [7] Lemmon, M. y Schlessinger, J. Regulation of signal transduction and signal diversity by receptor oligomerization. *TIBS*, 19(11):459–463, 1994.
- [8] O’Neill, L y Bowie, A. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat. Rev. Immunol.*, 7(5):353–364, 2007.
- [9] Takeuchi, O. *et al.* Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J. Immunol.*, 169(1):10–14, 2002.
- [10] Hoshino, K. *et al.* Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the *Ips* gene product. *J. Immunol.*, 162(7):3749–3752, 1999.
- [11] Fitzgerald, K. *et al.* Mal (MyD88-adaptor-like) is required for Toll-like receptor-4 signal transduction. *Nature*, 413(6851):78–83, September 2001. ISSN 0028-0836.
- [12] Oshiumi, H *et al.* TIR-containing adapter molecule (TICAM)-2, a bridging adapter recruiting to Toll-like receptor 4 TICAM-1 that induces interferon-beta. *J. Biol. Chem.*, 278(50):49751–49762, 2003.
- [13] Protein Crystallography [En línea]. URL <<http://www.proteincrystallography.org/>>. [Consulta: Noviembre 2011]
- [14] Halperin, I. *et al.* Principles of docking: an overview of search algorithms and a guide to scoring functions. *PROTEINS*, 47(4):409–443, 2002.
- [15] Duhovny, D. *et al.* Efficient unbound docking of rigid molecules. *Algorithms in Bioinformatics*, pages 185–200, 2002.
- [16] Taylor, R. *et al.* A review of protein-sMALL molecule docking methods. *J. Comput. Aided Mol. Des.*, 16:151–166, 2002.
- [17] Yuriev, E. *et al.* Challenges and advances in computational docking: 2009 in review. *J. Mol. Recogn.*, 24(2):149–164, 2011. ISSN 1099-1352.
- [18] Janin, J. y Wodak, S. The third capri assessment meeting toronto,Canada, april 20-21, 2007, July 2007. ISSN 0969-2126.
- [19] Moult, J. *et al.* Critical assessment of methods of protein structure prediction—round vii. *PROTEINS*, 69(S8):3–9, 2007. ISSN 1097-0134.
- [20] Nuñez, R. *et al.* A dimer of the Toll-like receptor 4 cytoplasmic domain provides a specific scaffold for the recruitment of signalling adaptor proteins. *PLoS One*, 2(8):e788, 2007.
- [21] Basith, S. *et al.* In silico approach to inhibition of signaling pathways of Toll-like receptors 2 and 4 by st2l. *PloS one*, 6(8):e23989, 2011.
- [22] Ohnishi, H. *et al.* Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling. *Proc. Natl. Acad. Sci. USA*, 106(25):10260, 2009.
- [23] Dunne, A. *et al.* Structural complementarity of Toll/interleukin-1 receptor domains in Toll-like receptors and the adaptors MAL and MyD88. *J. Biol. Chem.*, 278(42):41443–41451, 2003.

- [24] Andrusier, N. *et al.* Principles of flexible protein–protein docking. *PROTEINS*, 73(2):271–289, 2008.
- [25] Beveridge, D. *et al.* Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu Rev Biophys Biophys Chem.*, 18(1):431–492, 1989.
- [26] Simonson, T. *et al.* Free energy simulations come of age: protein-ligand recognition. *Accounts of chemical research*, 35(6):430–437, 2002.
- [27] Zacharias, M. *Protein-protein complexes: analysis, modeling and drug design.* Imperial College Press, 2da Edición. 2010.
- [28] Medline plus – artritis reumatoidea. [En línea] URL <<http://www.nlm.nih.gov/medlineplus/spanish/ency/article/000431.htm>>. [Consulta: Noviembre 2011]
- [29] Massardo, L. - Corticosteroides [En línea] URL <<http://escuela.med.puc.cl/publ/reumatologia/apuntes/17Corticoesteroides.html>>. [Consulta: Noviembre 2011]
- [30] Ohto, U. *et al.* Crystal structures of mouse and human rp105/md-1 complexes reveal unique dimer organization of the Toll-like receptor family. *J. Mol. Biol.*, 2011.
- [31] Kumar, H. *et al.* Toll-like receptors and innate immunity. *Biochemical and biophysical research communications*, 388(4):621–625, 2009.
- [32] RCSB - Protein Data Bank, Toll Like Receptors, Molecule of The Month [En línea] URL <<http://www.rcsb.org/pdb/101/motm.do?momID=143>>. [Consulta: Enero 2012]
- [33] Nyman, T. *et al.* The crystal structure of the human Toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *J. Biol. Chem.*, 283(18):11861–11865, 2008.
- [34] Toshchakov, V. *et al.* Differential involvement of bb loops of Toll-IL-1 resistance (TIR) domain-containing adapter proteins in TLR4-versus TLR2-mediated signal transduction. *J. Immunol.*, 175(1): 494–500, 2005.
- [35] Jiang, Z. *et al.* Details of Toll-like receptor: adapter interaction revealed by germ-line mutagenesis. *Proc. Natl. Acad. Sci. USA*, 103(29):10961–10966, 2006.
- [36] O’Neill, L y Bowie, A The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat. Rev. Immunol.*, 7(5):353–364, 2007.
- [37] Horng, T. *et al.* TIRAP: An adapter molecule in the Toll signaling pathway. *Nat Immunol*, 2(9): 835–841, September 2001.
- [38] Horng, T. *et al.* The adaptor molecule TIRAP provides signalling specificity for Toll-like receptors. *Nature*, 420(6913):329–333, ISSN 0028-0836. 2002
- [39] Yamamoto, M. *et al.* Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4. *Nature*, 420(6913):324–329. ISSN 0028-0836. 2002
- [40] Rowe, D *et al.* The myristoylation of TRIF-related adaptor molecule is essential for Toll-like receptor 4 signal transduction. *Proc. Natl. Acad. Sci. USA*, 103(16):6299–6304, 2006.
- [41] Poltorak, X. *et al.* Defective LPS signaling in C3H/HEJ and C57BL/10SCCR mice: mutations in TLR4 gene. *Science*;282(5396):2085-8. 2012.
- [42] Gong, J. *et al.* Inhibition of Toll-like receptors TLR4 and 7 signaling pathways by SIGIRR: A computational approach. *Journal of structural biology*, 169(3):323–330, 2010.

- [43] Kagan, J. *et al.* Phosphoinositide-mediated adaptor recruitment controls Toll-like receptor signaling. *Cell*;125(5):943-55. 2006
- [44] Gay, N. *et al.* What the myddosome structure tells us about the initiation of innate immunity. *Trends Immunol*, 32(3):104–109, Mar 2011.
- [45] Kagan, J. *et al.* TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. *Nat Immunol*;9(4):361-8. doi: 10.1038/ni1569. 2008
- [46] Motshwene, P. *et al.* An oligomeric signaling platform formed by the Toll-like receptor signal transducers MyD88 and IRAK-4. *J. Biol. Chem.*, 284(37):25404–25411, 2009.
- [47] Valkov, E. *et al.* Crystal structure of Toll-like receptor adaptor MAL/TIRAP reveals the molecular basis for signal transduction and disease protection. *Proc Natl Acad Sci U S A*;108(36):14879-84. doi: 10.1073/pnas.1104780108. 2011.
- [48] Karplus, M. Molecular dynamics simulations of biomolecules. *Acc Chem Res*, 35(6):321–3. 2002.
- [49] Adcock, S. y McCammon, J. Molecular Dynamics: survey of methods for simulating the activity of proteins. *Chemical Reviews*, 106(5):1589–1615, 2006.
- [50] Bijvoet Center For Biomolecular Research. NMR Spectroscopy Research Group – Molecular modeling practical [En línea] URL <<http://nmr.chem.uu.nl/~tsjerk/course/molmod/index.html#introduction>>. [Consulta: Enero 2012]
- [51] Venken, T. *et al.* An optimized MM/PBSA virtual screening approach applied to an hiv-1 gp41 fusion peptide inhibitor. *PROTEINS*, 2011.
- [52] Fogolari, F. *et al.* MM/PBSA analysis of molecular dynamics simulations of bovine beta-lactoglobulin: free energy gradients in conformational transitions? *PROTEINS*, 59(1):91–103, 2005.
- [53] Guillot, B. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.*, 101(1-3):219–260, 2002.
- [54] TCBG - Theoretical and Computational Biophysics Group, University of illinois. Autoionize plugin, version 1.3 [En línea] URL <<http://www.ks.uiuc.edu/Research/vmd/plugins/autoionize/>>. [Consulta: Marzo 2012]
- [55] GROMACS - Groningen machine for chemical simulations, manual version 4.5.4. Royal Institute of Technology, Uppsala University, Suecia. [En línea] URL <<http://www.gromacs.org/Documentation/Manual>>. [Consulta: Enero 2012]
- [56] Ganoth, A. A molecular dynamics study and free energy analysis of complexes between the mlc1p protein and two iq motif peptides. *Biophysical journal*, 91(7):2436–2450, 2006.
- [57] Swanson, J. *et al.* Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophysical journal*, 86(1):67–74, 2004.
- [58] Lee, M. Calculation of absolute protein-ligand binding affinity using path and endpoint approaches, *Biophys J.*;90(3):864-77. 2006
- [59] Tomasi, J. *et al.* Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chemical Reviews*, 94(7):2027–2094, 1994.
- [60] Weis, A. *et al.* Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *J. Med. Chem.*, 49(22):6596–6606, 2006.

- [61] Kongsted, J. y Ryde, U. An improved method to predict the entropy term with the MM/PBSA approach. *J. Comput. Aided Mol. Des.*, 23(2):63–71, 2009.
- [62] Protein Data Bank. Solution NMR structure of human myeloid differentiation primary response (MyD88). northeast structural genomics target hr2869a. [En línea] URL <<http://www.rcsb.org/pdb/explore/explore.do?structureId=2JS7>>. [Consulta: Agosto 2011]
- [63] Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.*, 2(12):15–32, 2007.
- [64] Laskowski, R. *et al.* PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, 26(2):283–291, 1993.
- [65] Rodriguez, R. Homology modeling, model and software evaluation: three related resources. *Bioinformatics*, 14(6):523–528, 1998.
- [66] Eisenberg, D. *et al.* Method to identify protein sequences that fold into a known three-dimensional structure. *Biotechnology Advances*, 14(4):505–505, 1996.
- [67] Colovos, C. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*, 2(9):1511–1519, 1993. ISSN 1469-896X.
- [68] NIH MBI Laboratory for Structural Genomics and Proteomics - Structural Analysis and VERification Server [En línea] URL <<http://nihserver.mbi.ucla.edu/SAVES/>>. [Consulta: Agosto 2012]
- [69] Wallner, B. *et al.* Can correct protein models be identified? *Protein Science*, 12(5):1073–1086, 2003.
- [70] McGuffin, L. The MODFOLD server for the quality assessment of protein structural models. *Bioinformatics*, (1367-4803 (Linking)), 2008.
- [71] Centos - The Community ENterprise Operating System [En línea] URL <<http://www.centos.org>>. [Consulta: Enero 2012]
- [72] UBUNTU - Official Site [En línea] URL <<http://www.ubuntu.com>>. [Consulta: Enero 2012]
- [73] Condor - High Throughput Computing [En línea] URL <<http://research.cs.wisc.edu/condor/index.html>>. [Consulta: Enero 2012]
- [74] Katchalski-Katzir, E. *et al.* Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A*, 89(6):2195–2199, Mar 1992.
- [75] Ritchie, D. *Parametric Protein Shape Recognition*. Tesis presentada para la obtención del grado *Doctor of Philosophy*, Departments of Computing Science and Molecular and Cell Biology, University of Aberdeen, 1998.
- [76] Kuntz, I. *et al.* A geometric approach to macromolecule-ligand interactions. *J Mol Biol*, 161(2): 269–288, 1982.
- [77] Chen, R. *et al.* Zdock: an initial-stage protein-docking algorithm. *Proteins*, 52(1):80–87, 2003.
- [78] Cheng, T. *et al.* Pydock: electrostatics and desolvation for effective scoring of rigid-body protein-protein docking. *Proteins*, 68(2):503–515, 2007.
- [79] Schrodinger, LLC. The Pymol Molecular Graphics System, version 1.3. 2011.

- [80] Pettersen, E. *et al.* UCSF Chimera: a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.
- [81] Humphrey, W. *et al.* VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33 – 38, 1996.
- [82] Chimera's user guide - Dock prep [En línea] URL <<http://www.cgl.ucsf.edu/chimera/1.2199/docs/UsersGuide/framecontrib.html>>. [Consulta: Enero 2012]
- [83] Ronni, T. *et al.* Common interaction surfaces of the Toll-like receptor 4 cytoplasmic domain stimulate multiple nuclear targets. *Molecular and Cellular Biology*, 23(7):2543–2555, 2003.
- [84] Toshchakov, V. *et al.* Targeting TLR4 signaling by TLR4 Toll/IL-1 receptor domain-derived decoy peptides: Identification of the TLR4 Toll/IL-1 receptor domain dimerization interface. *J. Immunol.*, 186(8):4819–4827, 2011.
- [85] Monie, T. *et al.* Structure and regulation of cytoplasmic adapter proteins involved in innate immune signaling. *Immunological Reviews*, 227(1):161–175, 2009.
- [86] Berggren, A. *et al.* The surface exposed amino acid residues of monomeric proteins determine the partitioning in aqueous two-phase systems. *Biochim Biophys Acta*;1596(2):253-68. 2002
- [87] Salgado, S. *et al.* Is it possible to predict the average surface hydrophobicity of a protein using only its amino acid composition? *Journal of Chromatography A*, 1075(1–2):133 – 143, 2005.
- [88] Hubbard, J. y Thornton, S. Naccess. 1993
- [89] Kyte, J. y Doolittle, R. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105 – 132, 1982.
- [90] Jones, S. y Thornton, J. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA*, 93(1): 13–20, 1996.
- [91] Baptista, A. *et al.* Constant-pH molecular dynamics using stochastic titration. *J. Chem. Phys.*;117, 4184, 2002
- [92] Anandkrishnan, R. *et al.* H++ 3.0: automating pk prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research*, 40(W1):W537–W541, 2012.
- [93] Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *PROTEINS*, 65(3):712–725, 2006.
- [94] Nguyen, T. *et al.* Study of tamiflu sensitivity to variants of a/h5n1 virus using different force fields. *J. Chem. Inf. Model*, 51(9):2266–2276, 2011.
- [95] Kuhn, B. *et al.* Validation and use of the MM/PBSA approach for drug discovery. *Journal of Medicinal Chemistry*, 48(12):4040–4048, 2005.
- [96] Jorgensen, W. *et al.* Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, 1996.
- [97] Dougherty, D. Cation-pi interactions involving aromatic amino acids. *The Journal of Nutrition*, 137 (6):1504S–1508S, June 2007.
- [98] Chakrabarti, P. *et al.* Geometry of nonbonded interactions involving planar groups in proteins. *Progress in Biophysics and Molecular Biology*, 95(1–3):83 – 137, 2007.

- [99] Torshin, I. *et al.* Geometric criteria of hydrogen bonds in proteins and identification of ‘bifurcated’ hydrogen bonds. *Protein Engineering*, 15(5):359–363, 2002.
- [100] Barlow, J. *et al.* Ion-pairs in proteins. *J Mol Biol*;168(4):867-8, 1983
- [101] Donald, J. *et al.* Salt bridges: Geometrically specific, designable interactions. *PROTEINS*, 79(3): 898–915, 2011.
- [102] Walker, R. Amber Tutorials - MM/PBSA [En línea] URL <<http://ambermd.org/tutorials/advanced/tutorial13/>>. [Consulta: Enero 2012]
- [103] Goodsell, D. *et al.* Automated docking of flexible ligands: Applications of Autodock. *J. Mol. Recogn.*, 9(1):1–5, 1996.
- [104] Simossis, V. y Heringa, J. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Research*, 33(suppl 2): W289–W294, 2005.
- [105] McFarland, B. *et al.* Cutting edge: A single, essential hydrogen bond controls the stability of peptide-mhc class ii complexes. *J. Immunol.*, 163(7):3567–3571, 1999.
- [106] Sant, C. *et al.* Individual hydrogen bonds play a critical role in mhc class ii: peptide interactions: implications for the dynamic aspects of class ii trafficking and dm-mediated peptide exchange. *Immunol Rev*;172:239-53, 1999.
- [107] Jin, M. *et al.* Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, 130(6):1071–1082, Sep 2007.
- [108] Khan, J. *et al.* Crystal structure of the Toll/interleukin-1 receptor domain of human IL-1rapl. *J Biol Chem*, 279(30):31664–31670, Jul 2004.
- [109] Vakser, I. *et al.* A systematic study of low-resolution recognition in protein–protein complexes. *Proc Natl Acad Sci U S A*, 96(15):8477–8482, Jul 1999.
- [110] Keskin, O. *et al.* Principles of protein-protein interactions: what are the preferred ways for proteins to interact? *Chem Rev*, 108(4):1225–1244, 2008.
- [111] Fernández-Recio, J. Identification of protein-protein interaction sites from docking energy landscapes. *J Mol Biol*, 335(3):843–865, 2004.
- [112] Li, X. *et al.* Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *J Mol Biol*, 344(3):781–795, Nov 2004.
- [113] Morris, G. *et al.* Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 19(14):1639–1662, 1998.
- [114] Torrie, G. y Valleau, J. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187 – 199, 1977.
- [115] Sugita Y. y Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1–2):141 – 151, 1999.
- [116] Sugita, Y. *et al.* Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.*, 113:6042–6051, October 2000.

- [117] Kleinjung, J. *et al.* Leap-dynamics: efficient sampling of conformational space of proteins and peptides in solution. *FEBS Letters*, 470(3):257 – 262, 2000.
- [118] Lin, Z. *et al.* Structural insights into TIR domain specificity of the bridging adaptor MAL in TLR4 signaling. *PLoS ONE*, 7(4):e34202, 04 2012.
- [119] Crowley, P. y Ubbink, M. Close encounters of the transient kind: protein interactions in the photosynthetic redox chain investigated by NMR spectroscopy. *Accounts of Chemical Research*, 36 (10):723–730, 2003.
- [120] Nooren, I. y Thornton, J. Diversity of protein-protein interactions. *EMBO J*, 22(14):3486–3492, 2003.
- [121] Xiu, X. *et al.* Nicotine binding to brain receptors requires a strong cation-[pgr] interaction. *Nature*, 458(7237):534–537, 2009.
- [122] Celie, P. *et al.* Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in achbp crystal structures, March 2004. *Neuron*;41(6):907-14, 2004.
- [123] Wang, J. *et al.* How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, 21(12):1049–1074, 2000.
- [124] Singh, N. y Warshel, A. Absolute binding free energy calculations: On the accuracy of computational scoring of protein–ligand interactions. *PROTEINS*, 78(7):1705–1723, 2010.
- [125] Li, Y. *et al.* Test MM/PBSA on true conformational ensembles of protein–ligand complexes. *J. Chem. Inf. Model*, 50(9):1682–1692, 2010.
- [126] Hou, T. *et al.* Assessing the performance of the MM/PBSA and MM/GBSA methods. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model*, 51(1):69–82, 2011.
- [127] Homeyer, N. y Gohlke, H. Free energy calculations by the molecular mechanics poisson–boltzmann surface area method. *Molecular Informatics*, 31(2):114–122, 2012.
- [128] Singh, N. y Warshel, A. Toward accurate microscopic calculation of solvation entropies: Extending the restraint release approach to studies of solvation effects. *J. Phys. Chem. C B*, 113(20):7372–7382, 2009.
- [129] Tsui, V. y Case, A. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*, 56(4):275–291, 2000.
- [130] Hutchinson, M. *et al.* Opioid activation of toll-like receptor 4 contributes to drug reinforcement. *J Neurosci.*;32(33):11187-200, 2012.

# Anexos A

## Anexo

### A.1. Secuencias del receptor TLR4 y adaptadores MAL, TRAM y MyD88

Tabla A.1: Secuencias de aminoácidos de los modelos utilizados para TLR4, MAL, TRAM y MyD88. *La numeración que muestra esta tabla es la utilizada a lo largo de este trabajo.*

Proteína	N° Residuo	Secuencia de aminoácidos
TLR4 <sup>1</sup>	662	AGCIKYGRGENIYDAFVIYSSQDEDWVRNELVKNLEEGVPPFQLCLHYRD
	712	FIPGVAIAANIIHEGFHKSARKVIVVVSQHFIQSRWCIFEYEIAQTWQFLS
	762	SRAGIIFIVLQKVEKTLLRQQVELYRLLSRNTYLEWEDSVLGRHIFWRRL
	812	RKALLDGKSWNPE
MAL	74	ASDSGSSRWSKDYDVCVCHSEEDLVAAQDLVSYLEGSTASLRCFLQLRDA
	124	TPGGAIVSELCQALSSSHCRVLLITPGFLQDPWCKYQMLQALTEAPGAEG
	174	CTIPLLSGLSRAAYPPELRFMYVVDGRGPDGGFRQVKEAVMRCKLLQEGE <sup>2</sup> <sup>1</sup>
MyD88	157	HMPERFDAFICYCPSDIQFVQEMIRQLEQTNYRLKLCVSDRDVLPGTCVW
	207	SIASELIEKRCRRMVVVVSDDYLSKECDFQTKFALSLSPGAHQKRLIPI
	257	KYKAMKKEFPSILRFITVCDYTNPCTKSWFWTRLAKALSLP
TRAM	65	VEEMFEEEEAEVFLKFVILHAEDDTDEALRVQNLLQDDFGIKPGIIFAE
	115	MPCGRQHLQNLDDAVNGSAWTILLTENFLRDTWCNFFQYTSLMNSVNRQ
	165	HKYNSVIPMRPLNPLPRERTPFALQTINALEEESRGFPTQVERIFQESV
	215	YKTQQTIVKETRNMVQRQFIA

Secuencia de cada uno de los monómeros que conforman el homodímero TLR4

<sup>2</sup> En el modelo propuesto los últimos aminoácidos son YLQTLSWX en lugar de CKLLQEGE

### A.2. Características del software de acoplamiento estudiado



Tabla A.2: Principales características de los algoritmos de acoplamiento estudiados

Nombre	Representación	Método de Búsqueda	Tipo de Licencia	Plataforma	Versión	Año	Soporte	Referencia
3D Dock	Discreta	Sistemática	Libre Académica	Descargable	2.0	2001	-	Sternberg et al. (1998)
ATTRACT (Ptools)	Granos	Guiada	Libre Académica	Descargable	1.0.4c	2009	+	Saladin et al. (2008)
Autodock	Discreta	Sistemática	Libre Académica	Descargable	4.2.3	2010	+	Morris et al. (2010)
BiGGER	Discreta	Sistemática	No Info	No disponible	—	—	-	Palma et al. (2000)
CLUSPRO	Discreta	Sistemática	Libre Académica	Web/Descargable	2.0	2009	+	Comeau et al. (2003)
DOCK	Discreta	Guiada	Libre Académica	Descargable	6.4	2010	+	Kuntz et al. (1982)
DOT	Discreta	Sistemática	Libre Académica	Descargable	2.0	2008	+	Mandell et al. (1995)
Escher NG	Discreta	Sistemática	Libre	Descargable	1.2	2007	-	Ausiello et al. (1997)
FRODOCK	Discreta	Sistemática	Libre	Descargable	1.04	2010	+	Chacon et al. (2009)
GRAMM	Discreta	Sistemática	Libre Académica	Descargable/Web	1.03	1998	-	Vakser (1991)
HADDOCK	Atómica	Guiada	Libre Académica	Descargable	2.1	2010	+	Bonvin et al. (2003)
HEX	Discreta	Sistemática	Libre Académica	Descargable/Web	6.3	2010	+	Ritchie (2000)
ICM	Atómica	Guiada	Comercial/Libre	Descargable	2.8	—	+	Abagyan (1985)
MolFit	Discreta	Sistemática	Libre Académica	Descargable	2.0	2006	+	Redington (2001)
PatchDock	Superficie	Sistemática	Libre Académica	Web	1.3	2005	+	Duhovny (2002)
RosettaDock	Atómica/Granos	Guiada	Libre Académica	Descargable/Web	3.2.1	2011	+	Gray et al. (2003)
SKE-Dock	Superficie	Sistemática	Libre Académica	Email	—	—	-	Terachi et al. (2005)
ZDOCK	Discreta	Sistemática	Libre Académica	Descargable	3.0.1	2007	+	Chen, Weng (2002)

+ : El software presenta soporte técnico

- : El software no presenta soporte técnico

Tabla A.3: Aplicaciones de los algoritmos de acoplamiento utilizados

Nombre	Aplicaciones
3D Dock	Perez Cano – Structural prediction of protein-RNA interaction by computational docking with propensity-based statistical potentials. (2010)
ATTRACT (Ptools)	Orry - Modeling and Docking the Endothelin G-Protein-Coupled Receptor (2000)
BiGGER	Zacharias, Fiorucci - Binding site prediction and improved scoring during flexible protein-protein docking with ATTRACT (2010)
CLUSPRO	Bustos - A model for the interaction between plant GAPN and 14-3-3zeta using protein-protein docking calculations, electrostatic potentials and kinetics.(2005)
DOCK	Dell'Acqua - Electron Transfer Complex between Nitrous Oxide Reductase and Cytochrome c552 (2008)
DOT	Katebi - Structural interpretation of protein-protein interaction network (2009)
Escher NG	Pons - FlgM anti-sigma factors: identification of novel members of the family, evolutionary analysis, homology modeling (2006)
GRAMM	Teotico et al. - Docking for fragment inhibitors of AmpC -lactamase (2008)
HADDOCK	Roberts and Pique - Definition of the Interaction Domain for Cytochrome c on Cytochrome c Oxidase: III. Prediction of the Docked Complex (1999)
HEX	Fahhan - Simulation of different truncated p16(INK4a) forms and in silico study of interaction with Cdk4.(2009)
ICM	Backos et al. - Posttranslational modification and regulation of glutamate-cysteine ligase by the $\alpha,\beta$ -unsaturated aldehyde 4-hydroxy-2-nonenal (2011)
MolFit	Canals et al. - Adenosine A2A-Dopamine D2 Receptor-Receptor Heteromerization (2003)
PatchDock	Bhattacharya - Interaction analysis of TerX/Y two component system from Mycobacterium tuberculosis. (2010)
RosettaDock	Xu - Model for the three-dimensional structure of vitronectin: predictions for the multi-domain protein from threading and docking (2001)
ZDOCK	Williams et al. - Structural Insights into the Interaction of Insulin-like Growth Factor 2 with IGF2R Domain 11 (2008)
	Bonvin et al. - Modelling protein-protein complexes involved in the cytochrome c oxidase copper-delivery pathway (2007)
	Fahhan - Simulation of different truncated p16(INK4a) forms and in silico study of interaction with Cdk4 (2009)
	Ritchie - Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2 (2003)
	Pecic - Affinity of aporphines for the human 5-HT2A receptor: insights from homology modeling and molecular docking studies (2010)
	Alzate Morales - Computational study on the interaction of N1 substituted pyrazole derivatives with B-raf kinase (2010)
	Berchansky - Prediction of HIV-1 entry inhibitors neomycin-arginine conjugates interaction with the CD4-gp120 binding site (2007)
	Heffetz - Protein-protein docking: progress in CAPRI rounds 6-12 using a combination of methods: the introduction of steered solvated molecular dynamics (2007)
	Fong et al - Protein-Protein Recognition between Acyltransferases and Acyl Carrier Proteins in Multimodular Polyketide Synthases (2010)
	Benyamini - A model for the interaction between NF- $\kappa$ B and ASP2 suggests an I-kappa-B-like binding mechanism (2009)
	Sivasubramanian et al. - Structural model of the mAb 806-EGFR complex using computational docking followed by computational and experimental mutagenesis (2006)
	Sivasubramanian et al. - Modeling the structure of mAb 14B7 bound to the anthrax protective antigen (2008)
	Dell'Orco - In silico screening of mutational effects on enzyme-proteic inhibitor affinity: a docking-based approach (2007)
	Weng et al. - The performance of ZDOCK and ZRANK in rounds 6-11 of CAPRI (2007)

Tabla A.4: Número de citas que recibe el paper de referencia de cada algoritmo de acoplamiento

Nombre	Paper de referencia	Nº Citas*
3D Dock	A computational system for modelling flexible protein-protein and protein-DNA docking. Sternberg et al. (1998)   11	27
ATTRACT (Ptools)	Protein-protein docking using 3D-Dock in rounds 3, 4, and 5 of CAPRI PTools: an opensource molecular docking library. Saladin et al. (2008)	8
BiGGER	BiGGER: A New (Soft) Docking Algorithm for Predicting Protein Interactions. Nuno Palma et al. (2000)	<b>230</b>
CLUSPRO	ClusPro: a fully automated algorithm for protein-protein docking. Comeau et al. (2004)	89
DOCK	A geometric approach to macromolecule-ligand interactions. Kuntz et al. (1982)	<b>1231</b>
DOT	Surveying Molecular Interactions With DOT. Mandell et al. (1995)	42
Escher NG	Escher: A new docking procedure applied to the reconstruction of protein tertiary structure. Aussello et al. (1997)	55
FRODOCK	FRODOCK: a new approach for fast rotational protein-protein docking. Garzón et al. (2009)	6
GRAMM	Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Vakser et al. (1992)	<b>599</b>
HADDOCK	HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information. Bonvin et al. (2002)	<b>582</b>
HEX	Protein Docking Using Spherical Polar Fourier Correlations. Ritchie (2000)	<b>276</b>
ICM	ICM - a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. Abagyan et al. (1994)	<b>612</b>
MolFit	MOLFIT: A computer program for molecular superposition. Redington (2001)	6
PatchDock	PatchDock and SymmDock: servers for rigid and symmetric docking. Duhovny et al. (2002)	147
RosettaDock	Protein-Protein Docking with Simultaneous Optimization of Rigid-Body Displacement and Side-Chain Conformations. Gray et al. (2003)	<b>329</b>
ZDOCK	ZDOCK - An Initial Stage Protein docking algorithm. Weng et al. (2003)	<b>363</b>

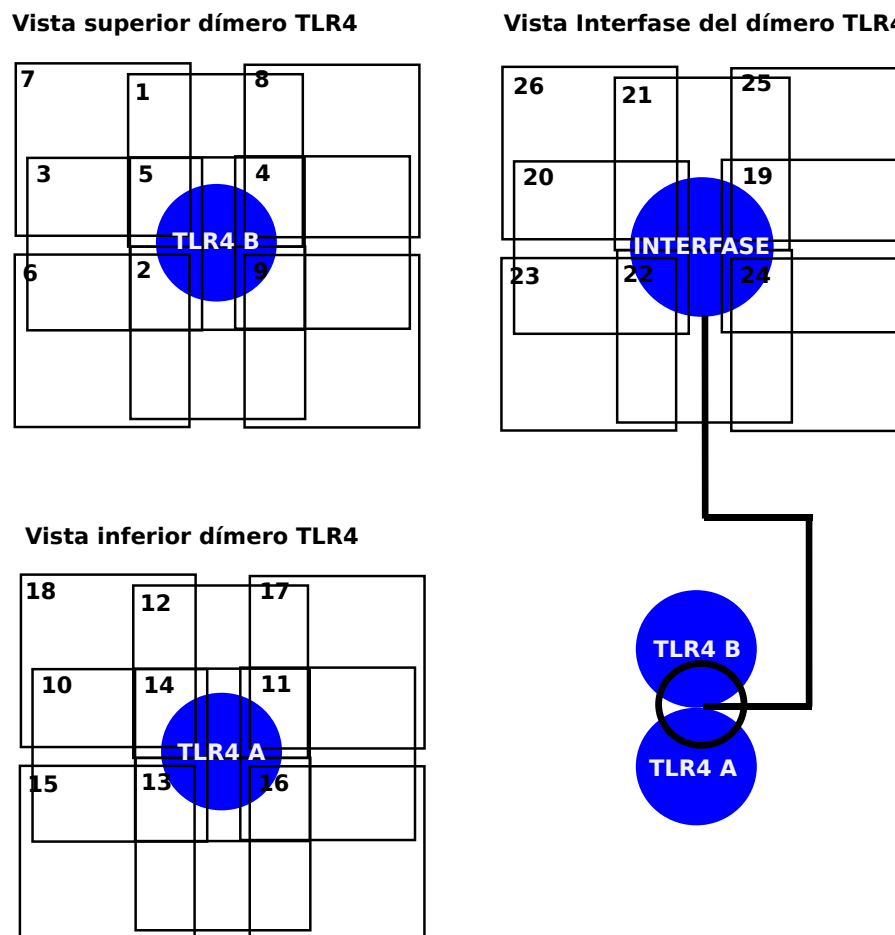
\*Información extraída de Google Scholar

Tabla A.5: Tiempo promedio de ejecución de algoritmos de acoplamiento

Nombre	Duración Simulación	Referencias
3D Dock	No se encontró información	—
ATTRACT (Ptools)	19 horas - single core Intel Xeon 1.86 Ghz - 53.000 configuraciones iniciales RB Docking – Bovine alpha chymotrypsinogen A/ Pancreatic secretory inh	Saladin et al. (2009)
BiGGER	2 a 8 horas - 25 casos de estudio - Intel Pentium II 450 Mhz	Palma et al. (2000)
CLUSPRO	4 horas - 16 Procesadores IBMpSeries 690 1.3 Ghz. 32 Gb memoria compartida (Filtrado limitante)	Comeau et al. (2004)
DOCK	No se encontró información	—
DOT	No se encontró información	—
Escher NG	No se encontró información	—
PRODOCK	30 minutos - 2.2 Mhz - HyHel-5/Lisozima	Garzón et al. (2009)
GRAMM	Horas (High Resolution) - Segundos (Low resolution)	Vakser et al. (1999)
HADDOCK	2 Días - 10 estaciones 1.3 GHz Procesador AMD - EIN-HPr complex (247 / 85 aa, 25 AIRs)	Bonvin et al. (2002)
HEX	20 minutos - 2.2 Mhz - HyHel-5/Lisozima	Ritchie et al. (2008)
ICM	5 minutos por ligando por procesador	molsoft.com/gui/start-dock.html
MolFit	No se encontró información	—
PatchDock	20 a 60 minutos - Pentium II 500 Mhz 512 MB - Set de proteínas	Duhovny et al (2002)
RosettaDock	No se encontró información	—
ZDOCK	4 minutos por complejo - 16 procesadores IBM-SP4	Weng et al. (2003)

### A.3. Acoplamiento sitio dirigido

Figura A.1: Esquema de la estrategia sitio dirigido utilizada con Autodock y Vina



## A.4. Parámetros utilizados en la implementación del protocolo de dinámica molecular en GROMACS

Tabla A.6: Parámetros del archivo de configuración (.mdp) para cada uno de las etapas de MD

Parámetro	Minimización	Posición Restringida	NPT	Producción
define	-DFLEXIBLE	-DPOSRES	-DFLEXIBLE	-DFLEXIBLE
integrator	cg	md	md	md
emtol	10	—	—	—
dt	—	0,001	0,001	0,002
nsteps	-1	50000	100000	3000000
nstcgsteep	100	—	—	—
nstxout	—	1000	1000	10000
nstvout	—	1000	1000	10000
nstfout	—	0	0	0
nstlog	—	1000	1000	10000
nstxtcout	—	1000	1000	10000
nstenergy	1	1	1	10000
nstcalcenergy	—	10	10	10
energygrps	System	Protein Non-Protein	Protein Non-Protein	Protein Non-Protein
ns_type	grid	grid	grid	grid
nstlist	—	10	10	10
rlist	1,4	1,4	1,4	1,4
coulombtype	PME-Switch	PME-Switch	PME-Switch	PME-Switch
rcoulomb	1,2	1,2	1,2	1,2
rcoulomb_switch	1,0	1,0	1,0	1,0
fourierspacing	0,12	0,12	0,12	0,12
pme_order	4	4	4	4
ewald_rtol	0,00001	0,00001	0,00001	0,00001
vdw-type	Shift	Shift	Shift	Shift
rvdw	1,2	1,2	1,2	1,2
rvdw_switch	0	0	0	0
constraints	none	all-bonds	all-bonds	all-bonds
constraints-algorithm	—	Lincs	Lincs	Lincs
lincs-order	—	4	4	4
lincs-iter	—	1	1	2
lincs-warnangle	—	30	30	30
pbc	xyz	xyz	xyz	xyz
tcoupl	—	V-rescale	V-rescale	V-rescale
Tc-grps	—	Protein Non-Protein	Protein Non-Protein	Protein Non-Protein
tau_t	—	0,1 0,1	0,1 0,1	0,1 0,1
ref_t	—	200 200	300 300	300 300
Pcoupl	—	—	Berendsen	Berendsen
Pcoupltype	—	—	Isotropic	Isotropic
tau_p	—	—	0,5	0,5
compressibility	—	—	4,50E-05	4,50E-05
ref_p	—	—	1,0	1,0
gen_vel	—	yes	no	no
gen_temp	—	200	—	—
gen_seed	—	9999	—	—
continuation	—	—	yes	yes

A continuación se describe de manera breve la función que cumple cada uno de los parámetros señalados:

- *define*: Permite controlar el uso de los archivos de topología. La opción -DFLEXIBLE indica a *grompp* que debe usar moléculas de agua flexibles en la topología. Por su parte la opción -DPOSRES puede ser utilizada para generar restricciones de posición en los átomos del sistema.
- *integrator*: Permite determinar el modo de ejecución de modo de realizar minimización del sistema, dinámica molecular o análisis de modos normales, entre otras opciones. La opción *cg* se refiere al uso del algoritmo de gradiente conjugado para la minimización de energía. Por su parte *md* se refiere a ejecución de dinámica molecular utilizando un algoritmo *leap-frog* para integrar las ecuaciones de movimiento de Newton.
- *emtol* [ $\text{kJ mol}^{-1} \text{ nm}^{-1}$ ]: Corresponde a la tolerancia utilizada para definir la precisión deseada durante la etapa de minimización. El valor limita de la convergencia si la máxima fuerza del sistema es menor que este valor.
- *dt* [ps]: Paso de integración para el tiempo. Se utiliza solo cuando se utilizada la opción *md* en el parámetro *integrator*.
- *nsteps*: Número máximo de pasos utilizados para la integración o minimización.
- *nstcgsteep* [pasos]: Frecuencia para la cual se desarrolla una iteración del método del gradiente al utilizar el método de gradiente conjugado al minimizar la energía de un sistema.
- *nstxout* [pasos]: Frecuencia con que se escribe trayectorias de salida, las últimas coordenadas del sistema siempre se escriben como salida.
- *nstvout* [pasos]: Frecuencia con que se escriben las velocidades en el archivo de salida.
- *nstfout* [pasos]: Frecuencia con que se escriben las fuerzas en el archivo de salida del sistema.
- *nstlog* [pasos]: Frecuencia con que se escriben las energías del sistema en el archivo *log*.
- *nstxtcout*: Frecuencia con que se escriben las trayectorias del sistema en el archivo de salida .xtc
- *nstenergy*: Frecuencia con que se escriben las energías del sistema en el archivo de energía .edr
- *nstcalcenergy*: Frecuencia con que se calculan las energía del sistema. Esta opción es relevante al desarrollar simulaciones de dinámica molecular, ya que afecta el desempeño al realizar cálculos en paralelo. Esto se debe a que el cálculo de energía requiere de la comunicación global de todos los procesos lo que se puede convertir en el cuello de botella en las simulaciones con alta paralelización.
- *energygrps*: Grupos del sistema para los cuales se almacena los archivos de energía de salida.
- *ns\_type*: Parte de los cálculos que realiza Gromacs se llevan a cabo utilizando una lista de átomos vecinos o próximos a cada átomo del sistema. Este parámetro define la manera en la que se actualiza esa lista. Para el caso de la opción *grid* se genera una grilla en la caja definida para la simulación y se chequean aquellos átomos que se encuentra en las grillas cercanas, en cambio para la opción *simple* se chequean todos los átomos del sistema.
- *nstlist* [pasos]: Frecuencia con que se actualiza la lista de átomos vecinos.
- *rlist* [nm]: Distancia de corte para la lista de átomos vecinos de corto alcance.

- *coulombtype*: Tratamiento utilizado para el cálculo de energía electrostáticas. La opción PME-Switch implica una combinación de PME y una función switch para el cálculo de las energías electrostáticas. La dimensión de la grilla de cálculo es controlada con *fourierspacing* y el orden de la interpolación utilizada con *pme\_order*.
- *rcoulomb* [nm]: Distancia de corte para el cálculo de energía electrostáticas
- *rcoulomb\_switch* [nm]: Distancia para la cual se comienza a implementar la función Switch sobre el cálculo de energía electrostática.
- *fourierspacing* [nm]: Corresponde al mayor espaciado que puede ser utilizado en la grilla para FFT al usar PME para el cálculo de la energía electrostática.
- *pme\_order*: Orden de la interpolación para PME
- *ewald\_rtol*: Tolerancia relativa usada en PME. Disminuir este valor da una suma más precisa, pero requiere de memoria adicional.
- *vdw-type*: Tratamiento para el cálculo de interacciones de no enlace. La opción Shift utiliza un potencial de tipo Lennard-Jones. Las fuerzas decaen suavemente a cero para distancias entre *rvdw\_switch* y *rvdw*.
- *rvdw*: Distancia de corte para la función de Lennard-Jones
- *rvdw\_switch*: Distancia para la cual comienza el decaimiento de la función de Lennard-Jones
- *constraints*: Tipo de restricciones impuestas sobre los enlaces del sistema. En el caso de la opción *none*, no se usan restricciones con excepción de aquellas definidas explícitamente en la topología. Esto implica que los enlaces y ángulos son representados por un potencial armónico (u otro tipo). La opción *all-bonds* implica que todos los enlaces se transforman en restricciones del sistema
- *constraints-algorithm*: Corresponde al tipo de algoritmo utilizado para mantener las restricciones sobre el sistema. En este caso se usa el algoritmo LINCS (*LINear Constraint Solver*)
- *lincs-order*: Define la precisión del algoritmo por medio de la definición del número de matrices usados en la expansión durante la inversión de matriz. Para una dinámica molecular normal se usa un orden 4.
- *lincs-iter*: Número máximo de iteraciones implementadas para la corrección de los largos rotacionales en LINCS.
- *lincs-warnangle* [°]: Máximo ángulo que puede rotar un enlace sin que LINCS lo note.
- *pbcs*: Parámetro que define el uso de condiciones de borde periódicas. *xyz* implica el uso de condiciones de borde periódicas en todas las direcciones.
- *tcoupl*: Define el algoritmo utilizado para acoplar la temperatura del sistema. En este caso se acopla la temperatura mediante V-rescale, un algoritmo de acoplamiento similar al de un termostato de Berendsen pero que incluye el rescalamiento de la velocidades del sistema, por medio de un término estocástico.
- *Tc-grps*: Grupos para los cuales se desea acoplar de manera separada la temperatura
- *tau\_t* [ps]: Constante de tiempo para el acoplamiento de temperatura
- *ref\_t* [K]: Temperatura de referencia para el acoplamiento de temperatura, una por cada grupo definido.



- *Pcoupl*: Define el algoritmo utilizado para acoplar la presión del sistema. La opción *Berendsen* implica la relajación exponencial del acoplamiento de la presión con una constante de tiempo  $\tau_p$ .
- *Pcoupltype*: Define el tipo de acoplamiento de la presión. En este caso la opción *isotropic* implica un acoplamiento a una presión isotrópica. La compresibilidad y la presión de referencia del sistema se definen respectivamente con *compressibility* [ $\text{bar}^{-1}$ ] y *ref-p* [ $\text{bar}^{-1}$ ].
- *tau-p*: Constante de tiempo para el acoplamiento de la presión.
- *compressibility*[ $\text{bar}^{-1}$ ]: Compresibilidad del sistema.
- *ref-p* [ $\text{bar}^{-1}$ ]: Presión de referencia para el sistema.
- *gen\_vel*: Parámetro determina la generación de las velocidades para los átomos del sistema al inicio de la dinámica molecular. En este caso, la opción *yes* implica la determinación de las velocidades iniciales del sistema de acuerdo a una distribución de Maxwell a una temperatura *gen\_temp*, con una semilla aleatoria *gen\_seed*
- *gen\_temp* [K]: Temperatura para genera la distribución Maxwell.
- *gen\_seed*: Parámetro usado para la determinación aleatoria de las velocidades.
- *continuation*: Parámetro seteado en *yes* en los casos que sea continuar con una simulación por etapas. Las condiciones de salida de la última etapa son considerados como parámetros de entrada para la nueva etapa.

## A.5. Diferencias entre modelos del receptor TLR4 y adaptadores propuestos por Nuñez *et al.* (2007) y Salgado (2011)

Tabla A.7: RMSD entre estructuras del dominio TIR de TLR4 y adaptadores propuestas por Nuñez (2007) y Salgado (2010)

Estructura	RMSD Local			
	C-alpha	Backbone	Heavy	All
TLR4 (Cadena A)	0.83	0.89	1.82	1.82
TLR4 (Cadena B)	0.53	0.57	1.8	1.8
MAL	2.12	2.15	2.56	2.65
TRAM	1.65	1.61	2.54	2.74
Estructura	RMSD Global			
	C-alpha	Backbone	Heavy	All
TLR4 (Cadena A)	0.83	0.89	1.82	1.82
TLR4 (Cadena B)	0.53	0.57	1.8	1.8
MAL	4.86	4.83	5.4	5.55
TRAM	4.55	4.51	5.49	5.65

$\text{RMSD}_{local}$ : RMSD calculado entre el modelo de Nuñez y el de Salgado con la superposición directa de la región no alineada. Mide la precisión de la conformación local del modelo.

$\text{RMSD}_{global}$ : RMSD entre el modelo de Nuñez y Salgado luego de la superposición de al menos 5 regiones vecinas en cada extremo de los modelos. Mide tanto la precisión de la conformación local como la orientación global entre los modelos.

## A.6. Histogramas RMSD C $\alpha$ durante simulaciones de dinámica molecular

Figura A.2: Histograma de RMSD C $\alpha$  de TLR4 y MAL durante la simulación del complejo tetramérico TLR4-MAL

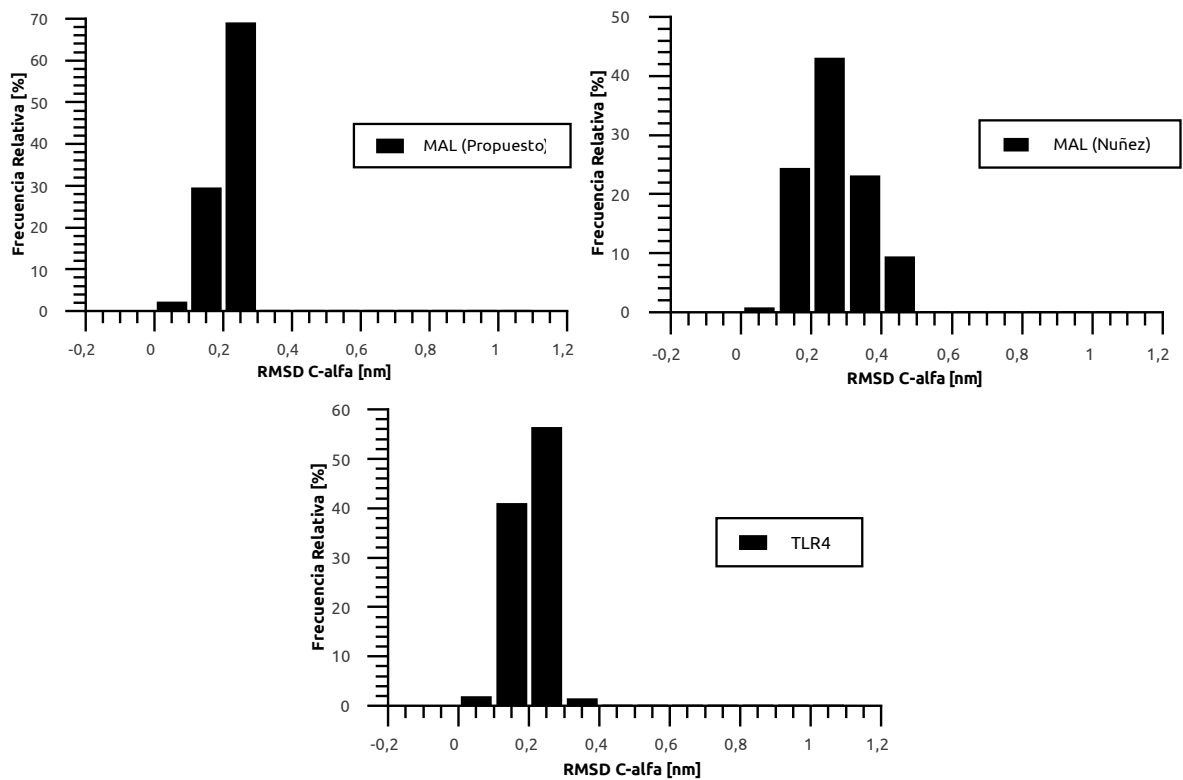


Figura A.3: Histograma de RMSD C $\alpha$  de TLR4, MAL y MyD88 generado a partir de la simulación del complejo TLR4-MAL-MyD88 (Modelo 1)

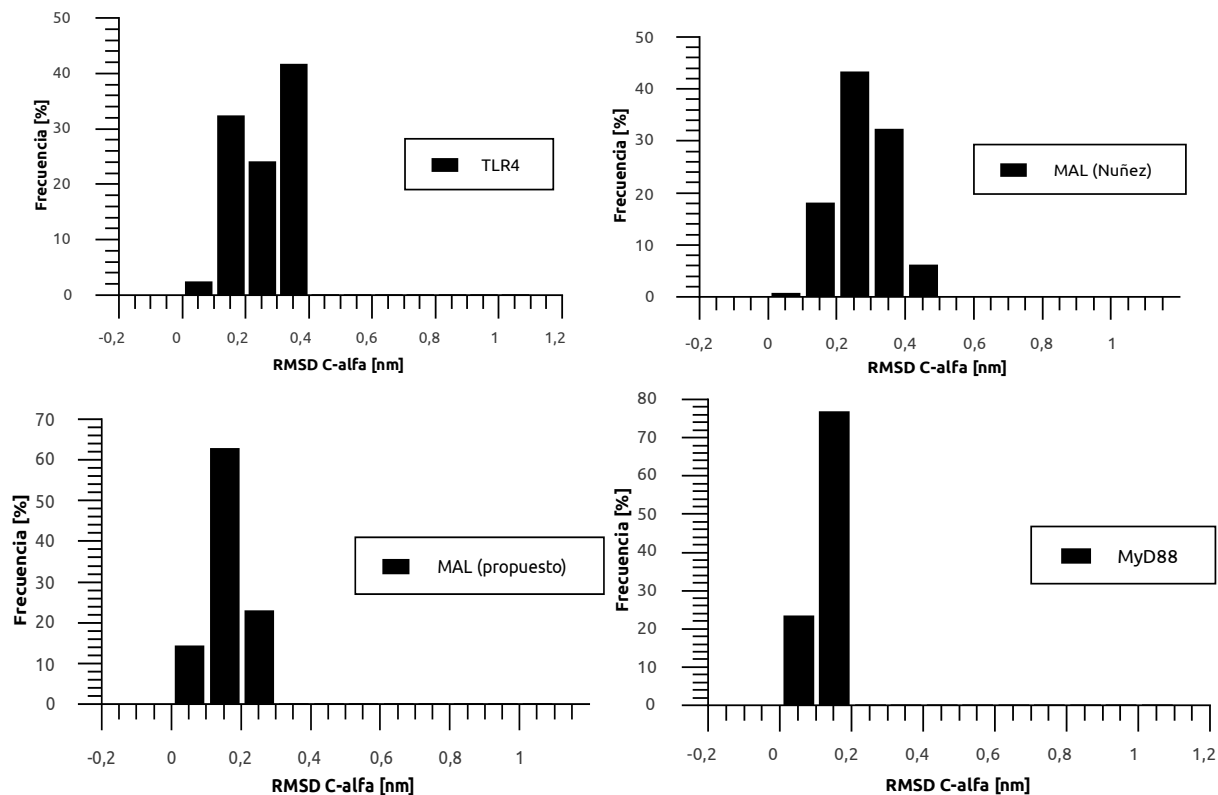
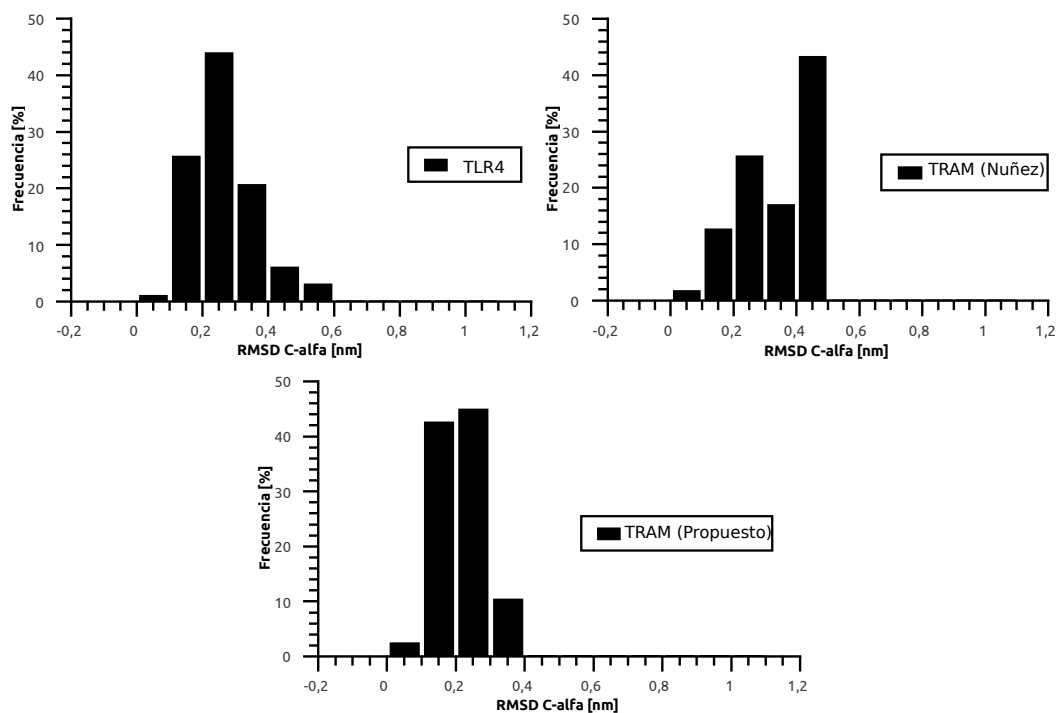


Figura A.4: Histograma de RMSD C $\alpha$  de TLR4 y TRAM durante simulación del complejo TLR4-TRAM



## A.7. RMSF de los residuos de TLR4 y adaptadores durante las simulaciones de dinámica molecular

Figura A.5: Fluctuación cuadrática media de TLR4 y MAL durante dinámica molecular del complejo TLR4-MAL

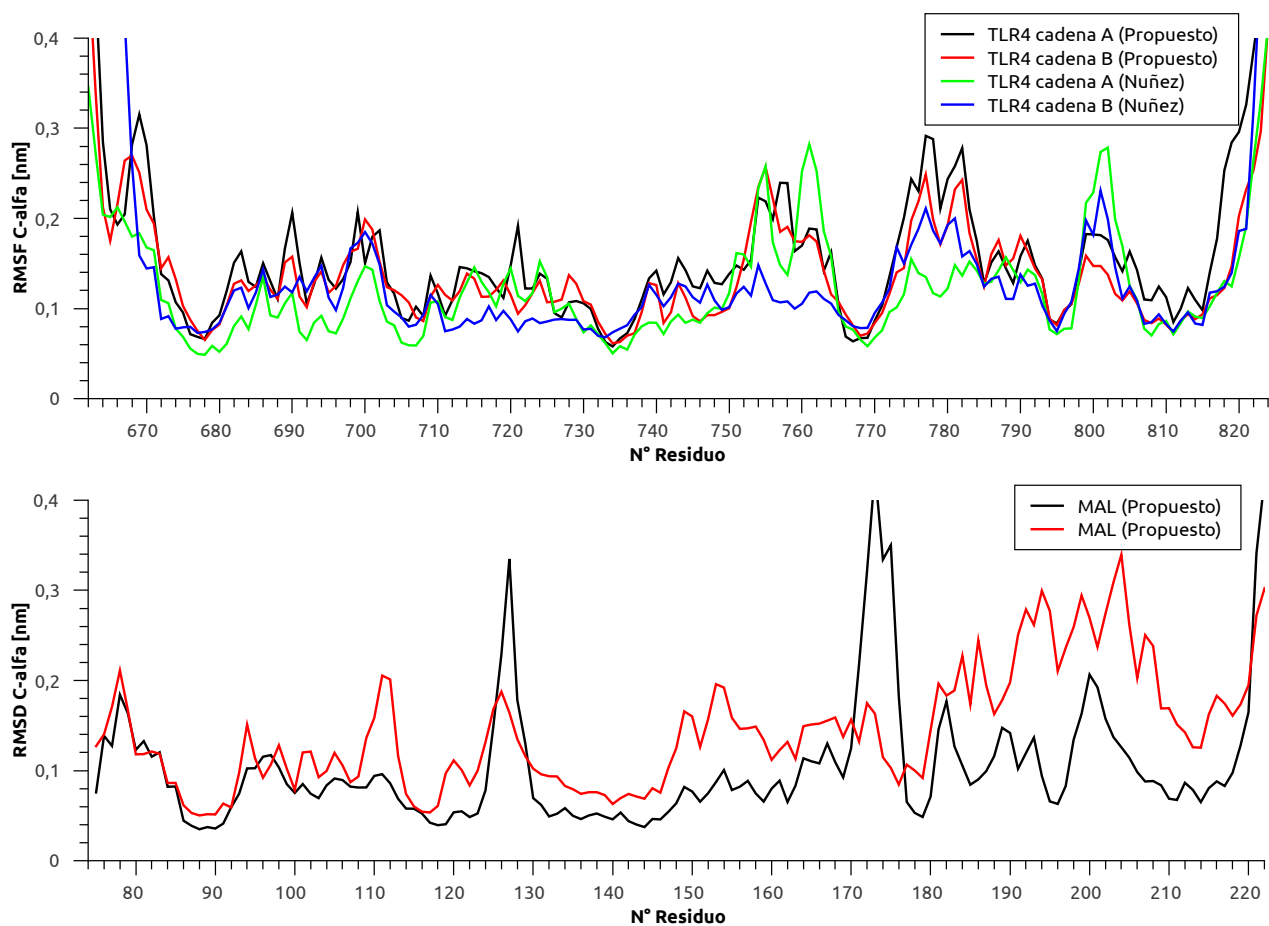


Figura A.6: Fluctuación cuadrática media de TLR4, MAL y MyD88 en dinámica molecular dle complejo TLR4-MAL-MyD88

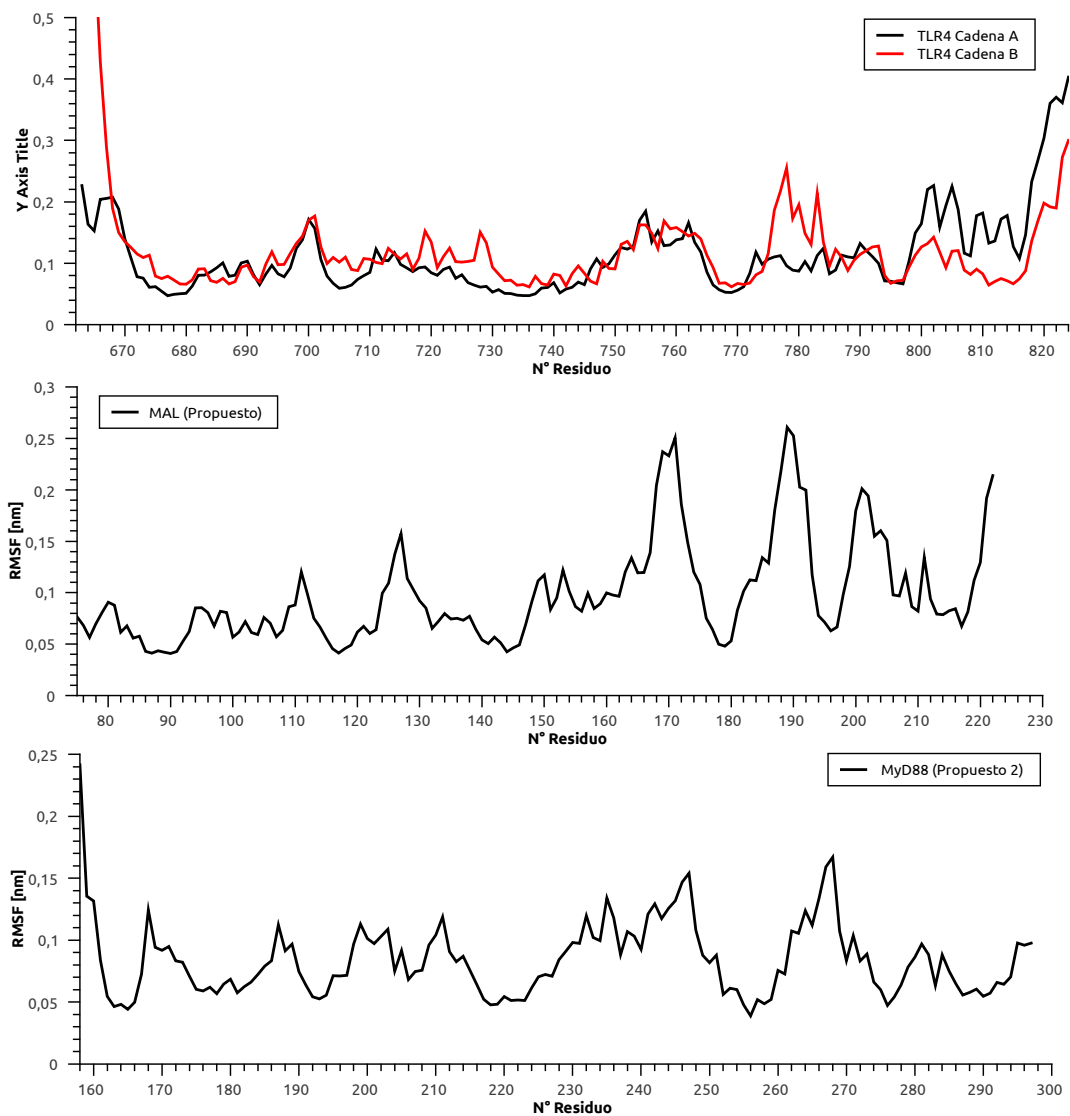
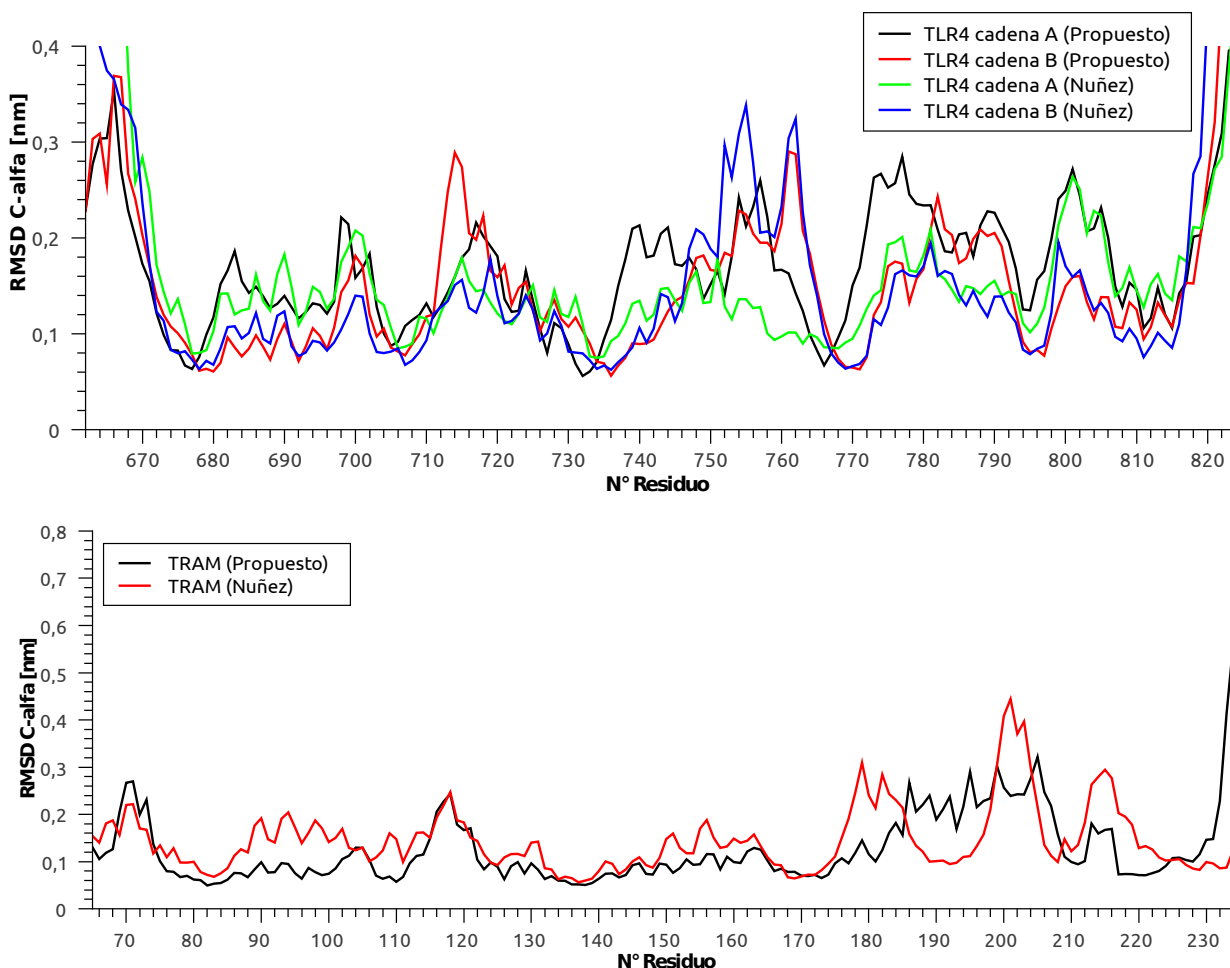


Figura A.7: Fluctuación cuadrática media de los residuos de TLR4 y TRAM durante dinámica molecular del complejo TLR4-TRAM



### A.8. Otros contactos determinados a partir de la dinámica molecular de los complejos estudiados

Tabla A.8: Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-MAL

Tipo de Enlace	TLR4-MAL (Nuñez)		TLR4-MAL (Propuesto)	
	Promedio	Desv. Estándar	Promedio	Desv. Estándar
Puentes de Hidrógeno	11	3	4	2
Cation- $\pi$	3	1	1	1
Hidrofóbico	5	2	10	3
Iónicos	1	1	0	0
Aromáticos	1	1	0	0

[H]

Tabla A.9: Contactos encontrados en la zona de interacción TLR4-MAL vía MD según modelo de acoplamiento. A: TLR4 cadena A, B: TLR4 cadena B, N: MAL modelo Nuñez, G: MAL modelo propuesto

Contactos TLR4-MAL						
Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
THR124	N	GLU685	A	40.9		
GLN135	N	SER744	A	28.2	SR744-745AA	I
GLU167	N	GLN758	B	44.9	Q758A	II
GLN220	N	GLN755	B	35.9	Q755A	I
ARG669	B	GLU190	N	20.3		
ARG763	B	GLU167	N	41.5	R763A	II
ARG787	B	GLU221	N	32.3	RL787-788AA	II
LYS819	B	ALA164	N	30.2		
SER682	A	GLU132	N	44.2	YSS680-682AAA	I
GLN683	A	ASP122	N	25.6	Q683A	II
GLN683	A	TYR86	N	39.5	Q683A	II
GLN739	A	ASP85	N	30.6	QH739-740AA	III
HIS740	A	ASP85	N	33.6	QH739-740AA	III
GLN743	A	SER139	N	42.2	IQ742-743AA	III
GLN683	B	SER113	G	36.5	Q683A	II
ARG689	B	GLU108	G	41.9		
ARG669	A	GLU167	G	20.9		
HIS724	A	SER83	G	24.6	H724A	II
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ILE665	A	LEU162	G	45.7		
TYR674	A	ALA128	G	36.8		
ILE673	B	MET161	N	24.8		
Contactos Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ARG115	G	HIS740	B	46.0	QH739-740AA	I
ARG669	B	TYR187	N	27.8		

Efecto de las mutaciones:

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial)

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo

Tabla A.10: Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-MAL-MyD88

Tipo de Enlace	TLR4-MyD88		MAL-MyD88	
	Promedio	Desv. Estándar	Promedio	Desv. Estándar
Puentes de Hidrógeno	5	2	4	1
Cation- $\pi$	0	0	1	0
Hidrofóbico	0	0	1	1
Iónicos	1	0	0	0
Aromáticos	0	0	0	0

[H]

Tabla A.11: Contactos encontrados vía MD en la zona de interacción TLR4-MAL-MyD88 según modelo de acoplamiento. *A: TLR4 cadena A, B: TLR4 cadena B, N: MAL modelo Nuñez, G: MAL modelo propuesto, M: MyD88 (Modelo1)*

Contactos TLR4-MAL-MyD88						
Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
ARG181	M	GLN743	B	31.9	IQ742-743AA	III
ARG197	M	GLN755	A	41.2	Q755A	II
GLN739	A	GLU178	M	20.6	QH739-740AA	III
GLN177	M	SER75	G	40.5		
GLN177	M	ALA74	G	21.9		
TRP222	G	ASP198	M	25.6		
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ILE147	G	PRO170	M	20.9		

Efecto de las mutaciones:

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial)

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

Tabla A.12: Número de contactos promedio observados por frame durante la dinámica molecular del complejo TLR4-TRAM

Tipo de Enlace	TLR4-TRAM (Nuñez)		TLR4-TRAM (Propuesto)	
	Promedio	Desv. Estándar	Promedio	Desv. Estándar
Puentes de Hidrógeno	15	3	7	2
Cation- $\pi$	11	4	2	1
Hidrofóbico	4	2	0	0
Iónicos	0	0	0	0
Aromáticos	1	0	0	0



Tabla A.13: Contactos encontrados vía MD en la zona de interacción TLR4-TRAM según modelo de acoplamiento. A: *TLR4* cadena A, B: *TLR4* cadena B, N: *TRAM* modelo Nuñez, G: *TRAM* modelo propuesto

Contactos TLR4-TRAM						
Puentes de Hidrógeno						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación [20]
LYS107	G	GLN739	A	31.6	QH739-740AA	III
GLN123	G	GLU725	A	36.9	E725A	III
ASN159	G	GLU671	B	30.9		
ASN159	G	SER820	B	46.5		
ASN822	B	ASN159	G	28.6		
ASN822	B	SER156	G	21.9		
SER682	A	ALA113	G	43.2	YSS680-682AAA	I
GLN683	A	GLU114	G	37.2	Q683A	II
CYS117	N	GLN683	B	23.6	Q683A	II
SER156	N	SER730	A	26.2		
ARG163	N	TRP821	A	22.6	W821STOP	II
GLN743	B	GLU74	N	20.6	Q742-743AA	III
ARG780	B	GLU74	N	33.2		
ARG780	B	GLU72	N	21.6		
ARG780	B	PHE69	N	38.2		
GLN755	A	GLU74	N	28.6	Q755A	II
LEU760	A	SER160	N	29.2	FL759-769AA	II
ARG763	A	PHE233	N	32.6	R763A	II
Contactos Hidrofóbicos						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ALA764	A	VAL161	N	42.1		
LEU760	A	ILE234	N	46.7	FL759-760AA	II
TRP821	A	VAL161	N	24.5	W821STOP	II
Contactos Cation- $\pi$						
Residuo	Cadena	Residuo	Cadena	% Existencia	Mutación	Efecto Mutación
ARG119	G	HIS724	B	28.8		
HIS740	A	PHE78	G	33.4	QH739-740AA	I
ARG163	N	TRP821	A	42.4	W821STOP	II
ARG669	A	TYR154	N	35.1		

Efecto de las mutaciones:

I: Mutaciones que generan receptores no funcionales

II: Mutaciones que generan problemas en la unión con adaptadores

III: Mutaciones que interrumpen interacciones favorables y provocan pérdida de estabilidad

IV: Reducción de la señal (NS: pérdida total, RS: pérdida parcial)

La numeración de residuos presentada en la tabla corresponde a la que muestra el Anexo A.1

## A.9. Cálculo de hidrofobicidad superficial promedio de TLR4, adaptadores y complejos

Tabla A.14: Hidrofobicidad superficial promedio según estructura para diversas escalas de hidrofobicidad

Estructura	Hidrofobicidad Superficial Promedio		
	Kyle-Dolittle	Cowan	Berggren
TLR4	-1.748	0.466	0.246
MAL	-0.838	0.582	0.276
TRAM	-1.266	0.531	0.243
MyD88	-1.611	0.464	0.213
TLR4-MAL	-1.098	0.534	0.269
TLR4-TRAM	-0.972	0.554	0.272

## A.10. Detalles de la interacción TLR4-TLR4

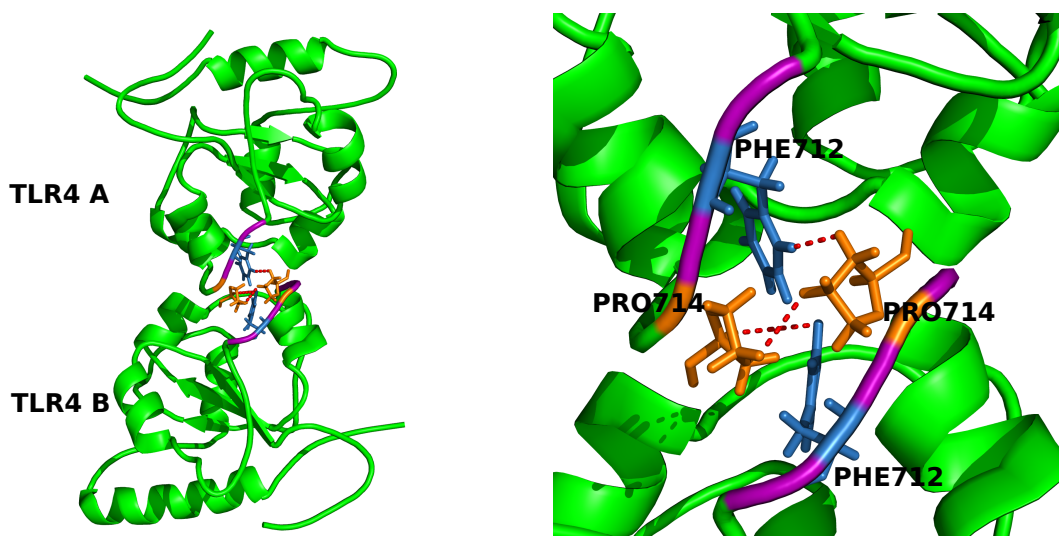


Figura A.8: Interacciones hidrofóbicas en el complejo TLR4-TLR4. La figura de la izquierda muestra la conformación adoptada por el homodímero cuando se producen las interacciones PHE712-PRO714, PRO714-PHE712, PRO714-PRO714 entre los dominios TIR de dos monómeros de TLR4. A la derecha de la figura se muestra la zona de interacción entre los monómeros de TLR4. En particular se observan las interacciones anteriormente mencionadas. Se plantea que la interacción PRO714-PRO714 permite la correcta orientación del resto de los residuos del lazo BB facilitando la interacción de otros residuos del lazo.

## A.11. Estimación de energía libre en función de contactos observados

Tabla A.15: Estimación de energía libre para complejo TLR4-MAL de acuerdo a los contactos presentados

Tipo de Enlace	TLR4-MAL(N)	TLR4-MAL(P)	Energía Enlace	Referencia
Puentes de Hidrogeno	11	4	-3.6	Pace <i>et al.</i> (2009)
Cation Pi	3	1	-8.0	Martis <i>et al.</i> (2008)
Hidrofobico	5	10	-8.8	Pace <i>et al.</i> (2011)
Ionicos	1	0	-5.0	Gurry <i>et al.</i> (2010)
Aromaticos	1	0	-9.0	Hunter <i>et al.</i> (1991)
Energía Estimada [kcal/mol]	-121,6	-110,4		

La energía de enlace se calcula como:  $\Delta(\Delta G) = \Delta G_{Mutante} - \Delta G_{Nativo}$

Valores de energía en [kcal/mol]

Tabla A.16: Energía de unión según tipo de interacción para complejo TLR4-TRAM de acuerdo a los contactos presentados

Tipo de Enlace	TLR4-TRAM(N)	TLR4-TRAM(P)	Energía Enlace	Referencia
Puentes de Hidrogeno	15	7	-3,6	Pace <i>et al.</i> (2009)
Cation Pi	11	2	-8	Martis <i>et al.</i> (2008)
Hidrofobico	4	0	-8,8	Pace <i>et al.</i> (2011)
Ionicos	0	0	-5	Gurry <i>et al.</i> (2010)
Aromaticos	1	0	-9	Hunter <i>et al.</i> (1991)
Energía Estimada [kcal/mol]	-186,2	-41,2		

La energía de enlace se calcula como:  $\Delta(\Delta G) = \Delta G_{Mutante} - \Delta G_{Nativo}$

Valores de energía en [kcal/mol]

## A.12. Representación de las superficies electrostáticas en los complejos TLR4-MAL y TLR4-TRAM.

Figura A.9: Representación de las superficies electrostáticas de TLR4-MAL y TLR4-TRAM. Para cada uno de los complejos TLR4-MAL y TLR4-TRAM se muestra una vista frontal a la zona de interacción de cada una de los dominios TIR participantes en la interacción.

