

UNIVERSIDAD DE CONCEPCION
FACULTAD DE FARMACIA



Profesor Patrocinante y Guía Dr(a). Marta Bunster
Laboratorio Biofísica Molecular
Departamento Bioquímica y Biología Molecular
Facultad Ciencias Biológicas
Universidad de Concepción

Modelo de conducción de luz en una varilla del ficobilisoma de *Gracilaria chilensis*

Tesis para optar al título profesional de Bioquímico

Maximiliano Francisco Figueroa Yévenes

Primer Semestre de 2005

INDICE GENERAL

ÍNDICE DE FIGURAS	iv
ÍNDICE DE TABLAS	viii
RESUMEN	1
SUMMARY	3
1.- INTRODUCCIÓN	5
2.- HIPÓTESIS Y OBJETIVOS	15
3.- METODOLOGÍA	16
Antecedentes	16
3.1 Estandarización de una metodología de docking proteico para ficobiliproteínas	18
3.1.1 Estructura de PC y PE	18
3.1.2 Selección del programa de docking	19
3.1.3 Protocolo de docking	20
3.1.3.a Restricción del espacio de búsqueda	20
3.1.3.b Evaluación de resultados	22
3.1.4 Construcción de un primer modelo de varilla	23
3.2 Estandarización de un protocolo de dinámica molecular para optimizar resultados de docking proteico	24
3.2.1 Generación de topologías	26
3.2.2 Evaluación de campos de fuerza para reconstitución de complejos creados mediante docking de proteínas	31
3.2.3 Dinámica molecular del modelo de varilla rígida	35
3.3 Postulación de vías preferenciales de conducción de luz en el modelo generado	36

4.- RESULTADOS	40
4.1 Estandarización de una metodología de docking proteico para ficobiliproteínas	40
4.1.1 Elección del software de docking	40
4.1.2 Docking de ficobiliproteínas	41
4.1.2.1 Docking ficocianina – ficocianina	43
4.1.2.2 Docking ficocianina – ficoeritrina	44
4.1.2.3 Docking ficoeritrina – ficoeritrina	45
4.2 Estandarización de un protocolo de dinámica molecular para optimizar resultados de docking proteico	49
4.2.1 Creación de topologías	49
4.2.2 Evaluación de campos de fuerza para reconstitución de complejos creados mediante docking de proteínas	50
4.2.3 Dinámica molecular del modelo de varilla de rígida	55
4.3 Vías preferenciales de conducción de luz	59
4.3.1 Vías preferenciales de luz intra ficoeritrina	59
4.3.2 Vías preferenciales de luz inter ficoeritrinas	62
4.3.2.a Vías inter ficoeritrinas en el modelo rígido	62
4.3.2.b Vías inter ficoeritrinas en el modelo post dinámica molecular	65
4.3.3 Vías preferenciales de luz inter ficoeritrina y ficocianina	70
4.3.3.a Vías inter ficoeritrina y ficocianina en el modelo rígido	70
4.3.3.b Vías inter ficoeritrina y ficocianina en el modelo post dinámica molecular.	72
4.3.4 Vías preferenciales de luz intra ficocianina	75
4.3.5 vías preferenciales de conducción de luz inter ficocianina	76
5.- DISCUSIÓN	79
6.- CONCLUSIONES	94

7.- APÉNDICES	95
Apéndice 1: Instrucciones de usa para ZDOCK	95
Apéndice 2: Instrucciones de uso para GROMACS	99
Apéndice 3: Edición de archivos para crear topologías de compuestos nuevos	106
Apéndice 4: Cálculo de constantes de transferencia de luz de acuerdo a la Ley de Förster	117
8.- BIBLIOGRAFÍA	123

INDICE DE FIGURAS

Figura 1	7
Esquema de un ficobilisoma de cianobacteria <i>Synechocystis sp.</i> 6701 adaptado de Anderson et al, 1998	
Figura 2	8
Estructura de ficobiliproteínas	
Figura 3	17
Diagrama de flujo del proyecto	
Figura 4	18
Diagrama de flujo de estandarización de metodología de docking de Ficobiliproteínas	
Figura 5	19
Ficobiliproteínas de <i>Gracilaria chilensis</i> componentes de la porción de varillas	
Figura 6	21
Micrografía electrónica de varillas de ficobilisomas	
Figura 7	25
Diagrama de flujo que muestra el trabajo a realizar para llevar a cabo la optimización del modelo rígido por medio de una dinámica molecular	
Figura 8	27
Esquema del cromóforo ficocianobilina	
Figura 9	28
Esquema de un enlace junto a su gráfico de energía potencial	
Figura 10	29
Esquema de un ángulo formado por 3 átomos enlazados junto a su gráfico de energía potencial	
Figura 11	30
Esquema de las 3 formas en las cuales podemos encontrar un ángulo dihedral impropio, junto a su gráfico de energía potencial	
Figura 12	31
Esquema de un ángulo dihedral propio junto a su gráfico de energía potencial	
Figura 13	33
Esquema del docking dirigido	

Figura 14	42
Script en bash “Calculator2.0”	
Figura 15	47
Complejos seleccionados en cada docking	
Figura 16	48
Modelo de la varilla construido a partir del alineamiento estructural de los complejos obtenidos en los 3 "docking" individuales	
Figura 17	51
Complejo 7, con un rmsd de 5.19 Å	
Figura 18	52
Gráfico del rmsd del esqueleto polipeptídico del heterodímero $\alpha\beta$ de ficocianina de <i>Gracilaria chilensis</i> para los 5 campos de fuerza	
Figura 19	53
Gráfico de energía potencial del sistema a lo largo de los 400 ps de dinámica molecular para el heterodímero $\alpha\beta$ de ficocianina de <i>Gracilaria chilensis</i> en el campo de fuerza ffG43b1 (Gromos96 para vacío)	
Figura 20	56
Gráfico del rmsd del esqueleto polipeptídico y la temperatura de éste a lo largo de la dinámica molecular del modelo de varilla de ficobilisoma de <i>Gracilaria chilensis</i>	
Figura 21	58
Script en bash “Cromoforos”	
Figura 22	59
Superposición de los cromóforos antes (color rojo) y después de la dinámica molecular (color azul)	
Figura 23	61
A: Esquema de una ficobiliproteína donde se denota esquemáticamente la ubicación de cada sub unidad con su nombre en nomenclatura tradicional (α y β) y con su respectivo nombre usado en el archivo de coordenadas PDB. B: Esquema de una de las tres vías preferenciales de conducción de luz intra ficoeritrina	
Figura 24	63
Vía preferencial de conducción de luz inter ficoeritrina para el primer tercio de la ficobiliproteína	
Figura 25	64
Vía preferencial de conducción de luz inter ficoeritrina para el segundo tercio de la ficobiliproteína	

Figura 26	65
Vía preferencial de conducción de luz inter ficoeritrina para el tercer tercio de la ficobiliproteína	
Figura 27	67
Vía preferencial de conducción de luz inter ficoeritrina para el primer tercio de la ficobiliproteína	
Figura 28	68
Vía preferencial de conducción de luz inter ficoeritrina para el segundo tercio de la ficobiliproteína	
Figura 29	69
Vía preferencial de conducción de luz inter ficoeritrina para el tercer tercio de la ficobiliproteína	
Figura 30	71
Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el primer tercio de ficoeritrina	
Figura 31	71
Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el segundo tercio de ficoeritrina	
Figura 32	72
Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el tercer tercio de ficoeritrina	
Figura 33	74
Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para un tercio de ficoeritrina	
Figura 34	74
Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el segundo tercio de ficoeritrina	
Figura 35	76
Esquema de una vía preferencial de conducción de luz intra ficocianina. Adaptado de Almonacid, 2004	
Figura 36	78
Esquema de las vías preferenciales de conducción de luz inter ficocianina. Adaptado de Almonacid, 2004	
Figura 37	105
Archivo “mdp” para realizar dinámica molecular con protocolo de simulated annealing usado en la varilla rígida de esta tesis	
Figura 38	108
Formato de un compuesto en el archivo rtp	

Figura 39	110
Formato de un compuesto en el archivo hdb	
Figura 40	112
Formato del archivo ffgmxbon.itp	
Figura 41	114
Formato del archivo ffgmxnb.itp	
Figura 42	118
Encabezado de un archivo .gjf	
Figura 43	119
Encabezado correcto en un archivo .gjf para realizar cálculos SP (single point)	
Figura 44	120
Ejemplo de un archivo .bat para ejecución seriada de archivos .gjf	
Figura 45	120
Líneas que describen a un cromóforo en el archivo nitrógenos-dipolo.txt	

INDICE DE TABLAS

Tabla 1	35
Temperaturas usadas (K) durante la dinámica molecular para cada uno de los componentes del sistema	
Tabla 2	43
Resultados de docking ficocianina – ficocianina y análisis de los aminoácidos presentes en la interfase del complejo generado	
Tabla 3	44
Resultados de docking ficocianina – ficoeritrina y análisis de los aminoácidos presentes en la interfase del complejo generado	
Tabla 4	45
Resultados de docking ficoeritrina – ficoeritrina y análisis de los aminoácidos presentes en la interfase del complejo generado	
Tabla 5	46
Caracterización de las interfases de los complejos seleccionados en cada docking	
Tabla 6	50
Complejos seleccionados en el docking enlazado	
Tabla 7	54
Δ energía potencial al término de la dinámica molecular	
Tabla 8	54
Evaluación estructural de los complejos luego de finalizada la dinámica molecular de 400 ps	
Tabla 9	55
Análisis de las interfases acabada la dinámica molecular de 400 ps	
Tabla 10	57
Análisis del modelo de antena de acuerdo a datos obtenidos a partir del Gráfico de Ramachandran	
Tabla 11	57
Análisis de la interfases del modelo final de interacción entre ficobiliproteínas para conformar la varilla mínima del ficobilisoma de <i>Gracilaria chilensis</i>	
Tabla 12	60
Se presentan los valores para cada par dador aceptor que participan en las vías intra ficoeritrina	

Tabla 13	62
Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina en el modelo rígido de varilla	
Tabla 14	66
Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina en el modelo una vez acabada la dinámica molecular	
Tabla 15	70
Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina y ficocianina	
Tabla 16	73
Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina y ficocianina	
Tabla 17	75
Vías de conducción de luz intra ficocianina Adaptado de Almonacid, 2004	
Tabla 18	77
Vías de conducción de luz inter ficocianina Adaptado de Almonacid, 2004	

RESUMEN

Los ficobilisomas son complejos multiproteicos presentes en algas rojas y cianobacterias, que cumplen una función accesoria en la captación y conducción de luz hacia los fotosistemas. Los ficobilisomas de *Gracilaria chilensis* están formados por ficobiliproteínas organizadas en un centro que contiene Aloficocianina(APC) y varillas que irradian desde él, formadas por Ficocianina(PC) y Ficoeritrina(PE).

Los ficobilisomas constituyen el ejemplo más eficiente(~100%), de un sistema de conducción de luz en el rango de 450-650nm, descrito hasta ahora en la naturaleza.

Este trabajo describe la obtención de un modelo estructural de una varilla mínima formada por dos unidades de Ficocianina y dos unidades de Ficoeritrina y la propuesta de vías preferenciales de conducción de luz a través de este complejo.

El procedimiento de obtención del modelo utilizó las estructuras tridimensionales de ambas proteínas; el uso del programa ZDOCK permitió obtener un modelo de encaje rígido entre PC-PC, PC-PE, y PE-PE. Este modelo se sometió a etapas de evaluación y refinamiento. El refinamiento se realizó mediante Dinámica Molecular utilizando GROMACS, después de un cuidadoso análisis de los campos de fuerza disponibles.

La transferencia de energía entre los grupos cromóforos responsables de las propiedades espectroscópicas de las ficobiliproteínas, se realiza por fluorescencia en resonancia, lo que permitió utilizar la Ecuación de Forster para determinar las velocidades de transferencia entre pares de grupos cromóforos. Los valores de las velocidades permitieron postular vías preferenciales de conducción en el modelo las que

concuerdan con los valores de velocidades de transferencia obtenidos experimentalmente disponibles en la literatura.

El análisis realizado en este trabajo permitió construir un modelo estructural-funcional de una varilla de ficobilisoma de *Gracilaria chilensis* que permite plantear una discusión que permita explicar la alta eficiencia de estos sistemas biológicos.

SUMMARY

Phycobilisomes are multiproteic complexes present in red algae and cyanobacteria; they perform an accessory function responsible of the light harvesting and conduction of light to Photosystems.

Phycobilisomes from *Gracilaria chilensis* are formed, as in other algae, by phycobiliproteins; they are organized in a core formed by Allophycocyanin(APC), and radially disposed rods formed by Phycocyanina(PC) and Phycoerythrin(PE).

This special arrangement of proteins constitutes the most efficient conductor of light (~100%) described in the literature, to this date. This work describes the building of an structural model of a minimum rod formed by two hexamers of PE and two hexamers of PC and proposes preferential pathways for light conduction able to explain the high efficiency of the system.

The methodology to obtain the model involved the three-dimensional structures of both phycobiliproteins and the use of ZDOCK to obtain a rigid docking model of the complexes PC-PC, PC-PE, PE-PE. The model was evaluated and refined. The refinement was performed by Molecular Dynamics after a careful evaluation of the force fields available in GROMACS.

The Forster theory was used to calculate the transfer rate between pairs of chromophores; this is possible because the energy transfer among chromophores occurs by fluorescence in resonance.

Preferential pathways for light conduction were proposed based on the values of the Transfer constants obtained for each pair of chromophores in the rod. The values obtained for the rate of conduction agree with the experimental data reported in the literature.

A structural- functional model of a minimum rod of phycobilisomes of *Gracilaria chilensis* was built. The pathways for the light conduction described in this work, introduces a discussion to explain the high efficiency of these biological systems.-

1.- INTRODUCCIÓN

Los organismos fotosintéticos necesitan de elementos inorgánicos, agua y dióxido de carbono, para sintetizar moléculas orgánicas y así obtener la energía necesaria para las actividades celulares. Este proceso de conversión de materia inorgánica a orgánica requiere energía luminosa, y se denomina fotosíntesis. Esta energía luminosa es obtenida del sol y es captada principalmente por pigmentos especializados, como la clorofila en plantas superiores, algas y cianobacterias (Glazer, 1989). En organismos marinos fotosintéticos como algas y cianobacterias, en cuyo ambiente acuático disminuye la cantidad y calidad de la luz accesible a estos organismos, se ha debido desarrollar una estrategia de captación de luz más eficiente, tal como sistemas accesorios de captación de luz en el rango de longitudes de onda en los cuales la clorofila no es eficaz. (Samsonoff y colaboradores, 2001).

Los ficobilisomas son los principales complejos multiproteicos accesorios para la captación y conducción de luz hacia los centros fotosintéticos de algas rojas y cianobacterias (Ducret y colaboradores, 1994). Estos complejos están encargados de captar la luz en el rango visible en el cual la clorofila no lo hace (550 – 660 nm aproximadamente) y conducir esta energía luminosa hacia los centros fotosintéticos con una altísima eficiencia cuántica (cercana al 100%) (Glazer, 1989). Esta propiedad hace que los ficobilisomas sean las estructuras más interesantes de estudiar para entender las bases moleculares de esta alta eficiencia en la naturaleza.

Los ficobilisomas se encuentran anclados a la membrana tilacoidal, a diferencia de la clorofila que se encuentra inserta en la membrana de los tilacoides (Samsonoff y

colaboradores, 2001). Los ficobilisomas están constituídos principalmente por ficobiliproteínas (PBP), y en menor escala por proteínas de unión o “linkers”, que permiten el correcto ensamble del ficobilisoma (Ducret y colaboradores, 1994). Estos ficobilisomas poseen una conformación espacial común, en la cual se presenta un núcleo formado por la ficobiliproteína aloficocianina (APC), del cual irradian varillas compuestas por ficocianina, (principalmente en cianobacterias), y ficoeritrina o ficoeritrocianina (generalmente en algas rojas) (Figura 1) (Glazer, 1989). Las PBP que componen los ficobilisomas presentan una unidad mínima compuesta por un heterodímero de cadenas $\alpha\beta$, e *in vivo* se presentan a la forma de hexámeros $(\alpha\beta)_6$ (ficoeritrina, ficoeritrocianina y ficocianina) o de trímeros $(\alpha\beta)_3$ (aloficocianina) (Figura 2). Las subunidades α y β presentan un plegamiento similar y corresponde a la clase estructural $\alpha\alpha$ de la familia de las globulinas. Cada cadena además lleva bilinas derivadas de tetrapirroles lineales, covalentemente unidas a grupos de cisteínas, las cuales confieren las características espectroscópicas y por lo tanto están encargados de captar y conducir la luz en los ficobilisomas hacia los centros fotosintéticos (Glazer, 1989). El número de bilinas (o grupos cromóforos) unidas a cada cadena va a depender de la naturaleza de la ficobiliproteína, observándose así una disminución paulatina desde las ficobiliproteínas ubicadas en la periferia (entre 34 y 15 cromóforos por hexámero) hacia las ubicadas en el núcleo (6 cromóforos por trímero) (Glazer, 1989). Sin embargo, el tipo de cromóforo no determina el tipo de ficobiliproteína, observándose que, tanto ficocianina como aloficianina poseen el mismo grupo cromóforo: ficocianobilina. Las propiedades espectroscópicas están dadas entonces por la geometría y el ambiente que le brinde la ficobiliproteína (Glazer, 1989; Samsonoff, 2001).

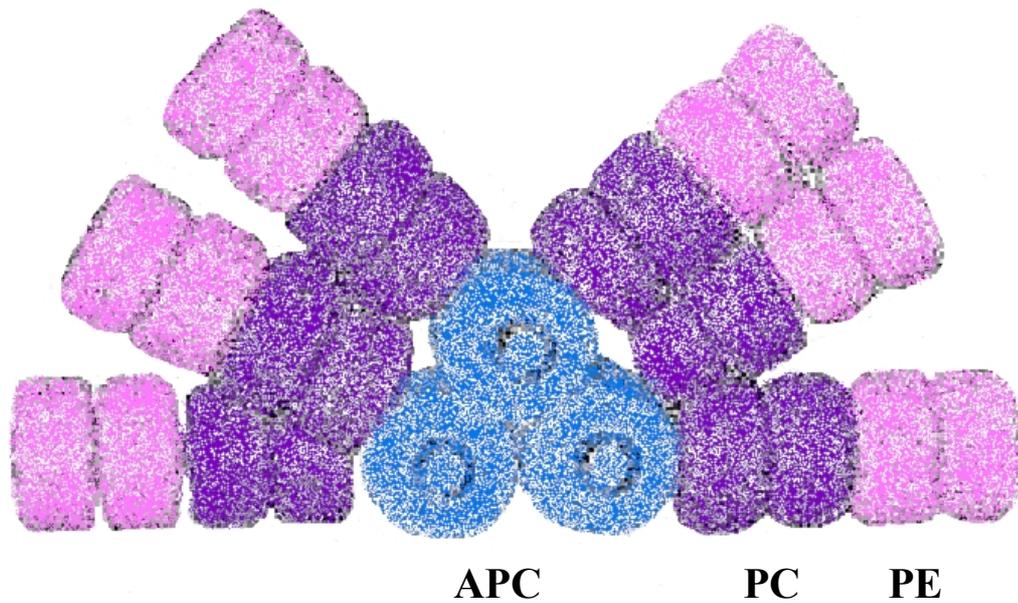


Figura 1. Esquema de un ficobilisoma de cianobacteria *Synechocystis sp.* 6701 adaptado de Anderson y colaboradores, 1998. Se observa un núcleo de aloficocianina (APC) del que irradian varillas compuestas en su parte proximal por ficocianina (PC) y en su parte distal por ficoeritrina (PE)

Los estudios hechos hasta la fecha están enfocados en su mayoría en algas rojas unicelulares y cianobacterias (Glazer, 1989; Ducret, 1996; Wang, 2001; Zolla, 2002), debido a su fácil manipulación en el laboratorio, pero son escasos los trabajos en algas pluricelulares (Jiang y colaboradores, 2001). El laboratorio de Biofísica Molecular de la Universidad de Concepción ha abordado el estudio de los ficobilisomas de algas pluricelulares y el espécimen usado como modelo de estudio es *Gracilaria chilensis* (Byrd, Mackachlan y Oliveira), conocida comúnmente como pelillo, y que podemos encontrar a lo largo de las costas chilenas y peruanas.

El alga roja eucariota *Gracilaria chilensis* posee ficobilisomas compuestos por un núcleo central constituido por la ficobiliproteína aloficocianina (APC), y sus varillas

están compuestas ficocianina (PC) en su parte proximal, y ficoeritrina (PE) en su parte más distal (Bunster y colaboradores, 1997 y 2000) (Figura 1).

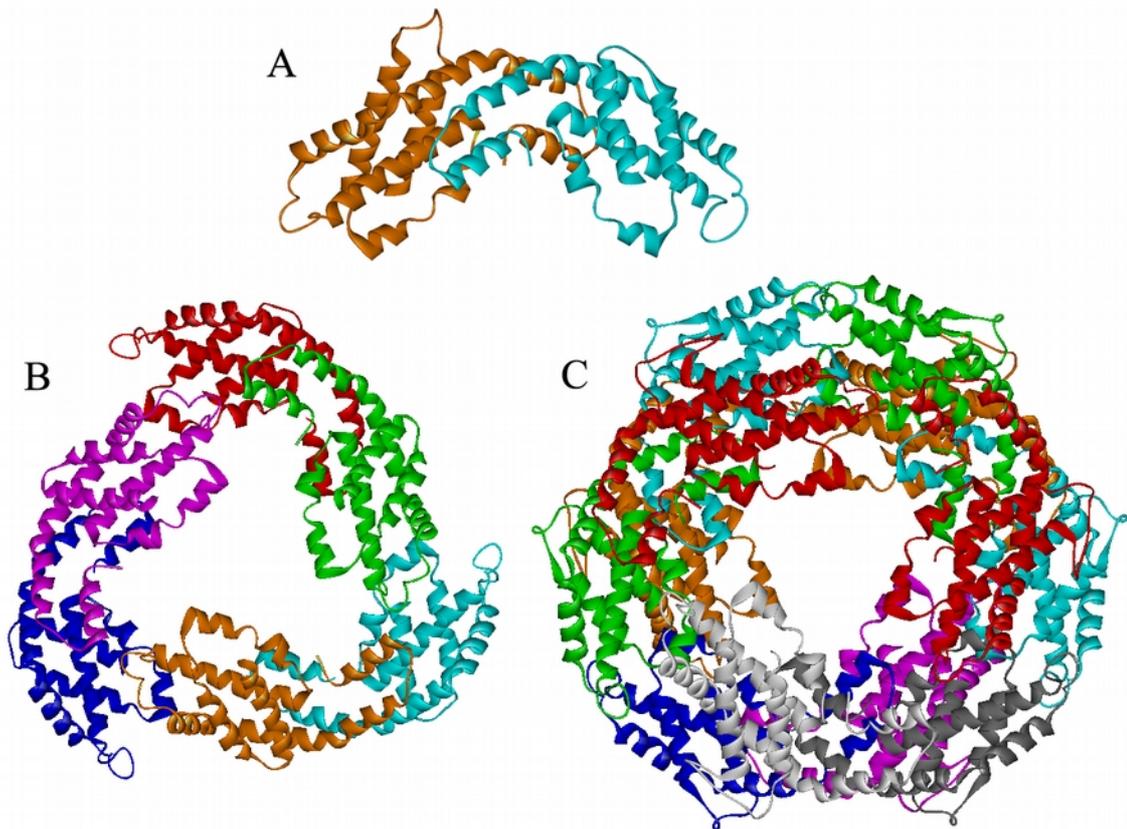


Figura 2. Estructura de ficobiliproteínas. A) Estructura del heterodímero $\alpha\beta$ de ficocianina de *Gracilaria chilensis* (en proceso de depósito en PDB). B) Estructura de un trímero, correspondiente a aloficocianina de *Porphyra yezoensis* (código PDB: 1KN1). C) Estructura de un hexámero, correspondiente a ficoeritrina de *Gracilaria chilensis* (código PDB: 1EYX). Las subunidades de cada una han sido coloreadas de forma diferencial.

En el laboratorio de Biofísica Molecular se ha logrado determinar la estructura de 2 de las 3 ficobiliproteínas presentes en *Gracilaria chilensis*: ficoeritrina (Contreras-Martel y colaboradores, 2001) y ficocianina (Bunster y colaboradores, en preparación), ambas mediante la técnica de difracción de rayos X. A su vez, otros autores han obtenido las estructuras de ficobiliproteínas de otras especies, y han formulado modelos de interacción entre ellas (Schimer y colaboradores, 1986; Stec y colaboradores, 1999;

Wang y colaboradores, 2001), sin embargo esos modelos son especulativos pues no generan un modelo estructural tangible, debido a que no se realiza un estudio estructural detallado de interacción entre las proteínas. Se conoce la estructura tridimensional (3D) de las ficobiliproteínas que forman la varilla y además se conoce a través de ensayos biológicos el tipo de asociación (microscopía electrónica) entre las ficobiliproteínas. Basados en esa información es posible realizar estudios bioinformáticos para predecir la asociación entre éstas 2 proteínas, mediante la técnica de ““docking” proteico”, que permite predecir de manera efectiva la interacción espacial de 2 proteínas como ya se ha visto en un número creciente de publicaciones (Elcock y colaboradores, 2001; Chen y colaboradores, 2002; Halperin y colaboradores, 2002; Janin y colaboradores, 2003). Además esta técnica hace que sea un requisito el conocimiento de la estructura tridimensional de las proteínas involucradas.

Se han descrito varios programas de “docking” proteico (Janin y colaboradores, 2003), todos ellos fijan una de las proteínas en el espacio, llamada proteína receptora, y la segunda proteína, proteína ligando, es rotada y trasladada alrededor de ella para lograr el ajuste. Este proceso se realiza en 6 dimensiones, 3 dimensiones de traslación y 3 dimensiones de rotación de la proteína ligando (Chen y colaboradores, 2003). El número de soluciones generadas pueden ser infinitas por lo tanto los programas de ““docking”” realizan una primera aproximación limitando los movimientos de rotación y traslación (Janin y colaboradores, 2003). Los programas de “docking” han evolucionado desde sus inicios para disminuir el costo computacional y el tiempo de cálculo, utilizando para esto, por ejemplo, Transformada rápida de Fourier (Camacho y colaboradores, 2003) y reduciendo los grados de libertad de cada aminoácido, considerando entonces la proteína como un cuerpo rígido.

Uno de los aspectos más importantes en la utilización de estos métodos reside en el sistema de evaluación de los modelos generados, es decir en el algoritmo utilizado para evaluar cada modelo y el peso asignado a cada parámetro para entregar finalmente una lista con los modelos de interacción ordenados de acuerdo a puntaje. Con respecto a esto es entonces donde reside la gran diferencia entre los distintos programas de “docking” proteico.

Los programas de “docking” más simples utilizan solo un criterio geométrico (Ben-Zeev y colaboradores, 2003) para evaluar las soluciones considerando sólo la complementariedad geométrica, pues este parámetro es muy importante para complejos proteína-proteína estables en el tiempo (por ejemplo, subunidades de una proteína con estructura cuaternaria). Sin embargo este criterio no es muy adecuado para proteínas que no poseen estructura cuaternaria, o de interacción transiente. Nuevas funciones de evaluación involucran, además de la complementariedad geométrica, la complementariedad electrostática y la hidrofóbica (Chen y colaboradores, 2003; Camacho y colaboradores, 2002; Camacho y colaboradores, 2003), los cuales permiten simular más eficientemente los factores involucrados en la interacción proteína-proteína. Además de los parámetros ya mencionados, es necesario considerar la desolvatación (Baker y colaboradores 2003; Camacho y colaboradores, 2000; Camacho y colaboradores 2003; Chen y colaboradores, 2003, Mendez y colaboradores 2003), pues para que dos proteínas interactúen entre sí, las superficies de las interfases deben desolvatarsen. Para simular este evento, los programas de “docking” han afrontado el problema desde un punto de vista estadístico que consiste en hacer mediciones de energía libre de desolvatación en un conjunto de complejos de estructura conocida, ya

sea por métodos semiempíricos o *ab initio*, utilizando mecánica cuántica. En general este método se llama ACE (Atomic Contact Energy) y consiste en calcular la diferencia de ΔG de formación de la interacción entre el contacto átomo proteína – átomo agua, y el ΔG de formación de la interacción átomo proteína – átomo proteína (Zhang y colaboradores, 1997). Este estudio entrega un valor promedio de ACE para cada átomo de los 20 aminoácidos; desde donde es posible predecir qué parte de la superficie de la proteína es más fácil de desolvatar, y por lo tanto qué porción de la superficie tiene más posibilidades de interactuar con otra proteína.

Una metodología nueva e innovadora consiste en hacer estudios de desolvatación de modo más exhaustivo usando solvente explícito en las simulaciones. Aunque en este caso el gasto computacional es mayor, puede entregar resultados más asertivos que los por la metodología de ACE (Gonzalez-Nilo, en preparación).

Uno de los problemas más grandes de la metodología de “docking” es la cantidad de información generada. Resultados experimentales muchas veces permiten acotar el área de búsqueda e incorporar datos en los parámetros utilizados para el cálculo (por ejemplo, la importancia de ciertos aminoácidos para que la dimerización ocurra). La forma más evidente es forzar a que ciertos aminoácidos (definidos por ensayos experimentales) estén en la interfase, o hacer que nunca lo estén: ambas situaciones simultáneamente son usadas habitualmente para definir un área de búsqueda más acotada (Chen y colaboradores, 2003). Así, mezclando los datos experimentales juntos a los de probabilidad de desolvatación, es posible acotar el área de búsqueda, disminuyendo el tiempo de cálculo, y hacer un poco más exhaustivo la búsqueda.

Debido a que las proteínas son consideradas como cuerpos rígidos, una vez realizado el “docking”, se debe ejecutar una minimización energética del sistema, que nos entregue la conformación correspondiente al mínimo energético global. Para conseguir esto, existen varios protocolos y metodologías, las cuales van desde simples minimizaciones energéticas por “Steepest Descent” hasta complejas metodologías *ab initio* de mecánica cuántica (Van der Spoel, 2004). Sin embargo, el costo computacional es tan alto en este último caso que es virtualmente imposible utilizar mecánica cuántica para simular eventos en una proteína completa. Para solucionar este problema se utilizan aproximaciones que involucran simplificaciones de la metodología cuántica, las que consisten en reemplazar esta metodología por su antecesora: la mecánica clásica, que en este nuevo contexto, se denomina mecánica molecular (Steinbach, 2003).

Una de las grandes diferencias entre mecánica clásica y mecánica cuántica, es que esta última es probabilística y, como su nombre lo indica, cuantizada, es decir, tanto la energía como la posición de una partícula está definida en cuantos discretos, a diferencia de la mecánica clásica, en que todo es un continuo, por lo cual es fácil seguir la trayectoria de un evento, y conocer la energía que posee la partícula en cada uno de sus estados del movimiento. La mecánica cuántica, en cambio, según el principio de incertidumbre de Heisenberg, no permite conocer la posición y el momento de una partícula con precisión infinita en un mismo tiempo (Steinbach, 2003). Sin embargo, a temperaturas iguales o mayores que la temperatura ambiente (300 K) los cuantos discretos se acercan tanto unos a otros, que es posible considerarlos como continuos, y ésta es la principal razón por la cual es posible utilizar mecánica clásica en las simulaciones, pues a este nivel, un sistema analizado de forma clásica y cuántica se observaría, prácticamente, de igual forma. La metodología con mecánica cuántica será

sin duda la elección en sistemas con pequeño número de átomos, y con certeza en el futuro será también la elección para sistemas más complejos, siempre y cuando la tecnología de los procesadores de los computadores siga creciendo al ritmo con el cual se ha desarrollado en los últimos años. En la actualidad, es posible realizar estudios mezclando ambas metodologías, para lo cual, sólo el sitio de interés de la proteína, por ejemplo un sitio activo de una enzima, se trata con mecánica cuántica, y el resto de la proteína con mecánica clásica (Steinbach, 2003; Van der Spoel y colaboradores, 2004).

Con todo lo descrito hasta este punto, puede decirse que es posible conocer la manera estructural mediante la cual interaccionan dos proteínas, sin necesidad de tener la resolución estructural del complejo dada las estructuras de cada una de las proteínas que componen el sistema: la bioinformática, junto a datos experimentales, proporciona las herramientas para solucionar este problema. Entonces, conocidas las estructuras de las ficobiliproteínas ficocianina y ficoeritrina de *Gracilaria chilensis*, es posible formular un modelo de interacción entre ellas en una varilla del ficobilisoma, y así encontrar una explicación a nivel estructural, de la alta eficiencia cuántica observada en la conducción de luz a través de ella. El problema no es menor, pues el modelo consiste en lo que hemos definido como “varilla mínima”, compuesta por 2 ficoeritrinas y 2 ficocianinas, lo que involucra los 3 tipos de interacciones presentes en una varilla del ficobilisoma de *Gracilaria chilensis*: interacción ficocianina – ficocianina; ficocianina – ficoeritrina; y ficoeritrina – ficoeritrina. Este modelo de 4 proteínas debe ser afrontado en realidad como la interacción de 48 proteínas, pues este es el número de subunidades en total que participan. He aquí la gran diferencia con respecto a los modelos de interacción antes propuestos, que sólo consideraban interacciones de 2 ficobiliproteínas como cuerpos rígidos y entes independientes. Sin embargo en esta tesis, al modelo

rígido le sigue un proceso de minimización energética en el cual estamos considerando el efecto que puede generar una proteína sobre otra al estar interaccionando con lo cual se logra entonces obtener el modelo de interacción de ficobiliproteínas más completo, desde el punto de vista metodológico, que aquellos formulados hasta la fecha.

El modelo estructural propuesto debe dar cuenta de las observaciones experimentales desde el punto de vista, es decir, su alta eficiencia así como el tiempo que tarda la luz en circular por las ficobiliproteínas. El modo mediante el cual es transmitida la luz a través de esta varilla es por el fenómeno conocido como FRET (transferencia de energía por fluorescencia en resonancia) el cual es representado por la ley de Förster que nos permite calcular entonces valores para constantes de transferencia de luz entre pares de cromóforos. En todos los trabajos previos las vías de conducción de luz para ficoeritrina fue estimada íntegramente en base a distancias y no por el cálculo de estas constantes, lo que hace de este trabajo entonces el primer modelo completo desde el punto de vista estructural así como funcional.

2.- HIPÓTESIS Y OBJETIVOS

HIPÓTESIS

Dadas las estructuras cristalográficas de los componentes de las varillas de los ficobilisomas de *Gracilaria chilensis*, es posible generar un modelo de interacción entre ellas que expliquen la alta eficiencia cuántica de conducción de luz en éste sistema.

OBJETIVO GENERAL

Generar un modelo de interacción entre los componentes de las varillas de los ficobilisomas de *Gracilaria chilensis*.

OBJETIVOS ESPECÍFICOS

- Estandarización de una metodología de "“docking”" proteico para ficobiliproteínas
- Estandarización de una metodología de dinámica molecular para optimizar resultados de “docking” proteico
- Postulación de vías preferenciales de conducción de luz en el modelo generado

3.- METODOLOGÍA

Antecedentes

La bioinformática constituye la metodología de elección para el desarrollo de esta tesis ya que permite postular un buen modelo estructural de interacción de ficobiliproteínas en una varilla del ficobilisoma de *G. chilensis*, con el objeto de proponer una explicación estructural de la alta eficiencia en la conducción de luz a lo largo de esta varilla. Para ello es necesario utilizar programas computacionales que tienen la capacidad de generar modelos de interacción de proteínas (“docking” proteico), minimizaciones energéticas de modelos de proteicos (dinámicas moleculares) y postular vías preferenciales de conducción de luz (metodología mecano cuántica semi empírica).

Los programas utilizados en este estudio se ejecutaron bajo sistema operativo Linux (Linux Mandrake 10.0), excepto para el cálculo de vías preferenciales de conducción de luz que se ejecutó bajo sistema operativo Windows XP®.

Los objetivos específicos constituyen etapas que se realizaron de manera secuencial, por lo tanto se incluye en algunas de ellas, un diagrama de flujo con el objeto de esclarecer la interrelación entre las distintas etapas.

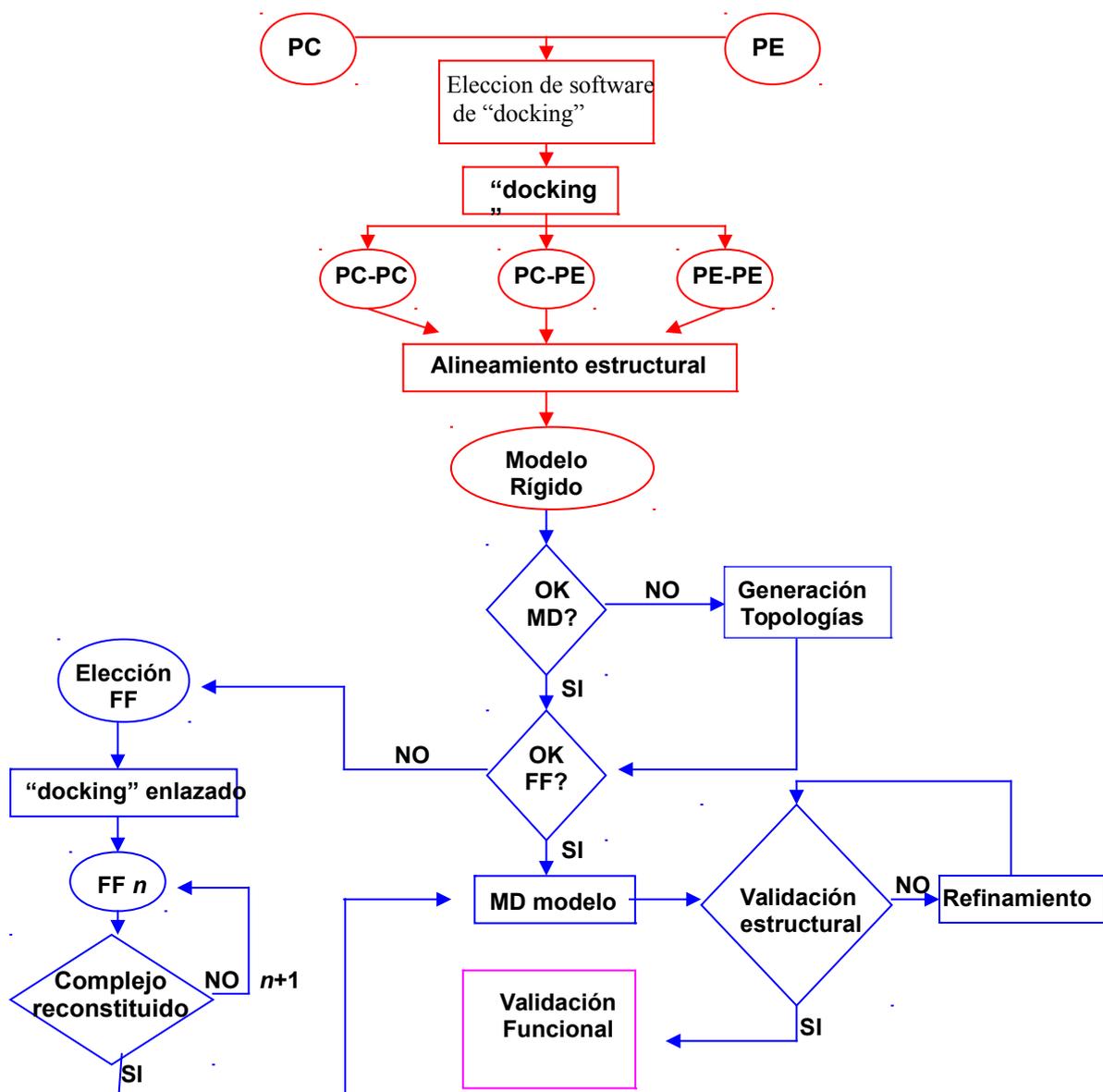


Figura 3: Diagrama de flujo del proyecto. En colores se distinguen los 3 objetivos específicos. En rojo se enmarca el primer objetivo específico “Estandarización de una metodología de “docking” proteico para ficobiliproteínas”. En azul el segundo objetivo específico “Estandarización de una metodología de dinámica molecular para optimizar resultados de “docking” proteico”, y en fucsia se observa el tercer objetivo específico “Postulación de vías preferenciales de conducción de luz en el modelo generado”. (PC: ficocianina; PE: ficoeritrina; MD: dinámica molecular; FF: campo de fuerza)

3.1 Estandarización de una metodología de “docking” proteico para ficobiliproteínas

Esta etapa está representada en la Figura 4.

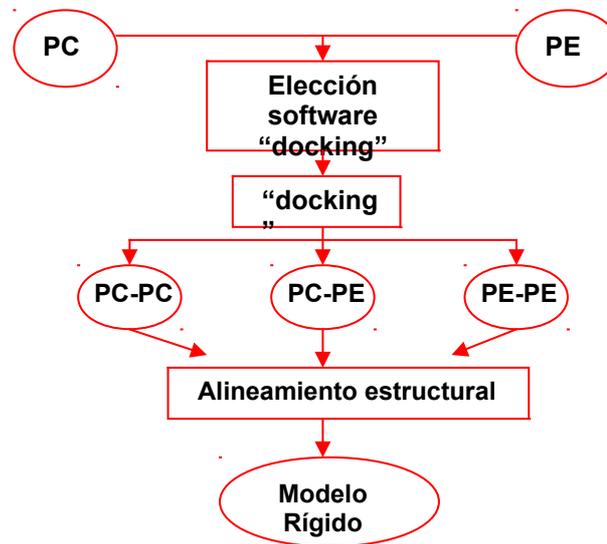


Figura 4. Diagrama de flujo de estandarización de metodología de “docking” de ficobiliproteínas.

3.1.1 Estructura de PC y PE

Se utilizaron las estructuras de las ficobiliproteínas R-Ficoeritrina (código PDB: 1eyx; Contreras-Martel y colaboradores, 2001) y C-Ficocianina (código PDB: 4TST; Bunster y colaboradores 2005 en preparación), ambas de *Gracilaria chilensis*, y determinadas en el laboratorio de Biofísica Molecular de la Universidad de Concepción.

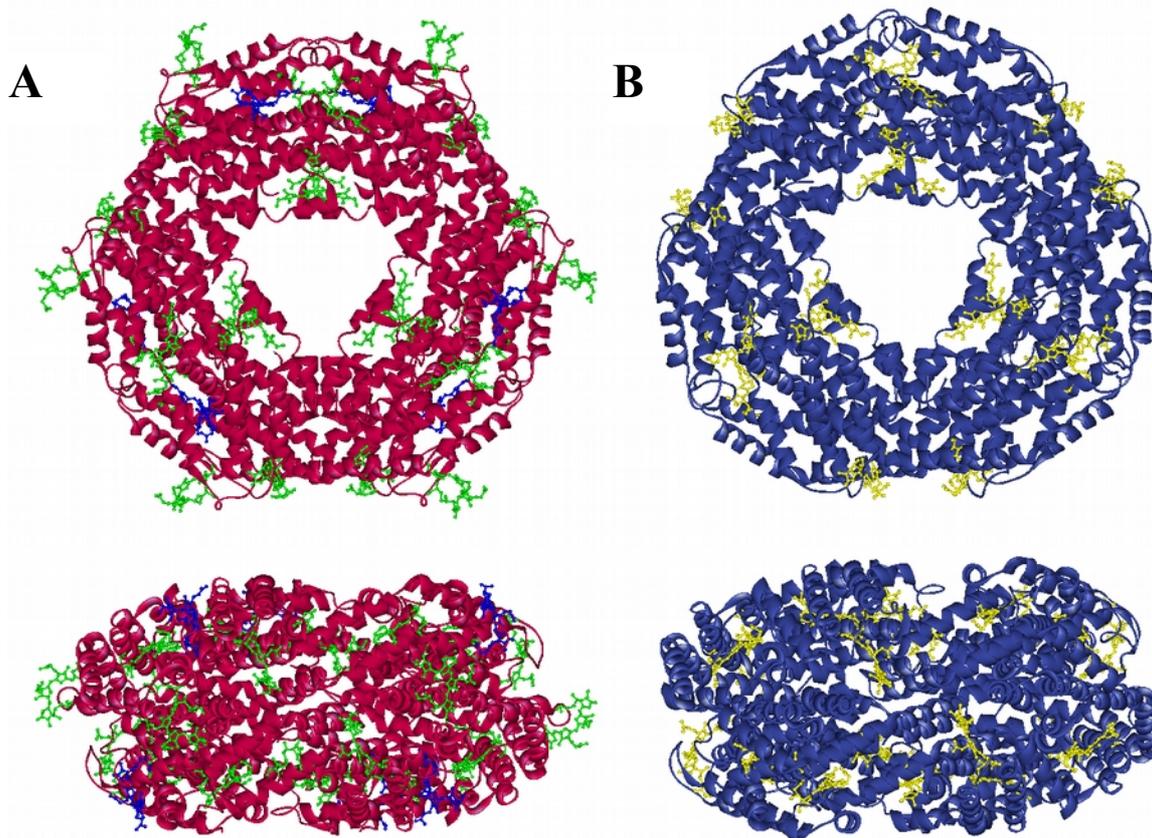


Figura 5. Estructura de las ficobiliproteínas de *Gracilaria chilensis* componentes de la porción de varillas. A: ficoeritrina; los cromóforos ficoeritrobilina se destacan en verde, y ficourobilina en azul. B: ficocianina; el cromóforo ficocianobilina se destaca en amarillo.

3.1.2 Selección del programa de “docking”

Para poder realizar un estudio de “docking” de proteínas, se seleccionó el programa de “docking” adecuado para el problema, considerando en la búsqueda la información generada en el último evento CAPRI (Critical Assessment of PRediction of Interactions) (<http://capri.ebi.ac.uk/>) (Janin y colaboradores, 2003), que evalúa a los software mediante una prueba donde se hacen competir a los distintos grupos que crean estos softwares, intentando predecir la correcta asociación entre 2 proteínas para generar un complejo proteína-proteína.

En base a los resultados obtenidos en este análisis, se tomó la decisión sobre que software de “docking” se utilizará en este trabajo. El software seleccionado es ZDOCK y el detalle de su elección se describe en el capítulo de “discusión”.

3.1.3 Protocolo de “docking”

El “docking” proteico es una técnica efectiva para predecir interacción entre proteínas, siempre y cuando se posean conocimientos previos sobre interacción entre las proteínas estudiadas. El procedimiento se realizó en dos etapas con el objeto de incorporar la información previa disponible y de este modo producir un modelo confiable de interacción de proteínas.

3.1.3.a Restricción del espacio de búsqueda

Una vez seleccionado el software, se utilizó la información experimental disponible para acotar el área de búsqueda, y luego para seleccionar los modelos apropiados de entre los resultados generados por un software de “docking”. Esta etapa es fundamental ya que un software de “docking” puede entregar miles de soluciones posibles para la interacción entre dos proteínas, y los criterios biológicos y resultados experimentales permiten una discriminación válida.

La restricción del área de búsqueda se hizo de acuerdo a las fotografías de microscopía electrónica de transmisión, que muestran una interacción tipo “back to back” entre las ficobiliproteínas (Lundell y colaboradores, 1981; figura 6), lo que junto

al análisis estructural realizado con SwissPDBViewer (Guex y colaboradores, 2003) en el cual se identificaron aminoácidos pertenecientes a la superficie, permitió seleccionar las regiones con posibilidades de interacción y por lo tanto reducir el tiempo de búsqueda.

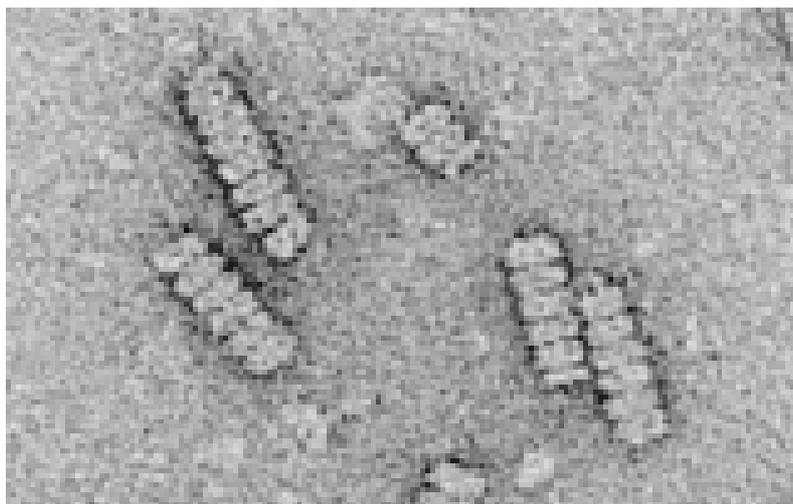


Figura 6. Micrografía electrónica de varillas de ficobilisomas. Se observan interacciones “back to back” entre las ficobiliproteínas (Lundell y colaboradores, 1981).

El software ZDOCK (Chen y colaboradores, 2003), seleccionado para llevar a cabo este estudio, permite agregar la información recopilada antes de realizar el protocolo de “docking” entre proteínas. La ejecución de ZDOCK se realizó en un computador Pentium IV con un procesador de 2.0 GHz de velocidad, 1 GB de memoria RAM, y con sistema operativo Linux Mandrake 10.0. A este software se puede también acceder a través de un servidor localizado en la Universidad de Boston (<http://zdock.bu.edu>).

El procedimiento de “docking” se realizó a un paso angular de 15°, con restricción del área de búsqueda y con un reporte de los 500 primeros modelos de

interacción. Más información sobre el uso del software ZDOCK se puede encontrar en el Apéndice 1 de esta tesis.

3.1.3.b Evaluación de resultados

Los modelos generados por ZDOCK se analizaron con el objetivo de seleccionar sólo aquellos consistentes con las observaciones experimentales y con criterios geométricos y energéticos.

Se usaron 3 criterios de evaluación:

- a) Evaluación visual
- b) Complementariedad geométrica
- c) Interacciones hidrofóbicas

Evaluación visual: utilizando el programa de visualización SwissPDBViewer, se seleccionaron sólo aquellos modelos que presentaban la conformación “back to back” observada experimentalmente por microscopía electrónica (Lundell y colaboradores, 1981). Todos los modelos seleccionados en base a este criterio fueron luego evaluados por los otros dos criterios.

Evaluación geométrica: permitió discriminar entre los modelos ya seleccionados considerando el número de aminoácidos integrados en la interfase que denotan una mayor área de interacción, y por ende mayor complementariedad geométrica.

Evaluación hidrofóbica: permitió determinar el porcentaje de los aminoácidos hidrofóbicos presentes en la interfase, dato importante si se considera que en complejos estables en el tiempo se observa una gran cantidad de interacciones del tipo hidrofóbicas (Vasker y colaboradores, 1994). Usando el software SwissPDBViewer se seleccionaron todos los aminoácidos encontrados en la interfase de la ambas proteínas, se creó un nuevo archivo pdb con estos datos y se cuantificó el porcentaje de aminoácidos hidrofóbicos en el total. A mayor porcentaje de aminoácidos hidrofóbicos, mayor es la interacción de este tipo.

Este procedimiento se realizó, como se mencionó anteriormente, para los 3 tipos de “docking” necesarios para generar una varilla mínima: PC – PC; PC – PE; y PE – PE.

3.1.4 Construcción de un primer modelo de varilla

Una vez obtenidos los 3 complejos ideales a partir del “docking” rígido, se procedió a armar el modelo “rígido” de la varilla, por simple alineamiento estructural de la ficobiliproteína redundante de cada complejo (una de las C-PC con la C-PC del complejo C-PC – R-PE, y luego la R-PE de este con otra del complejo R-PE – R-PE), utilizando el software SwissPDBViewer (Guex y colaboradores, 2003). Luego, se editó manualmente el nuevo archivo pdb generado para eliminar la proteína redundante en cada caso. A este modelo se le denominó “modelo rígido”, debido a que como ya se mencionó, en el “docking” de proteínas sus estructuras se consideran como cuerpos rígidos.

3.2 Estandarización de un protocolo de dinámica molecular para optimizar resultados de “docking” proteico

Esta etapa permite optimizar el modelo rígido desde un punto de vista estructural, ya que cuando dos proteínas interactúan, ellas sufren algunos cambios conformacionales, especialmente en los aminoácidos de la interfase, o en algunos casos, cambios más drásticos en zonas alejadas. Además es necesario refinar las distancias inter-atómicas generadas por el procedimiento de “docking”, que resultan a veces más cortas de lo permitido (Krippahl y colaboradores, 2003).

El procedimiento de dinámica molecular se aplicó al modelo rígido con el objetivo de buscar un mínimo energético global de acuerdo al diagrama de flujo indicado en la figura 7.

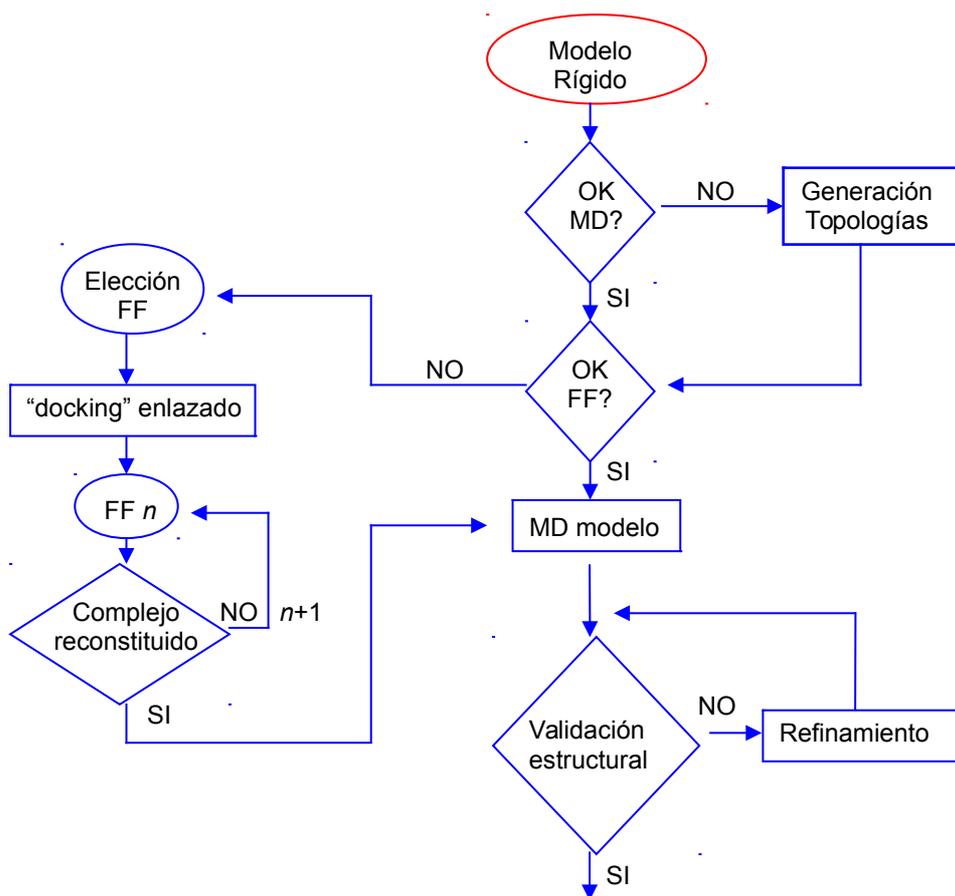


Figura 7. Diagrama de flujo que muestra el trabajo a realizar para llevar a cabo la optimización del modelo rígido por medio de una dinámica molecular (MD: dinámica molecular; FF: campo de fuerza).

Los campos de fuerza utilizados en esta optimización requieren la definición de todos los átomos participantes en el cálculo. En general estos están definidos para todos los aminoácidos y grupos que comúnmente están presentes en las proteínas. Sin embargo las ficobiliproteínas presentan los grupos cromóforos y una asparragina metilada en la posición $\beta 72$ que no se encuentran definidas. Por esta razón fue necesario crear las topologías para estos grupos para cada campo de fuerza utilizado en este estudio.

Una vez creadas las topologías requeridas se eligió el mejor campo de fuerza presente en GROMACS (Van der Spoel, 2004) capaz de optimizar modelos generados por “docking” de proteínas. Se procedió a realizar un “docking” enlazado de proteínas (Elcock y colaboradores, 2001) con el cual se discrimina entre los campos de fuerza encontrando aquel que es capaz de reconstituir de mejor manera un complejo cristalográfico a partir de un complejo generado por “docking”. Con estos resultados se procedió a realizar la dinámica molecular de la “varilla rígida” con el campo de fuerza adecuado. Una vez terminado el proceso, se evaluó estructuralmente el modelo generado. Una vez validado el modelo, se realizó la evaluación funcional, dada por la postulación de vías preferenciales de conducción de luz a través de la varilla. La metodología detallada de estas etapas es explicada a continuación.

3.2.1 Generación de topologías

La creación de topologías involucra la definición completa de cada compuesto, donde debemos definir:

- tipo de átomos que lo componen
- cargas parciales en los átomos del compuesto
- enlaces entre los átomos del compuestos
- ángulos entre los átomos
- ángulos dihedros propios e impropios
- interacciones tipo Van der Waals que no estén definidas

La Figura 8 muestra la estructura de uno de los cromóforos y la nomenclatura para cada uno de los átomos que requieren definición.

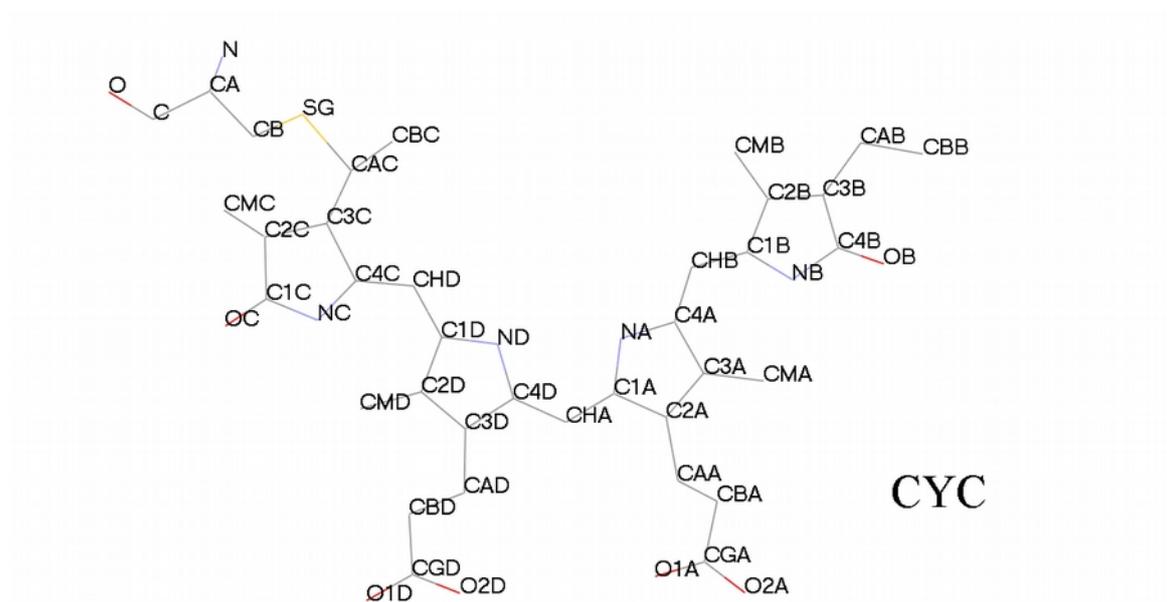


Figura 8. Esquema del cromóforo ficocianobilina. Se indican los nombres de los átomos (no hidrógenos) de acuerdo a como se denominan en los archivos de coordenadas (PDB)

Un archivo de topologías entrega toda la información estructural de este compuesto al programa, pues de otro modo, sólo lo consideraría como puntos discretos en el espacio sin relación unos con otros y sin saber cómo interactuarían al estar cercanos a otros.

Para la creación de topologías se utilizó el servidor PRODRG (<http://davapc1.bioch.dundee.ac.uk/programs/prodrg/prodrg.html>) (Schuettelkopf y colaboradores, 2004), al cual se ingresaron las coordenadas PDB del compuesto. Este servidor entregó la información topológica válida para una serie de softwares, entre ellos para GROMACS (Van der Spoel, 2004), para el cual entrega un archivo completo con información necesaria para definir topológicamente nuestro compuesto en el campo

de fuerza Gromos87 o ffgmx, como se denomina en GROMACS, pero no para el resto de los campos de fuerza, uno de los cuales, OPLS/AA (ffoplsaa) posee hidrógenos explícitos. Allí está contenida la información necesaria sobre cada átomo, tal como tipo de átomo y carga y la información enlazante necesaria para una definición topológica, donde es importante especificar:

- a) **enlaces:** se deben definir todos los enlaces covalentes que presente el compuesto. En la figura 8 se puede apreciar como ejemplo el enlace dado entre CAC y C3C. En este caso los enlaces se consideran como resortes o elásticos que unen a dos esferas las que pueden acercarse y alejarse una de otra, como se muestra en la figura 9.

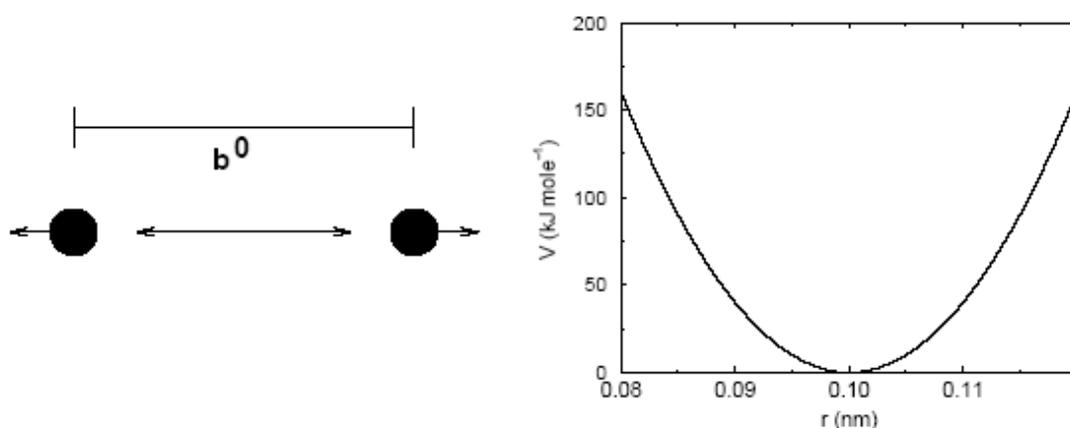


Figura 9. Esquema de un enlace junto a su gráfico de energía potencial. Se aprecia como existe un largo de enlace distinto de cero que está en el equilibrio, generando en dicha posición un mínimo de energía potencial (b^0 : distancia óptima de enlace; V : energía potencial del enlace; r : longitud del enlace). Imagen obtenida desde el manual de GROMACS.

- b) **Ángulos:** se deben definir todos los ángulos formados por los átomos unidos por enlaces covalentes, lo que ingresa la geometría que presenta la molécula. Un ejemplo es el ángulo formado por los enlaces entre los átomos CAC – C3C y C3C - C4C en la figura 8. Al igual que para los enlaces, se consideran 3 esferas unidas por un resorte o elástico (figura 10)

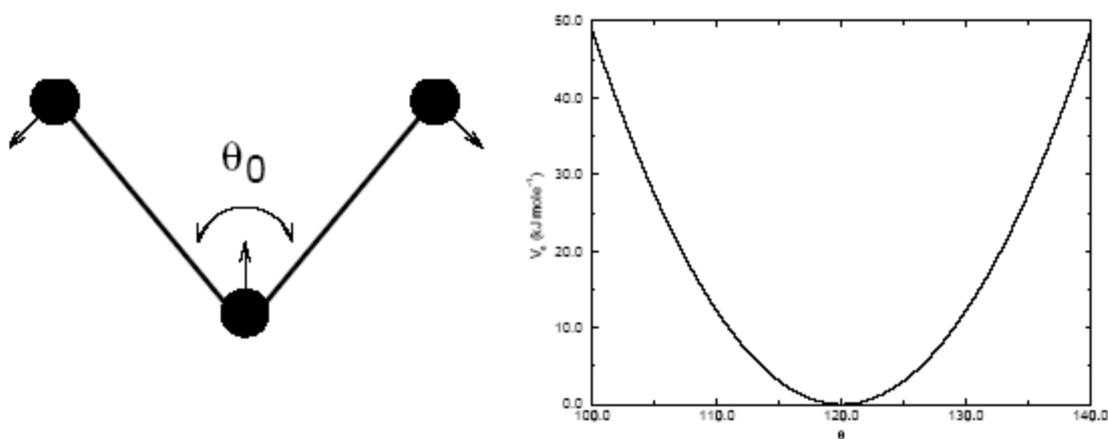


Figura 10. Esquema de un ángulo formado por 3 átomos enlazados junto a su gráfico de energía potencial. En el gráfico podemos observar como existe un ángulo en el cual el mínimo de energía potencial es alcanzado. Este valor del ángulo dependerá del tipo de átomos que estén participando (θ_0 : ángulo óptimo; V : energía potencial del ángulo; θ : ángulo medido). Imagen obtenida desde el manual de GROMACS.

- c) **Ángulos dihedros impropios:** se deben definir todos los ángulos dihedros impropios presentes en la estructura del compuesto (por ejemplo, el ángulo formado por los átomos C3B - C4B - C2B – CAB en la figura 8). Un ángulo dihedro impropio se define como el ángulo formado por los planos ijk y jkl (figura 11)

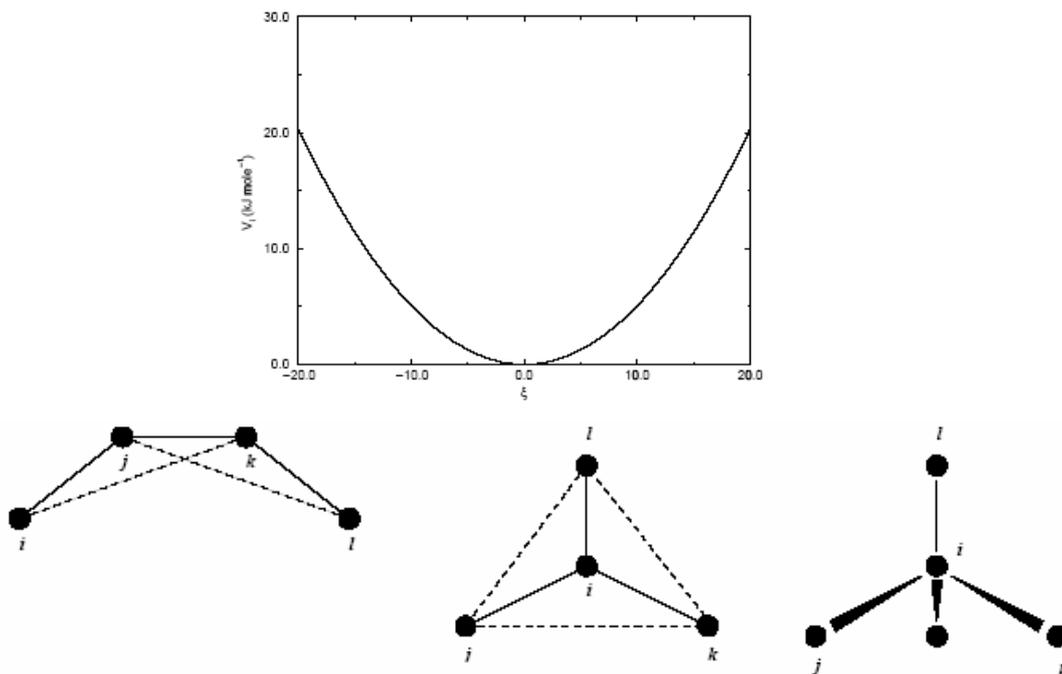


Figura 11. Esquema de las 3 formas en las cuales podemos encontrar un ángulo dihedro impropio, junto a su gráfico de energía potencial. El ángulo es la medida de los planos ijk y jkl . (V : energía potencial; ξ : medida del ángulo dihedro impropio). Imagen obtenida desde el manual de GROMACS.

- d) **Ángulos dihedros propios:** miden interacciones entre los átomos 1 y 4 que forman estos ángulos (figura 12, átomos i y l). Todas estas interacciones deben ser definidas, y la manera correcta de hacerlo es por inspección visual, rotando la estructura de manera de dejar eclipsados los átomos centrales (j y k), y ver si los átomos 1 y 4 (i y l) están hacia el mismo lado (*cis*) o en lados opuestos (*trans*). El ángulo está dado por los planos ijk y jkl . (Por ejemplo el ángulo formado por los átomos C4B - C3B - CAB - CBB de la figura 8.

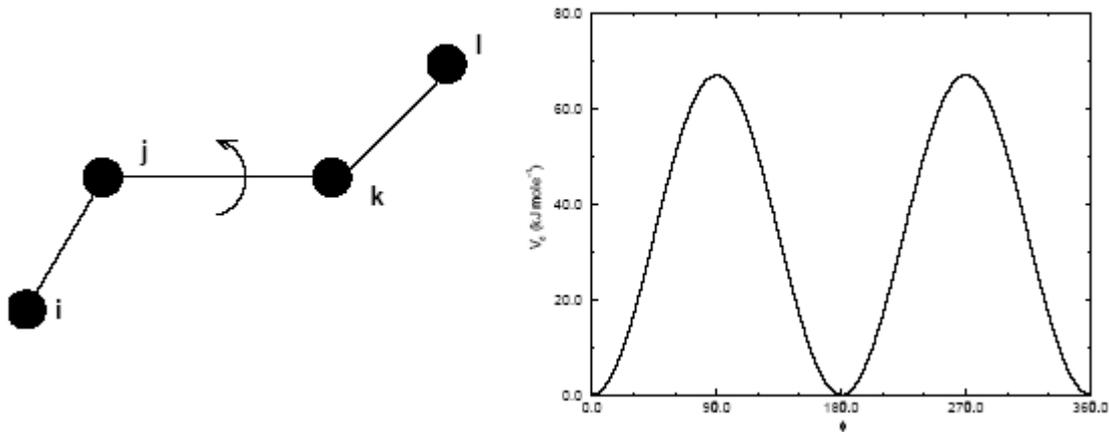


Figura 12. Esquema de un ángulo dihedro propio junto a su gráfico de energía potencial. La periodicidad está dada por el valor del $\cos \varphi$ (V : energía potencial; φ : ángulo formado por los planos ijk y jkl).

- e) **Interacciones de Van der Waals no definidas:** estas interacciones de tipo no enlazante deben ser definidas de acuerdo al tipo de átomo que participe en particular. Más información está disponible en el Apéndice 2 este documento.

3.2.2 Evaluación de campos de fuerza para reconstitución de complejos creados mediante “docking” de proteínas

En la literatura no existe información que apoye el uso de algún campo de fuerza específico para las simulaciones que involucran la optimización de modelos generados por “docking” de proteínas. Esto hizo necesaria la evaluación de los campos de fuerza disponibles con el objeto de seleccionar el más adecuado. Un campo de fuerza (ecuación 1) es aquella ecuación que permite evaluar el sistema en estudio desde el punto de vista de su energía, considerando los átomos como esferas unidas por resortes

y tratadas bajo las leyes de la mecánica clásica. El campo de fuerza involucra tanto las interacciones enlazantes como las no enlazantes.

Ecuación 1:

$$V(\vec{R}) = V_{\text{enlazado}}(\vec{R}) + V_{\text{noenlazado}}(\vec{R}),$$

$$V_{\text{enlazado}}(\vec{R}) = \left(\sum_{\text{enlace}} k_l (l - l_0)^2 + \sum_{\text{ángulos}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{impropios}} k_\omega (\omega - \omega_0)^2 + \sum_{\text{torsion}} A_n [1 + \cos(n\phi - \phi_0)] \right)$$

$$V_{\text{noenlazado}}(\vec{R}) = \sum_{i < j} \left(\epsilon_{ij} \left[\left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_r \epsilon_0 r_{ij}} \right)$$

Se trabajó con el software GROMACS (Van der Spoel, 2004), para realizar los cálculos de mecánica molecular. Este programa es muy versátil, pues permite una velocidad de cálculo 10 veces más rápida que otros programas, permite tratar cada aminoácido de una proteína por separado, y posee variadas herramientas de análisis. Se ejecutó bajo sistema operativo Linux (ver Apéndice 3), lo que abre la posibilidad de paralelizar el programa y en el futuro ejecutar el cálculo en un cluster si fuese necesario.

Como GROMACS permite la utilización de 5 campos de fuerza diferentes (Gromos87, Gromos96, Gromos96 con correcciones hidrofóbicas, Gromos96 para vacío y OPLS/AA; Van der Spoel, 2004), se hizo necesaria la elección del campo de fuerza adecuado al estudio aquí presentado.

El campo de fuerza se seleccionó realizando un ““docking” enlazado”, utilizando para ello el heterodímero $\alpha\beta$ de ficocianina de *Gracilaria chilensis*, es decir, un “docking” en el cual se conoce la solución (estructura del complejo) (figura 13). Se realizó el “docking” entre las subunidades α y β de ficocianina de *Gracilaria chilensis*

con el programa ZDOCK, exactamente igual a como lo haríamos para las ficobiliproteínas completas, siguiendo el protocolo descrito anteriormente (sección 2.1). Se asumió que ningún software de “docking” era capaz de generar un modelo idéntico al observado cristalográficamente. Por lo tanto un software de mecánica molecular con el campo de fuerza correcto debería ser capaz de regenerar la conformación observada experimentalmente. El motivo de usar las subunidades α y β de ficocianina de *Gracilaria chilensis* se debe a que es un sistema estable en el tiempo, y que además posee los grupos cromóforos. Para cada campo de fuerza se utilizó la topología específica.

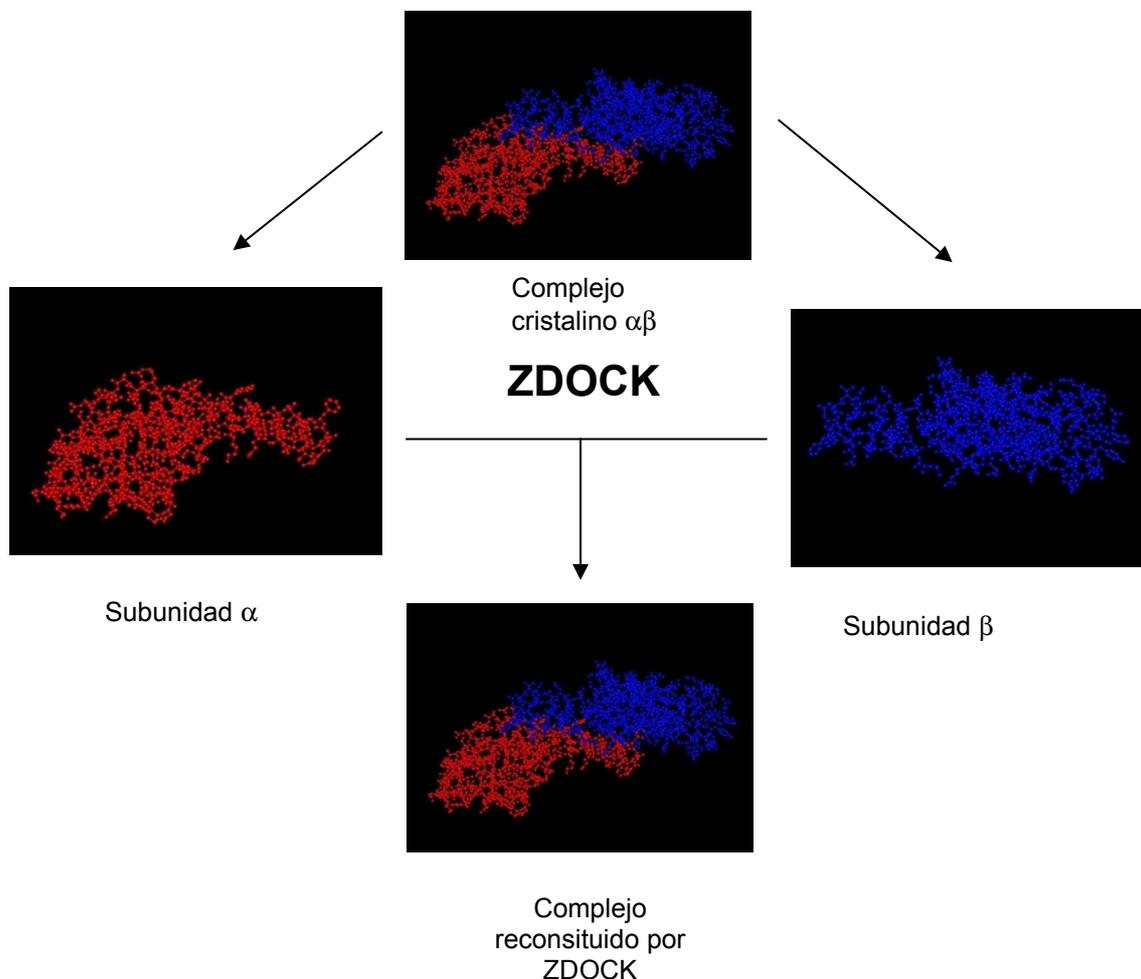


Figura 13. Esquema del “docking” dirigido. Se posee el complejo cristalográfico el cual se separa en sus dos componentes, y luego ZDOCK debe reconstituirlo.

Una vez obtenido el modelo de “docking” rígido de las subunidades alfa y beta de ficocianina de *Gracilaria chilensis* por ZDOCK, se realizó una dinámica molecular del complejo utilizando el siguiente protocolo (Van der Spoel, 2004):

- Minimización energética (“steepest descent”)
- Dinámica de pre equilibrio por 50 ps a 300 K
- Dinámica sin restricciones por 400 ps a 300 K
- Minimización energética (“steepest descent”)

Los procesos de dinámica molecular se monitorearon por su RMSD (“root mean square deviation”), que es la medida de comparación entre las posiciones de los átomos a cada instante con respecto a la estructura original (Maiorov y colaboradores, 1995), y por sus energías potenciales (Steinbach, 2003). Estos dos valores permiten evaluar si las estructuras han convergido a una que sea la más estable, es decir, a la estructura correspondiente al mínimo energético global.

Una vez analizados todos los campos de fuerza disponibles, se eligió el que proporcionó los mejores resultados para este problema de “docking” proteico, que resultó ser OPLS/AA o ffoplsaa como se denomina en GROMACS, y en el capítulo de “resultados”.

3.2.3 Dinámica molecular del modelo de varilla rígida

La Dinámica Molecular del modelo de varilla rígida se realizó utilizando el campo de fuerza “ffoplsaa” escogido en la etapa anterior. El protocolo seguido es el mismo utilizado en el estudio de reconstitución del “docking” enlazado, o sea:

- Minimización energética (“steepest descent”)
- Dinámica de pre equilibrio por 30 ps a 288 K
- Dinámica sin restricciones por 200 ps
- Minimización energética (“steepest descent”)

En este caso en la etapa de dinámica molecular sin restricciones, se desarrolló un protocolo de “simulated annealing” (Filizola y colaboradores, 1997; Van der Spoel, 2004), mucho más exhaustivo en la búsqueda del mínimo energético global. La temperatura de cálculo usada a los distintos tiempos se muestra en la Tabla 1.

	0 a 10 ps	10 a 150 ps	150 a 190 ps	190 a 200 ps
Proteína	288 - 303	303 - 303	303 - 288	288 - 288
CYC	288	288	288	288
CYE	288	288	288	288
CYU	288	288	288	288
SOL	288 - 303	303 - 303	303 - 288	288 - 288
Na+	288 - 303	303 - 303	303 - 288	288 - 288

Tabla 1. Temperaturas usadas (K) durante la dinámica molecular para cada uno de los componentes del sistema (CYC: cromóforo ficocianobilina; CYE: cromóforo ficoeritrobilina; CYU: cromóforo ficourobilina; SOL: solvente (agua); Na+: iones sodio para neutralizar la carga del sistema).

El protocolo de simulated annealing consiste en calentar todo el sistema, menos los cromóforos desde 288 a 303 K, para luego enfriar a la temperatura inicial, debido a

que los cromóforos, a temperaturas mayores a 290 K adquirirían una conformación distinta a la observada cristalográficamente, lo que potencialmente puede alterar los cálculos posteriores de vías preferenciales de conducción de luz. La temperatura utilizada, 288 K, es la temperatura promedio del agua de mar en las costas de la octava región, ambiente en el cual se encuentra el alga *Gracilaria chilensis*.

Al igual que en el caso de la búsqueda del campo de fuerza adecuado, la dinámica molecular se monitoreó a través del valor de RMSD (Maiorov y colaboradores, 1995) y su energía potencial (Steinbach, 2003) en busca de la convergencia hacia el mínimo global.

La estructura generada posterior a la Dinámica Molecular se evaluó estructuralmente, utilizando el gráfico de Ramachandran (Laskowski, 2004), donde se espera que al menos el 85% de los aminoácidos se encuentre en regiones permitidas (Laskowski, 2004). Los aminoácidos que se encontraron fuera de las zonas permitidas se ajustaron manualmente y se sometió la molécula a una minimización utilizando “Steepest Descent” en el mismo programa.

3.3 Postulación de vías preferenciales de conducción de luz en el modelo generado

El modelo estructural propuesto debe ser capaz de explicar la alta eficiencia de conducción de luz observada en una varilla del ficobilisoma.

La transferencia de luz en una varilla del ficobilisoma ocurre mediante el fenómeno conocido como FRET (transferencia de energía en resonancia) y puede ser

cuantificada de acuerdo a la teoría de Förster que define constantes de transferencia entre un par de cromóforos dador y aceptor a través de la ecuación 2

Ecuación 2:

$$\kappa_{DA} = \frac{9 \ln 10}{128\pi^5 N_A n^4} \times \frac{\kappa_{DA}^2}{R_{DA}^6} \times \frac{\Phi_D \epsilon_A}{\tau_D} \times \frac{\int_0^{\infty} F_D(\lambda) A_A(\lambda) \lambda^4 d\lambda}{\int_0^{\infty} F_D(\lambda) d\lambda}$$

(K_{DA} : constante de transferencia de energía; N_A : N° de Avogadro; n : índice de refracción del medio en el cual se encuentra el par de cromóforos dador y aceptor; κ_{DA} : factor de orientación; R_{DA} : distancia en Å del par de cromóforos dador y aceptor; Φ_D : rendimiento cuántico del cromóforo dador; ϵ_A : absorptividad molar máxima del cromóforo aceptor; τ_D : tiempo de vida media de fluorescencia del cromóforo dador en ausencia de un aceptor)

Esta ecuación puede ser resumida en cuatro términos:

$$K_{DA} = C * G * S * I$$

Donde C agrupa todas las constantes; el término G, corresponde al factor geométrico entre los pares de cromóforos dador y aceptor; S involucra las propiedades espectroscópicas del par de cromóforos dador aceptor; e I corresponde a las integrales de solapamiento de los espectros del dador y el aceptor.

El cálculo de la constante de energía en transferencia (K_{DA}) para las vías intra e inter ficoeritrina, así como inter ficoeritrina y ficocianina, se realizó usando una versión modificada de la ecuación de Förster (ecuación 3), ya que no se disponía de los valores de rendimiento cuántico, absorptividad molar máxima e integrales de solapamiento para Ficocurobilina y Ficocoeritrobilina.

Ecuación 3:

$$K_{DA} = (\kappa^2/\tau_D)(R_0/R)^6$$

κ : Factor de orientación

τ_D : Vida media del cromóforo dador

El factor de orientación κ^2 se calculó semi empíricamente en Gaussian98®

R : Distancia entre cromóforos

usando el método mecanocuántico semi-empírico PM3 (Frisch y colaboradores, 2002), de acuerdo a la ecuación 4:

Ecuación 4:

$$\kappa = \mu_D \cdot \mu_A - 3 (\mu_D \cdot r_{DA}) (\mu_A \cdot r_{DA})$$

donde μ_D y μ_A son vectores unidad que describen la dirección de los momentos dipolares del par de cromóforos dador y aceptor, y r_{DA} es el vector unidad que describe la dirección de la línea que conecta los centros de masa de los cromóforos que interactúan. El Radio de Förster, R_0 , está definido como la distancia entre el par de cromóforos dador y aceptor a la cual la eficiencia de la transferencia de la energía es igual al 50%. Este último valor ha sido aproximado a 50 Å en trabajos previos (Duerring y colaboradores, 1991) y se ha observado que el uso de este valor es aceptable y congruente con los resultados obtenidos usando la ecuación completa para ficociannina. No obstante, observaciones experimentales han mostrado que dicho valor de radio de Förster puede oscilar entre 44 y 63 Å (Debreczeny y col. 1995). La distancia entre los cromóforos, R , se obtuvo de sus coordenadas desde el archivo pdb. El tiempo de vida media de fluorescencia del cromóforo dador en ausencia de un aceptor, τ_D , se obtuvo de la literatura. (Otto y colaboradores, 2003).

El análisis de los valores de las constantes de transferencia entre cromóforos, con el software Maple 7®, permitió describir las vías preferenciales de conducción de luz en el modelo estructural propuesto. Para ello se tabularon los valores correspondientes a cada para Dador-Aceptor y se escogieron las vías que presentaban

mayores Constantes y menores tiempos de transferencia de energía en resonancia, considerando los valores reportados en la literatura para estos procesos *in vivo*.

NOTA: En el apéndice 4 de esta tesis está la información relacionada con el uso de los programas anteriormente mencionados.

4.- RESULTADOS

4.1 Estandarización de una metodología de “docking” proteico para ficobiliproteínas

4.1.1 Elección del software de “docking”

Como se mencionó en la metodología, la primera parte del trabajo consistió en la elección de un software adecuado para realizar “docking” proteico. Para este propósito se utilizó la información generada por el evento CAPRI (Janin y colaboradores, 2003), y las características propias de cada software.

Todo el software disponible y analizado se ejecutó bajo el sistema operativo Linux a excepción de Bigger (Krippahl y colaboradores 2003) que funciona en Windows. Para ello se instaló el sistema operativo Linux Mandrake 10.0 (www.mandrakelinux.com), que posee una serie de librerías matemáticas de cálculo, tales como transformada rápida de Fourier, pre compiladas y de fácil instalación a diferencia de otras distribuciones Linux.

El software seleccionado para realizar los estudios de “docking” de proteínas, fue el software ZDOCK (Chen y colaboradores, 2003). En el capítulo de “discusión” se puede encontrar mayor información acerca propiedades del software, así como la comparación con otros softwares similares; más detalles se incluyen en el Apéndice 1.

4.1.2 “docking” de ficobiliproteínas

Se realizaron los 3 procedimientos de “docking” con ZDOCK necesarios para obtener una conformación de una varilla formada por 4 ficobiliproteínas: 2 ficoeritrinas y 2 ficocianinas.

Un factor importante para la selección del mejor modelo de interacción lo constituye la evidencia experimental (figura 6), que indica una interacción del tipo “back to back” entre las ficobiliproteínas en una varilla (Lundell y colaboradores, 1981). Por este motivo, se seleccionó por inspección visual de los 500 resultados entregados por ZDOCK, con la ayuda del software SwissPDBViewer, sólo aquellos complejos que presentaban conformación “back to back”. Los complejos seleccionados se analizaron considerando sus interfases. Se desarrolló con el objeto de análisis, un script (archivo de comandos) para bash (intérprete de comandos en Linux y Unix) (figura 15), que permitió obtener el número de aminoácidos y el porcentaje de aminoácidos hidrofóbicos presentes en dicha superficie. No obstante, dicho script solo lee un archivo pdb que le entregamos, por lo que las superficies de interacción deben ser definidas antes, por ejemplo con SwissPDBViewer, definiendo, como en esta tesis, los aminoácidos integrados en la interfase a aquellos de proteínas distintas que estén a menos de 5 Å de distancia.

```

#!/bin/sh
### SCRIPT THAT CALCULATES THE NUMBER OF RESIDUES IN THE INTERFACE
### AND THE PERCENTAGE OF HIDROFOBIC RESIDUES
###
### ahora no es necesario tener un nombre especifico o archivo .pdb
### modo de uso:
### ./calculator2.0 archivo_para_analisis
### disfrutenlo
grep "^ATOM" ${1} > a
cut -c 18-26 a > b
uniq b > c
value1=`wc -l c | cut -c 1-3`
echo "#####" > ${1}_interface.txt
echo "Análisis de la interface del sistema ${1}" >> ${1}_interface.txt
echo "#####" >> ${1}_interface.txt
echo "El número de aminoácidos en la interface es" >> ${1}_interface.txt
echo " $value1" >> ${1}_interface.txt
cut -c 5 c > e
echo "de una proteína son" >> ${1}_interface.txt
grep A e > porsiac1
value100=`wc -l porsiac1 | cut -c 1-4`
echo " $value100" >> ${1}_interface.txt
echo "y de la otra son" >> ${1}_interface.txt
grep B e > porsiac2
value101=`wc -l porsiac2 | cut -c 1-4`
echo " $value101" >> ${1}_interface.txt
grep CYS c >> d
grep ALA c >> d
grep VAL c >> d
grep LEU c >> d
grep ILE c >> d
grep PHE c >> d
grep MET c >> d
grep PRO c >> d
value2=`wc -l d | cut -c 1-3`
echo "#####" >> ${1}_interface.txt
echo "de ellos son hidrofóbicos" >> ${1}_interface.txt
echo " $value2" >> ${1}_interface.txt
echo "de una proteína son" >> ${1}_interface.txt
cut -c 5 d > f
grep A f > porsiac3
value102=`wc -l porsiac3 | cut -c 1-3`
echo " $value102" >> ${1}_interface.txt
echo "y de la otra son" >> ${1}_interface.txt
grep B f > porsiac4
value103=`wc -l porsiac4 | cut -c 1-3`
echo " $value103" >> ${1}_interface.txt
echo "#####" >> ${1}_interface.txt
echo "### gracias por utilizar calculator2.0 ###" >> ${1}_interface.txt
echo "#####" >> ${1}_interface.txt
echo "##### by dalmonac y maxfigue #####" >> ${1}_interface.txt
rm -f a
rm -f b
rm -f c
rm -f d
rm -f e
rm -f f
rm -f porsiac1
rm -f porsiac2
rm -f porsiac3
rm -f porsiac4
cat ${1}_interface.txt

```

Figura 14. Script en bash “Calculator2.0”. Ejecutable en ambiente UNIX o LINUX es capaz de leer un archivo pdb y entregar los resultados tanto en pantalla como guardarlos en un archivo de texto.

A continuación se presentan los resultados para cada uno de los “docking” por separado.

4.1.2.1 “docking” ficocianina – ficocianina

Se analizaron los 500 resultados proporcionados por ZDOCK a un paso angular de 15°, seleccionándose sólo aquellos que tuvieran la conformación back to back. Los complejos seleccionados junto con una caracterización de sus interfases se muestran en la tabla 2.

Complejo	aa en la superficie de A	aa en la superficie de B	aa hidrófobos en la superficie de A	aa hidrófobos en la superficie de B	total aa en la superficie	% aa hidrófobos en la superficie
Complejo1	172	173	53	53	345	30,72
Complejo2	170	171	50	51	341	29,62
Complejo3	184	187	58	62	371	32,35
Complejo4	168	175	47	55	343	29,74
Complejo5	164	171	46	53	335	29,55
Complejo6	171	174	52	53	345	30,43
Complejo7	167	171	48	46	338	27,81
Complejo10	175	168	55	52	343	31,20
Complejo11	155	158	44	44	313	28,12
Complejo20	167	162	48	44	329	27,96
Complejo22	145	149	39	42	294	27,55
Complejo26	162	165	43	49	327	28,13
Complejo27	143	140	43	37	283	28,27
Complejo57	146	139	35	34	285	24,21
Complejo63	141	141	39	40	282	28,01

Tabla 2. Resultados de “docking” ficocianina – ficocianina y análisis de los aminoácidos presentes en la interfase del complejo generado (aa: aminoácidos; A y B: ficocianinas involucradas en el “docking”; la numeración de los complejos señala el ranking dado por ZDOCK). En negro se destaca el complejo con las mejores características para un complejo proteína – proteína estable en el tiempo.

Al observar estos resultados se puede apreciar claramente que las funciones de evaluación que posee ZDOCK en conjunto con las acotaciones del área de búsqueda,

son bastante eficiente, pues ya que se encontraron 15 complejos con la disposición “back to back”, y 8 de estos están dentro de los 10 primeros lugares, lo que demuestra su gran asertividad. El análisis de las interfases muestra que el complejo 3 destaca en el grupo, presentando el mayor número de aminoácidos integrados en la interfase, lo que denota una mayor complementariedad geométrica, con un mayor porcentaje de aminoácidos hidrofóbicos en la interfase, lo que demuestra una mayor estabilización por este tipo de interacciones. En consecuencia se eligió el complejo 3 como modelo de interacción entre ficocianinas.

4.1.2.2 “docking” ficocianina – ficoeritrina

Los resultados, obtenidos con un procedimiento similar al señalado en 3.1.2.1 se muestran en la tabla 3.

Complejo	aa en la superficie de A	aa en la superficie de B	aa hidrofobos en la superficie de A	aa hidrofobos en la superficie de B	Total de aa en la superficie	% aa hidrofobos en la superficie
Complejo2	127	134	49	44	261	35,63
Complejo4	122	137	45	43	259	33,98
Complejo6	117	132	47	44	249	36,55
Complejo382	93	80	38	20	173	33,53
Complejo390	84	83	34	30	167	38,32

Tabla 3. Resultados de “docking” ficocianina – ficoeritrina y análisis de los aminoácidos presentes en la interfase del complejo generado (aa: aminoácidos; A y B: ficocianina y ficoeritrina respectivamente involucradas en el “docking”; la numeración de los complejos señala el ranking dado por ZDOCK). En negro se destaca el complejo con las mejores características para un complejo proteína – proteína estable en el tiempo.

En este caso el número de complejos con la conformación “back to back” es menor. A pesar de la alta similitud estructural, se trata de proteínas distintas que

presentan superficies de interacción diferentes. Al mismo tiempo su gran simetría hace que se acentúen pequeñas diferencias, lo que restringe el número de modelos de interacción aceptables de acuerdo a las reglas de selección. Sin embargo, vemos que 3 de los 5 complejos seleccionados están dentro de los 10 primeros lugares, y que el complejo 2 presenta las mejores características sobresaliendo al resto. Se seleccionó al complejo 2 como modelo de interacción Ficocianina – Ficoeritrina.

4.1.2.3 “docking” ficoeritrina – ficoeritrina

El modo de proceder es similar a los dos casos anteriores, 3.1.2.1 y 3.1.2.2 . Los resultados se muestran en la tabla 4.

Complejo	aa en la superficie de A	aa en la superficie de B	aa hidrófobos en la superficie de A	aa hidrófobos en la superficie de B	Total aa en la superficie	% aa hidrófobos en la superficie
Complejo5	119	113	46	45	232	39,22
Complejo10	119	113	46	46	232	39,66
Complejo19	115	119	43	49	234	39,32
Complejo109	116	114	46	45	230	39,57
Complejo240	118	114	48	48	232	41,38
Complejo308	148	146	59	59	294	40,14
Complejo313	117	114	47	44	231	39,39
Complejo422	112	115	46	46	227	40,53

Tabla 4. Resultados de “docking” ficoeritrina – ficoeritrina y análisis de los aminoácidos presentes en la interfase del complejo generado (aa: aminoácidos; A y B: ficoeritinas involucradas en el “docking”; la numeración de los complejos señala el ranking dado por ZDOCK). En negro se destaca el complejo con las mejores características para un complejo proteína – proteína estable en el tiempo.

Se puede apreciar que de los 8 complejos seleccionados, el complejo 308 es el que destaca ya que presenta el mayor número de aminoácidos integrados en la interfase, lo que denota su mayor complementariedad. Utilizando información experimental de

estudios de interacciones entre complejos estables en el tiempo, la misma usada en los dos casos anteriores, se seleccionó el complejo 308 como modelo de interacción Ficoeritrina – Ficoeritrina. Una comparación entre los tres complejos seleccionados se muestra en la tabla 5..

Complejo	aa en A	aa en B	aa hidrófobos en A	aa hidrófobos en B	total aa	% de aa hidrófobos
PC – PC	182	185	67	71	367	37,60
PC – PE	127	134	49	44	261	35,63
PE – PE	148	146	59	59	294	40,14

Tabla 5. Caracterización de las interfases de los complejos seleccionados en cada “docking” (aa: aminoácidos; A y B: proteínas involucradas en el “docking”; PC: ficocianina; PE: ficoeritrina)

Utilizando los modelos de “docking” generados (PC – PC; PC – PE; PE – PE), se construyó la antena por alineamiento estructural de la ficobiliproteína redundante. La figura 15 ilustra estos 3 “docking” rígidos, donde en colores idénticos se muestra la ficobiliproteína redundante que debe alinearse estructuralmente con su homóloga:

a) PC – PC

b) PC – PE

c) PE – PE



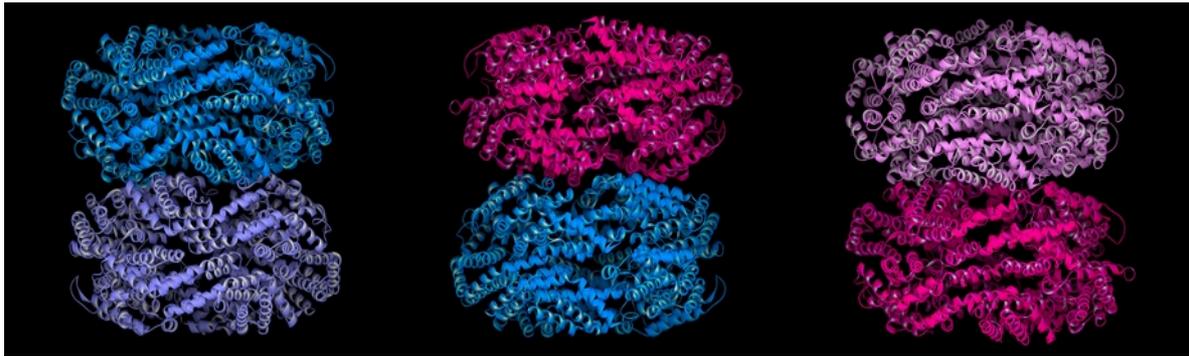


Figura 15. Complejos seleccionados en cada “docking”. a) Complejo Ficocianina-Ficocianina, b) Complejo Ficocianina - Ficoeritrina, c) Complejo Ficoeritrina-Ficoeritrina. Con flechas se indica la ficobiliproteína que debe ser alineada estructuralmente para formar el primer modelo de varilla.

Después de realizar el alineamiento estructural con el software SPDBV, se editó el archivo PDB generado con el objeto de eliminar una de las ficobiliproteínas redundantes en cada caso (una PC y una PE) y de este modo se construyó el modelo rígido de la varilla, generado a partir de 3 “docking” de cuerpo rígido, como se muestra en la figura 16.

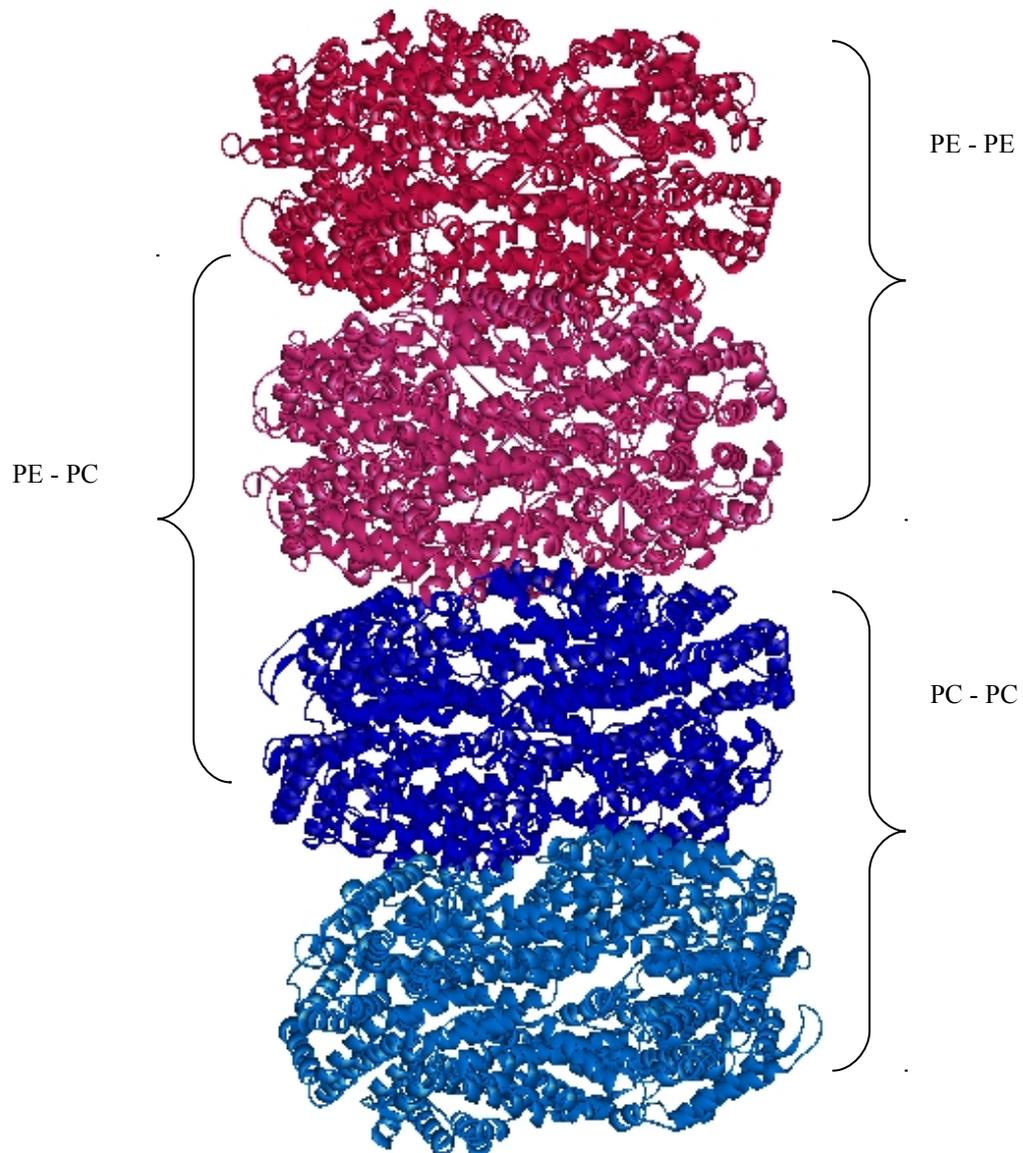


Figura 16. Modelo de la varilla construido a partir del alineamiento estructural de los complejos obtenidos en los 3 "docking" individuales. En el extremo superior en colores rojizos se observan R-ficoeritrina, y en colores azulados C-ficocianina.

4.2 Estandarización de un protocolo de dinámica molecular para optimizar resultados de “docking” proteico

El modelo rígido de varilla se optimizó por dinámica molecular y los resultados de las distintas etapas se muestran separadamente a continuación.

4.2.1 Creación de topologías

En los campos de fuerza que posee el software para mecánica molecular GROMACS, no están definidos topológicamente ninguno de los cromóforos, ni tampoco la asparragina metilada que poseen estas ficobiliproteínas en la posición $\beta 72$. Las topologías para la asparragina $\beta 72$ y ficocianobilina fueron creadas para los 5 campos de fuerza considerados. Las topologías para ficoeritrobilina y ficourobilina se definieron sólo para el campo de fuerza que se decidió ocupar para realizar la dinámica molecular de la varilla rígida. Los cromóforos fueron definidos entonces en el campo de fuerza OPLS/AA, debido a los resultados obtenidos en el estudio de “Evaluación de campos de fuerza” cuyos resultados son presentados a continuación en la sección 3.2.2. Estas topologías están disponibles en el laboratorio de Biofísica Molecular de la Universidad de Concepción.

4.2.2 Evaluación de campos de fuerza para reconstitución de complejos creados mediante “docking” de proteínas

Se evaluaron los 5 campos de fuerza presentes en GROMACS (Gromos87 (ffgmx), Gromos96 (ffG43a1), Gromos96 con correcciones hidrofóbicas (ffG43a2), Gromos96 para vacío (ffG43b1) y OPLS/AA (ffoplsaa). El procedimiento de prueba consistió en realizar un “docking” dirigido del heterodímero alfa-beta de ficocianina de *Gracilaria chilensis* al que se le introdujeron errores (desplazamiento de la subunidad β) que deberían corregirse mediante un proceso de dinámica molecular que reconstruya el complejo cristalográfico. El “docking” dirigido se realizó a un paso angular de 15°, dejando fija la subunidad α y rotando la β . Cada solución se evaluó de acuerdo al rmsd presentado por la subunidad β con respecto a su posición espacial en la estructura cristalina del complejo $\alpha\beta$. Se seleccionaron complejos que se asemejaban a la orientación cristalográfica sin ser idénticos. De estos 50 complejos observados se seleccionaron los 5 mejores (tabla6).

Complejo	rmsd (Å)
Complejo1	2,46
Complejo2	3,89
Complejo7	5,19
Complejo15	7,77
Complejo45	5,3

Tabla 6. Complejos seleccionados en el “docking” enlazado. El rmsd es calculado comparando la posición de la subunidad β con respecto a su posición en el complejo cristalino.

Aunque el complejo 1 (tabla 6) constituye un buen modelo, con sólo 2.46 Å de rmsd, se eligió trabajar con el complejo 7 (figura 17), pues presenta una desviación mayor respecto al complejo cristalino (rmsd = 5.19Å), con el objeto de identificar cuál

campo de fuerza utilizado por la dinámica molecular, es capaz de reconstituir mejor el complejo reportado experimentalmente.

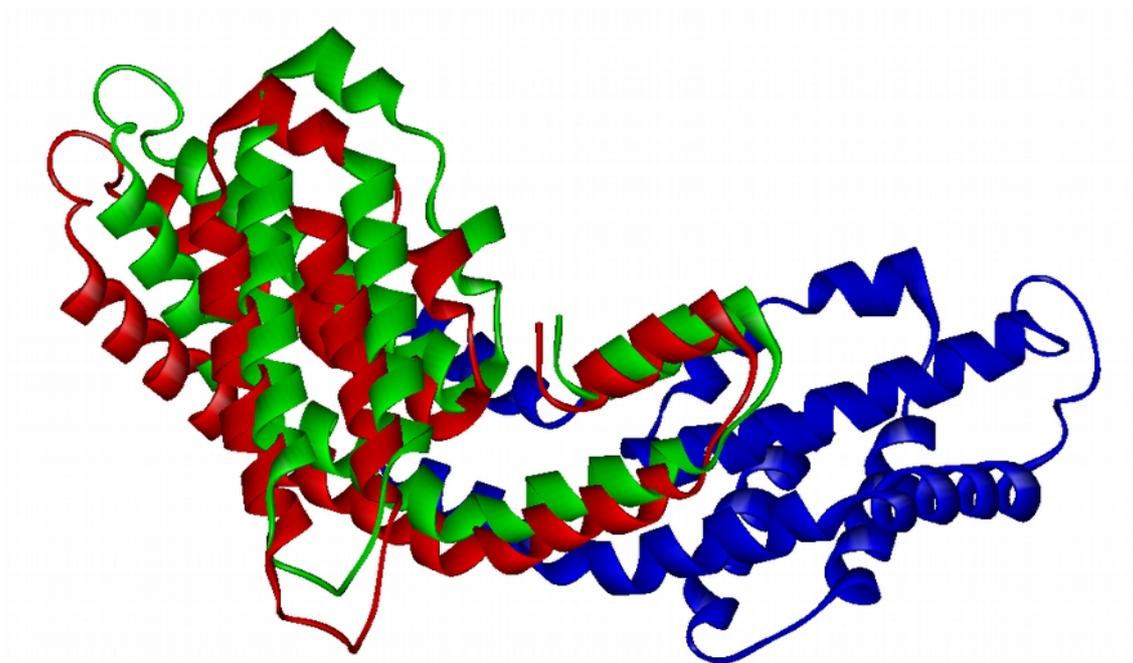


Figura 17. Complejo 7, con un rmsd de 5.19 Å. En azul subunidad α ; verde: subunidad β del complejo cristalino; rojo: subunidad β del complejo 7.

El complejo 7 se sometió al protocolo de dinámica molecular indicado en la sección metodología. A continuación se presenta el gráfico del rmsd del esqueleto polipeptídico a lo largo de los 400 ps de la dinámica (figura 18), donde se muestra la convergencia de estas estructuras:

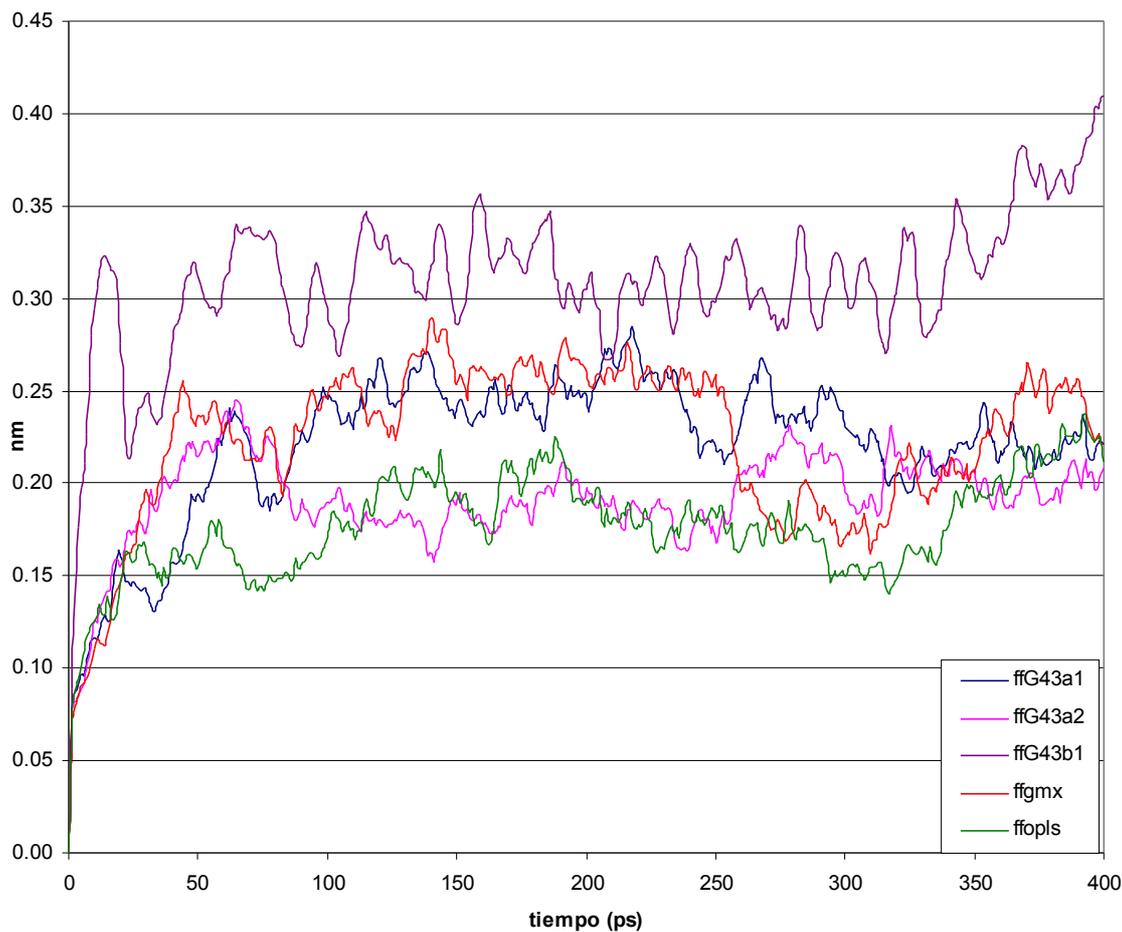


Figura 18. Gráfico del rmsd del esqueleto polipeptídico del heterodímero $\alpha\beta$ de ficocianina de *Gracilaria chilensis* para los 5 campos de fuerza. Se observa en el gráfico que a partir de los 50 ps aproximadamente casi no hay variación en el valor de rmsd, lo que indica que ha llegado a una conformación estable.

Cada una de las dinámicas además fue monitoreada por su energía potencial, hasta que alcanzó un plateau de mínima energía. Como ejemplo de este plateau, se muestra el gráfico de energía potencial (figura 19) para el campo de fuerza ffG43b1 (gromos 96 para vacío)

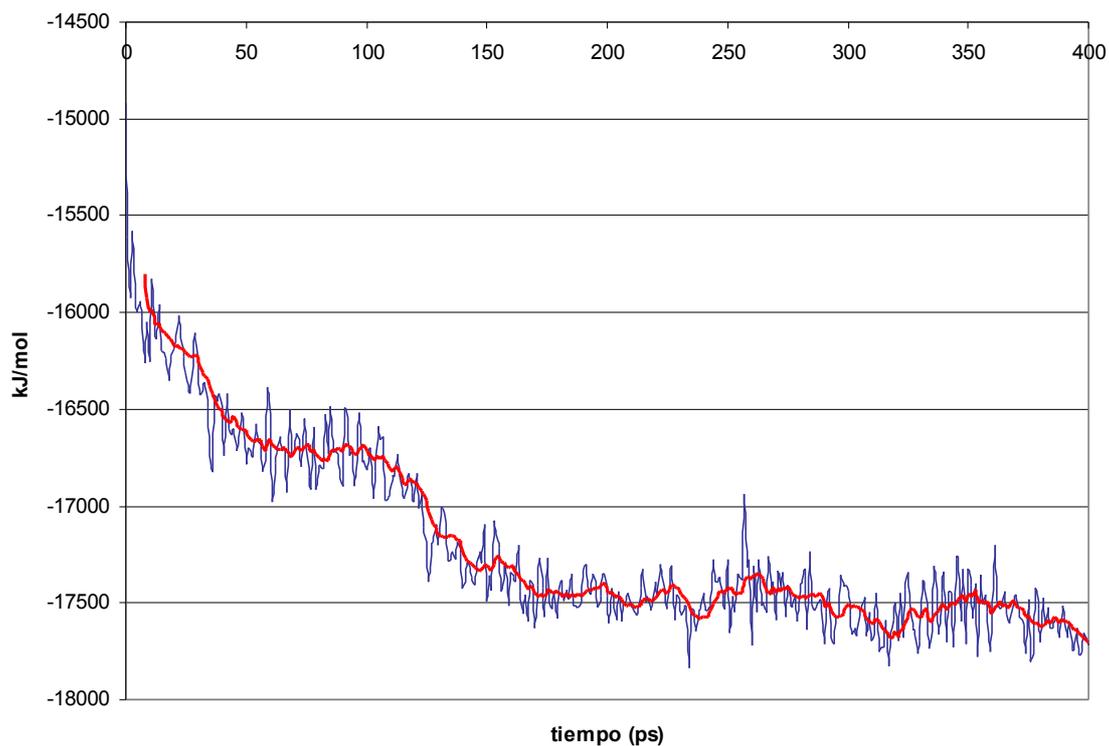


Figura 19. Gráfico de energía potencial del sistema a lo largo de los 400 ps de dinámica molecular para el heterodímero $\alpha\beta$ de ficocianina de *Gracilaria chilensis* en el campo de fuerza ffG43b1 (Gromos96 para vacío). En rojo se muestra la línea de promedio de las oscilaciones energéticas, observándose como a partir de los 150 ps se llega a un mínimo energético que es mantenido en el tiempo.

Finalmente se compararon las diferencias en energía potencial (tabla 7). Este ΔE se calcula considerando el valor de energía que presentaba la estructura al inicio de la dinámica de pre equilibrio y la energía al final de la MD (promedio de los últimos 25 ps).

Campo de fuerza	Δ energía potencial (kJ/mol)
ffG43a1	32834
ffG43a2	32733
ffG43b1	4096
ffgmX	33100
ffoplsaa	40698

Tabla 7. Δ energía potencial al término de la dinámica molecular. La comparación de estos valores da cuenta la capacidad de minimizar energéticamente un sistema.

Se evaluaron estructuralmente los resultados para la estructura de mínima energía obtenida con cada campo de fuerza, mediante el cálculo de rmsd del modelo respecto de la estructura cristalina (tabla 8). Se incluye además los resultados de rmsd para las subunidades α y β , previa una alineación estructural de las subunidades α , con el propósito de observar el acercamiento de la subunidad β hacia la posición cristalográfica.

Campo de Fuerza	rms total Å	alineando alfa	
		rms alfa Å	rms beta Å
ffG43a1	2,74	1,45	5,84
ffG43a2	2,57	1,57	5,94
ffG43b1	4,39	2,81	8,82
ffoplsaa	1,62	0,91	3,14
ffgmX	2,43	1,66	5,91
complex7	1,73	0	5,23

Tabla 8: Evaluación estructural de los complejos luego de finalizada la dinámica molecular de 400 ps.

La evaluación involucró además un estudio de las interfases, caracterizando el tamaño de éstas y las características hidrofóbicas (tabla 9). Estos datos se obtuvieron utilizando el servidor Protein-Protein Interaction Server (<http://www.biochem.ucl.ac.uk/bsm/PP/server/index.html>).

Campo de Fuerza	Área de interfase (Å²)	%átomos polares	%átomos no-polares
ffG43a1	1395	31,37	68,6
ffG43a2	1524	32,22	67,7
ffG43b1	1608	30,07	69,9
ffgmX	1645	31,33	68,6
ffoplsaa	1510	26,32	73,6
Cristalografico	1561	28,89	71,1

Tabla 9. Análisis de las interfases acabada la dinámica molecular de 400 ps.

De acuerdo a los resultados observados, se decidió proseguir con el estudio con el campo de fuerza OPLS/AA (ffoplsaa), debido a que éste presentó los mejores resultados estructurales (Tabla 8) y además presentó el mayor ΔE , lo que demuestra una estructura más estable.

4.2.3 Dinámica molecular del modelo de varilla de rígida

La varilla generada por “docking” rígido se sometió a un procedimiento de dinámica molecular utilizando el campo de fuerza OPLS/AA (ffoplsaa) evaluado como el mejor para este tipo de estudio.

El protocolo seguido en esta dinámica es similar al usado en el “docking” de las subunidades α - β , pero con menos tiempo en la dinámica de pre equilibrio (30 ps) con 200ps para la dinámica sin restricciones. El procedimiento se realizó a temperatura no constante de acuerdo a un protocolo de simulated annealing (Filizola y colaboradores 1997) calentando y enfriando la estructura (Figura 20).

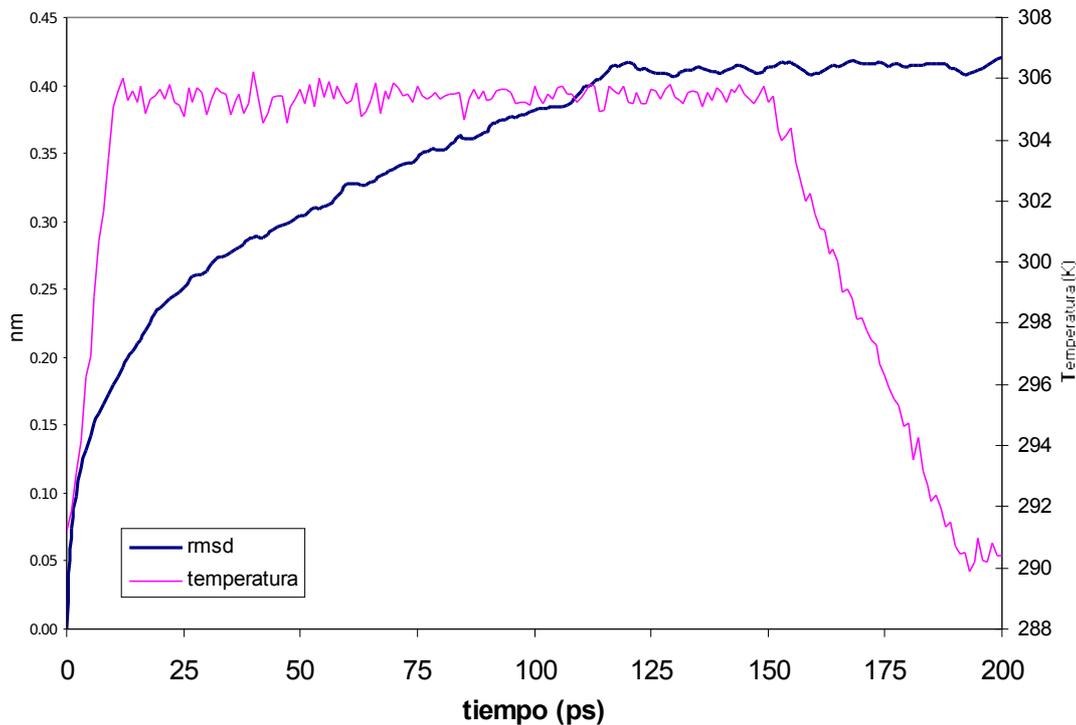


Figura 20. Gráfico del rmsd del esqueleto polipeptídico y la temperatura de éste a lo largo de la dinámica molecular del modelo de varilla de ficobilisoma de *Gracilaria chilensis*. Se observa que la temperatura no es constante (ver detalle en el texto), y la estructura es estable a partir de los 125 ps.

En la figura 20 podemos observar la convergencia de la estructura en un plateau de rmsd; donde esta estructura final posee un rms de 4.26 \AA con respecto a la inicial. En esta misma figura, en magenta, se aprecia la temperatura del esqueleto a lo largo de la dinámica, con 10 ps de calentamiento, 140 ps de mantención, 40ps de enfriamiento, y 10ps de mantención. Posterior a la dinámica molecular, el modelo fue analizado estructuralmente. La Tabla 10 muestra los datos geométricos de la estructura al término de la dinámica obtenidos a partir del gráfico de Ramachandran (servidor PROCHECK <http://biotech.ebi.ac.uk:8400/>)

regiones más favorables	5863 aa	85,27% (94,26%)
regiones permitidas	910 aa	13,23% (5,31%)
regiones generosamente permitidas	67 aa	0,97% (~ 0,00%)
regiones no permitidas	36 aa	0,52% (0,003%)
Total	6876 aa	100%

Tabla 10. Análisis del modelo de antena de acuerdo a datos obtenidos a partir del Gráfico de Ramachandran. Valores entre paréntesis indican el porcentaje de los aminoácidos en esas regiones antes de efectuarse la dinámica molecular.

Se observa que sobre el 85% de los aminoácidos están en las “regiones más favorables”. Sólo 36 aminoácidos están en regiones no permitidas, lo que en promedio se traduce que menos de 1 aminoácido están en regiones no permitidas por cadena polipeptídica, dado que el sistema está compuesto por 48 de estas cadenas. En la tabla 11 se entrega el análisis realizado a las interfases presentes en el modelo de varilla una vez finalizada la dinámica molecular y una comparación con las observaciones realizadas antes de efectuarla.

interfase	Nº aa en cada proteína	Nº aa hidrófobos	Total aa	% aa hidrófobos	ptes salinos	ptes H	Área (Å²)
PC – PC	139 y 148	48 y 46	287(367)	32.75	7 (3)	36 (31)	2655
PC – PE	124 y 128	35 y 48	252(261)	32.94	9 (2)	30 (16)	2447
PE – PE	128 y 126	45 y 49	254(294)	37.01	6 (1)	27 (8)	2281

Tabla 11. Análisis de la interfases del modelo final de interacción entre ficobiliproteínas para conformar la varilla mínima del ficobilisoma de *Gracilaria chilensis*. Valores entre paréntesis corresponden a los observado antes de efectuarse la dinámica molecular). PC: ficocianina; PE: ficoeritrina; aa: aminoácidos; ptes salinos: puentes salinos; ptes H: puentes hidrógeno

Se aprecia como a pesar de la disminución del número de aminoácidos integrados a la interfase, las interacciones se han optimizado dando una estructura más estable.

Para evaluar la posición de los cromóforos al terminar la dinámica molecular y obtener las coordenadas de los cromóforos desde el archivo de coordenadas del modelo completo al término de la dinámica molecular, se creó un script (archivo de comandos) para bash (interprete de comandos en Linux) que se muestra en la figura 21.

```
##script para obtener coordenadas
##de los cromoforos presentes
##usar asi:
##./cromoforos archivo.pdb
grep "CYC" $(1)> cromoforos_$(1)
grep "CYU" $(1)>> cromoforos_$(1)
grep "CYE" $(1)>> cromoforos_$(1)
echo "##### Gracias por usar Cromoforos #####"
echo "##### by Maxi #####"
```

Figura 21. “Script” en bash “Cromoforos”.

Este programa es ejecutable bajo ambiente UNIX o LINUX y es capaz de leer un archivo de coordenadas pdb y generar un nuevo archivo de coordenadas pdb sólo con las coordenadas de los cromóforos presentes en las ficobiliproteínas de *Gracilaria chilensis*. Utilizando el script antes mencionado, se extrajeron las coordenadas de los cromóforos tanto del modelo de varilla rígida, como del modelo final. (después de acabada la dinámica molecular). La figura 22 muestra la superposición de cromóforos en el estado inicial y final.

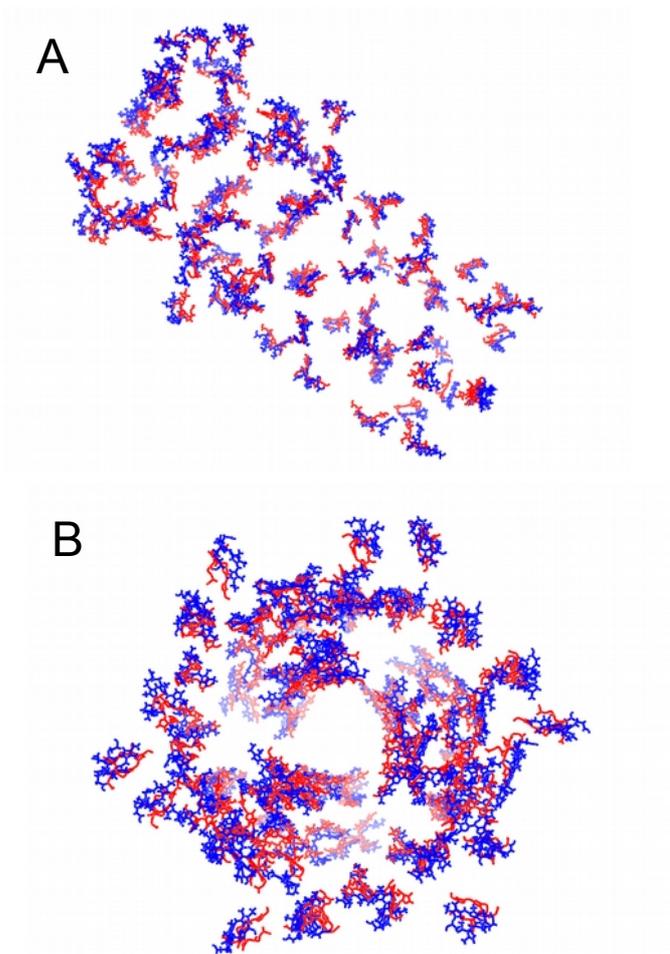


Figura 22. Superposición de los cromóforos antes (color rojo) y después de la dinámica molecular (color azul). Se observa la casi ausencia de desplazamiento de los cromóforos. A: vista lateral; B: vista superior

4.3 Vías preferenciales de conducción de luz

Los resultados de la evaluación funcional son presentados a continuación.

4.3.1 Vías preferenciales de luz intra ficoeritrina

Ficoeritrina de *Gracilaria chilensis* posee 30 cromóforos en total, de los cuales 24 corresponden a ficoeritrobilina y los 6 restantes a ficourobilina. El estudio se realizó utilizando la ecuación aproximada de la ley de Förster asumiendo un valor de tiempo de vida media de fluorescencia de 3.33 ps para cada uno de los 30 cromóforos, así como un radio de Förster de 61 Å. En base a la literatura, se asumieron como posibles pares

dador aceptor a aquellos cromóforos cuyos tiempos de transferencia sean menores a 50 ps. El inverso de la constante de transferencia denota el tiempo del evento (tabla 12).

Dador	Aceptor	Distancia (Å)	alfa	Kappa	1/Kt _{DA} (ps)	1/Kt _{DA} (ps)
cye_A139	cye_A82	26,53	49,95	0,91	27,09	27,09
cye_A82	cye_F82	19,36	150,01	-0,94	3,81	3,81
cye_L82	cye_A82	34,25	6,05	-1,50	46,18	46,18
cye_M82	cye_A82	27,25	144,73	1,80	8,12	8,12
cyu_F50	cye_A82	18,63	52,11	-0,40	16,62	16,62
cye_L158	cye_A139	24,26	38,04	-0,89	16,53	16,53
cye_C82	cye_B82	19,36	149,99	-0,94	3,81	3,81
cye_P82	cye_B82	32,93	156,04	1,55	33,78	33,78
cye_K139	cye_B158	23,89	139,73	0,50	47,25	47,25
cye_L158	cye_B158	27,18	82,74	1,29	15,59	15,59
cye_C139	cye_C82	26,53	49,94	0,91	27,06	27,06
cye_K82	cye_C82	27,25	144,69	1,80	8,12	8,12
cye_P82	cye_C82	34,25	6,06	-1,50	46,15	46,15
cyu_B50	cye_C82	18,63	52,05	-0,40	16,70	16,70
cye_P158	cye_C139	24,26	38,00	-0,89	16,55	16,55
cye_E82	cye_D82	19,36	150,01	-0,94	3,81	3,81
cye_N82	cye_D82	32,93	156,02	1,56	33,75	33,75
cye_O139	cye_D158	23,89	139,51	0,51	46,34	46,34
cye_P158	cye_D158	27,18	82,59	1,29	15,57	15,57
cye_E139	cye_E82	26,53	49,90	0,91	27,08	27,08
cye_N82	cye_E82	34,25	6,02	-1,50	46,20	46,20
cye_O82	cye_E82	27,25	144,74	1,80	8,12	8,12
cyu_D50	cye_E82	18,63	51,94	-0,40	16,94	16,94
cye_N158	cye_E139	24,26	37,96	-0,89	16,50	16,50
cye_L82	cye_F82	32,93	156,05	1,55	33,77	33,77
cye_M139	cye_F158	23,89	139,77	0,51	46,49	46,49
cye_N158	cye_F158	27,18	82,63	1,29	15,59	15,59
cye_P82	cye_K82	19,53	138,72	-1,10	2,96	2,96
cyu_P50	cye_K82	18,75	37,70	0,78	4,54	4,54
cye_M82	cye_L82	19,53	138,77	-1,10	2,96	2,96
cyu_L50	cye_M82	18,76	37,73	0,78	4,55	4,55
cye_O82	cye_N82	19,53	138,82	-1,10	2,96	2,96
cyu_N50	cye_O82	18,75	37,35	0,79	4,50	4,50

Tabla 12. Se presentan los valores para cada par dador aceptor que participan en las vías intra ficoeritrina. Se aprecian 3 vías idénticas de transferencia de luz repetidas simétricamente en ficoeritrina. En negro se destacan los pares dador aceptor que forman una de las vías antes señaladas. La nomenclatura usada es: cye = ficoeritroblina; cyu = ficourobilina, cuya letra seguida es el identificador de cadena a la cual pertenece y el número a la cisteína a la cual está unido covalentemente. “Distancia” es la distancia que separa al par de cromóforos; “alfa” es el ángulo formado por los vectores que describen el momento dipolar de cada cromóforo; “kappa” es el factor de orientación determinado semiempíricamente; “1/KT_{DA} y 1/KT_{AD}” son los inversos de las constantes de transferencia para el paso dador aceptor y aceptor dador respectivamente.

En la figura 23 se presenta un esquema en el cual se gráfica la vía preferencial de conducción de luz marcada en negrita en la tabla 12

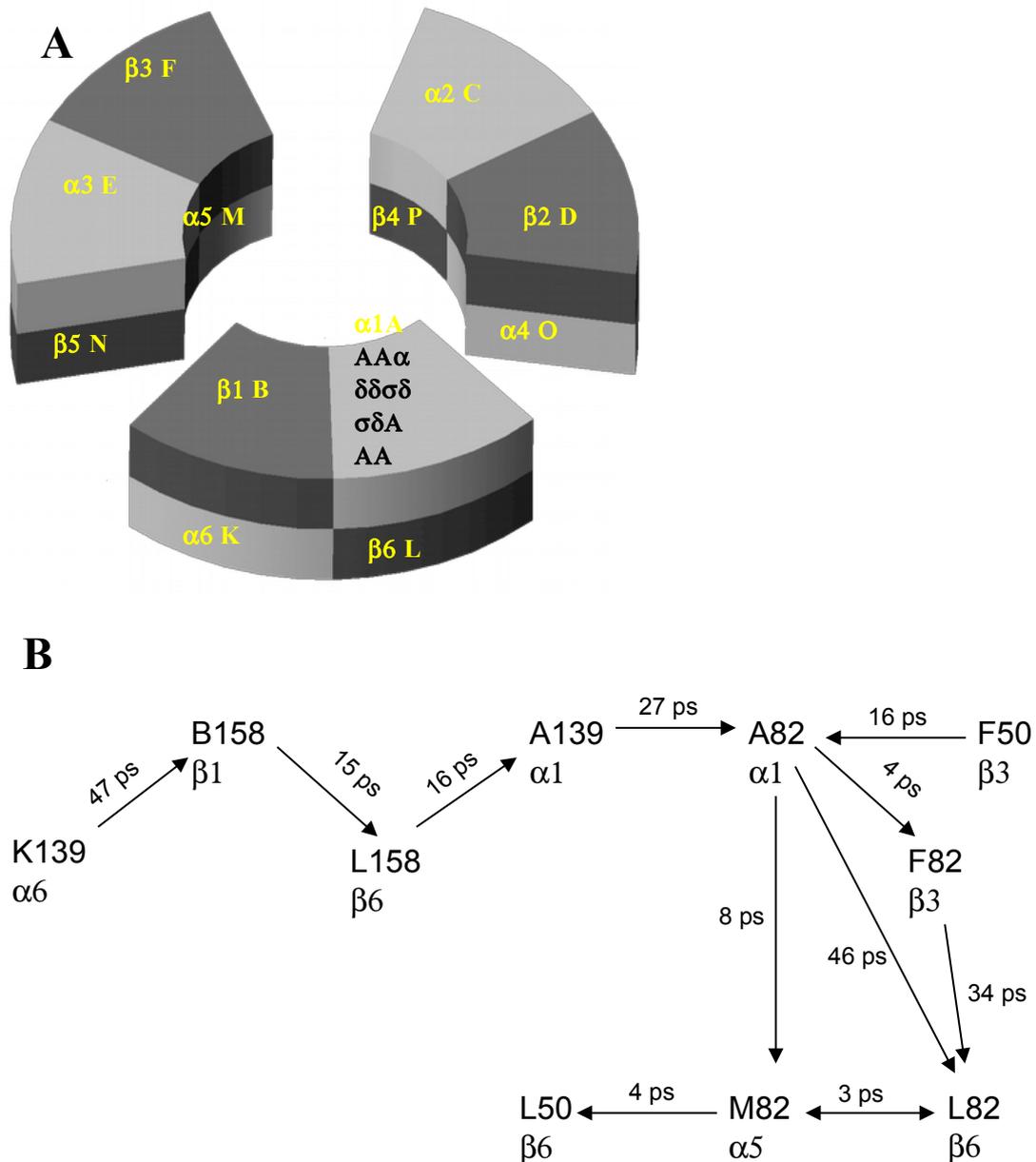


Figura 23. A: Esquema de una ficobiliproteína donde se denota esquemáticamente la ubicación de cada sub unidad con su nombre en nomenclatura tradicional (α y β) y con su respectivo nombre usado en el archivo de coordenadas PDB. **B:** Esquema de una de las tres vías preferenciales de conducción de luz intra ficoeritrina en el que señalan los tiempos de transferencia de luz entre los cromóforos. El número que acompaña al nombre de la subunidad, en el caso de la nomenclatura tradicional, hace alusión a la posición de la cisteína a la cual esta unido el cromoforo

4.3.2 Vías preferenciales de luz inter ficoeritrinas

4.3.2.a Vías inter ficoeritrinas en el modelo rígido

Usando la misma metodología que para las vías intra ficoeritrina se procedió a obtener las vías inter ficoeritrina. Dado que en el modelo de “docking” se genera un pequeño ángulo de inclinación entre las ficobiliproteínas, se presentan entonces 3 vías distintas de conducción, a diferencia de las vías intra que son 3 idénticas. En la tabla 13 se presentan los pares dador aceptor para las vías preferenciales de conducción de luz inter ficoeritrina

Dador	Aceptor	Distancia (Å)	alfa	Kappa	1/KT _{DA} (ps)	1/KT _{AD} (ps)
cye_K82	cye_a82	29,95	140,74	1,80	14,33	14,33
cye_K82	cye_f82	29,51	69,26	-1,15	32,04	32,04
cye_K139	cyu_f50	34,16	97,07	-1,21	69,99	69,99
cye_L82	cye_e82	29,51	5,52	-0,90	52,64	52,64
cye_M82	cye_e82	29,70	144,15	1,48	20,14	20,14
cye_N158	cyu_d50	38,67	36,65	-1,55	89,42	89,42
cye_O82	cye_c82	26,80	146,79	1,78	7,48	7,48
cye_O82	cyu_b50	30,80	118,87	-0,88	71,19	71,19
cye_P82	cye_f82	23,38	151,33	0,64	25,88	25,88
cye_P158	cyu_b50	34,54	42,05	-1,35	59,56	59,56
cyu_P50	cye_a82	28,09	120,53	0,69	67,08	67,08
cye_L82	cye_d82	23,11	155,56	1,36	5,26	5,26

Tabla 13. Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina en el modelo rígido de varilla. La nomenclatura usada es: cye = ficoeritrobilina; cyu = ficourobilina, cuya letra seguida es el identificador de cadena a la cual pertenece y el número a la cisteína a la cual está unido covalentemente. “Distancia” es la distancia que separa al par de cromóforos; “alfa” es el ángulo formado por los vectores que describen el momento dipolar de cada cromóforo; “kappa” es el factor de orientación determinado semiempíricamente; “1/KT_{DA} y 1/KT_{AD}” son los inversos de las constantes de transferencia para el paso dador aceptor y aceptor dador respectivamente.

Las figuras 24, 25 y 26 esquematizan las 3 vías preferenciales de conducción de luz inter ficoeritrina en el modelo rígido de “docking” a partir de cada una de las 3 vías intra ficoeritrina.

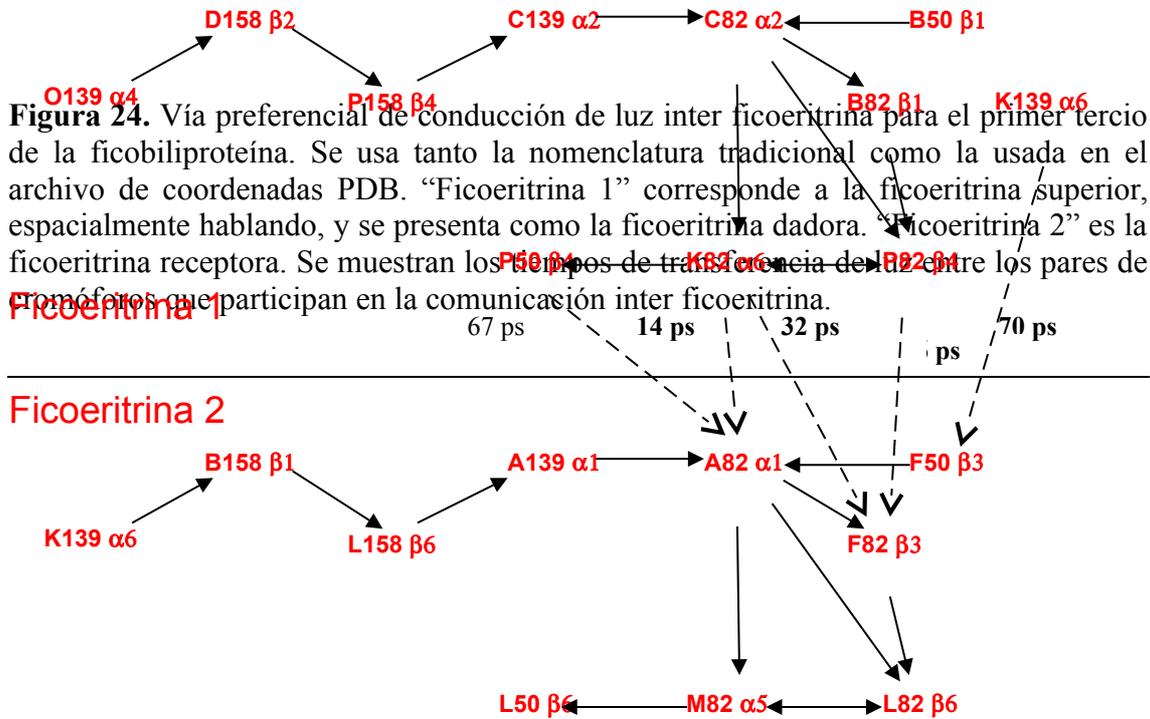


Figura 24. Vía preferencial de conducción de luz inter ficoeritrina para el primer tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina.

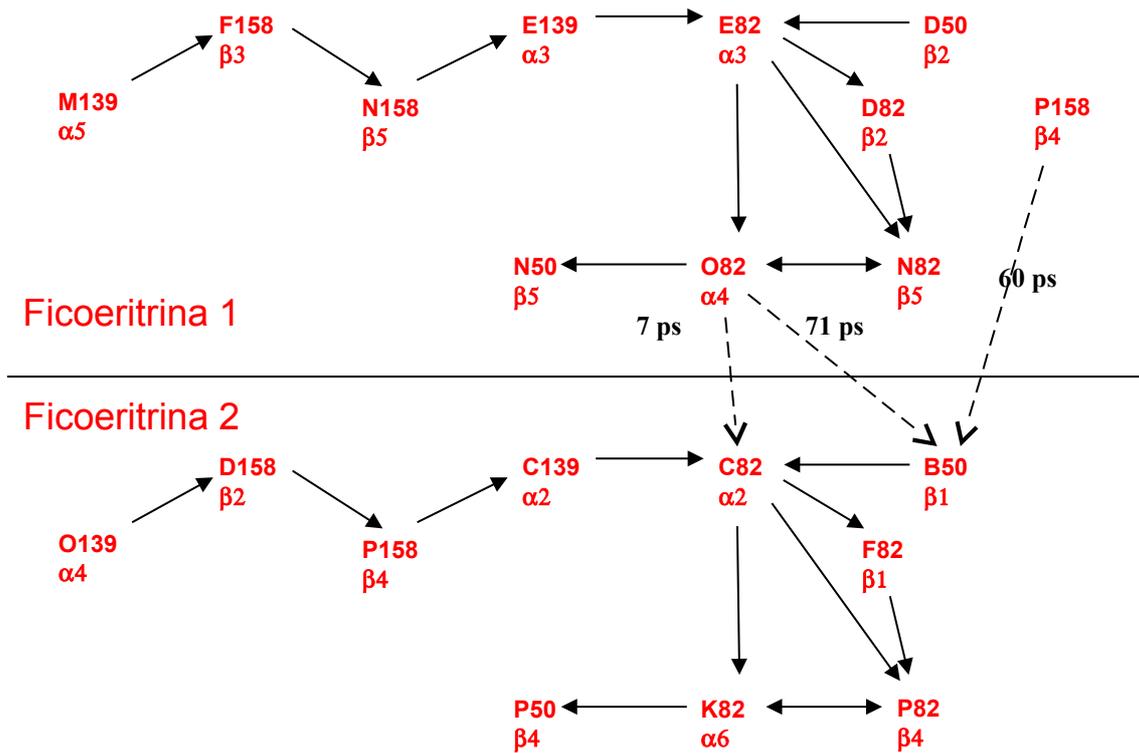


Figura 25. Vía preferencial de conducción de luz inter ficoeritrina para el segundo tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina.

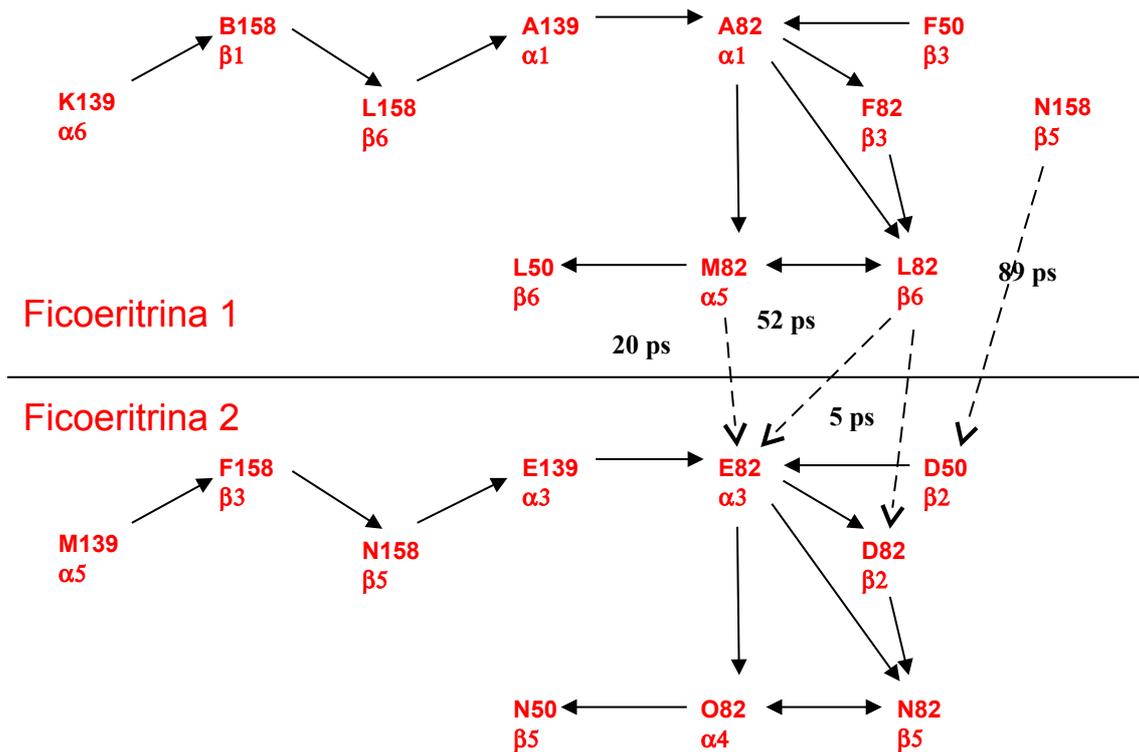


Figura 26. Vía preferencial de conducción de luz inter ficoeritrina para el tercer tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina.

4.3.2.b Vías inter ficoeritrinas en el modelo post dinámica molecular

El modelo de varilla estructuralmente válido, es el generado una vez acabada la dinámica molecular. A este modelo se le calcularon las vías preferenciales de conducción de luz inter ficoeritrina del mismo modo como se hizo con el modelo rígido,

comparando cada una de las 3 vías preferenciales de conducción de luz intra ficoeritrina con respecto a sus homólogas en la ficoeritrina receptora. En base a la literatura, se establece como par aceptor-dador a aquellos que presenten un tiempo de duración del evento menor a 100 ps (Gillbro y colaboradores, 1983; Glazer y colaboradores, 1985)

Dador	Aceptor	Distancia	alfa	Kappa	1/KT _{DA}	1/KT _{AD}
cye_K82	Cye_a82	31,36	21,48	-1,00	60,38	60,38
cye_K82	Cye_f82	31,15	165,86	1,50	26,03	26,03
cye_L82	Cye_d82	23,51	153,76	1,74	3,58	3,58
cye_L82	Cyu_d50	30,99	47,55	0,87	74,84	74,84
cye_M82	Cye_e82	33,08	118,99	1,32	48,46	48,47
cye_M82	Cyu_d50	31,57	90,91	1,35	34,97	34,97
cye_N82	Cye_b82	23,40	133,21	-0,77	17,80	17,80
cye_N82	Cye_d82	35,45	56,45	1,14	97,96	97,96
cye_O82	Cye_c82	28,20	90,17	1,31	18,92	18,92
cye_O139	Cye_c82	35,94	74,94	-1,30	81,73	81,73
cye_P82	Cye_f82	20,80	153,73	0,51	19,87	19,87
cyu_L50	Cye_e82	26,45	131,10	1,26	13,87	13,87
cyu_N50	Cye_d158	33,18	119,04	1,44	41,17	41,17
cyu_P50	Cye_b158	38,56	105,80	1,59	83,13	83,13

Tabla 14. Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina en el modelo una vez acabada la dinámica molecular. La nomenclatura usada es: cye = ficoeritrobilina; cyu = ficourobilina, cuya letra seguida es el identificador de cadena a la cual pertenece y el número a la cisteína a la cual está unido covalentemente. “Distancia” es la distancia que separa al par de cromóforos; “alfa” es el ángulo formado por los vectores que describen el momento dipolar de cada cromóforo; “kappa” es el factor de orientación determinado semiempíricamente; “1/KT AD y 1/KT DA” son los inversos de las constantes de transferencia para el paso dador aceptor y aceptor dador respectivamente.

En la tabla 14 se presentan los pares dador aceptor para las vías preferenciales de conducción de luz inter ficoeritrina en el modelo de varilla al término de la dinámica molecular. Las figuras 27, 28 y 29) esquematizan las 3 vías preferenciales de conducción de luz inter ficoeritrina en el modelo de varilla una vez acabada la dinámica molecular, a partir de cada una de las 3 vías intra ficoeritrina

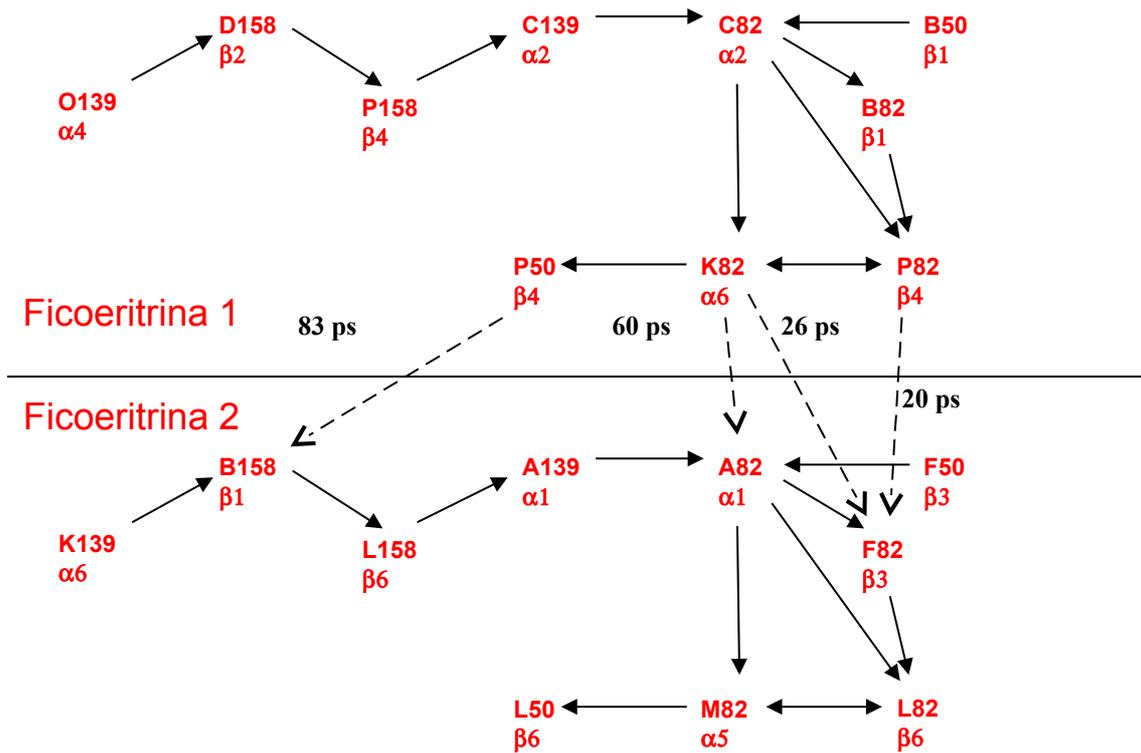


Figura 27. Vía preferencial de conducción de luz inter ficoeritrina para el primer tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina.

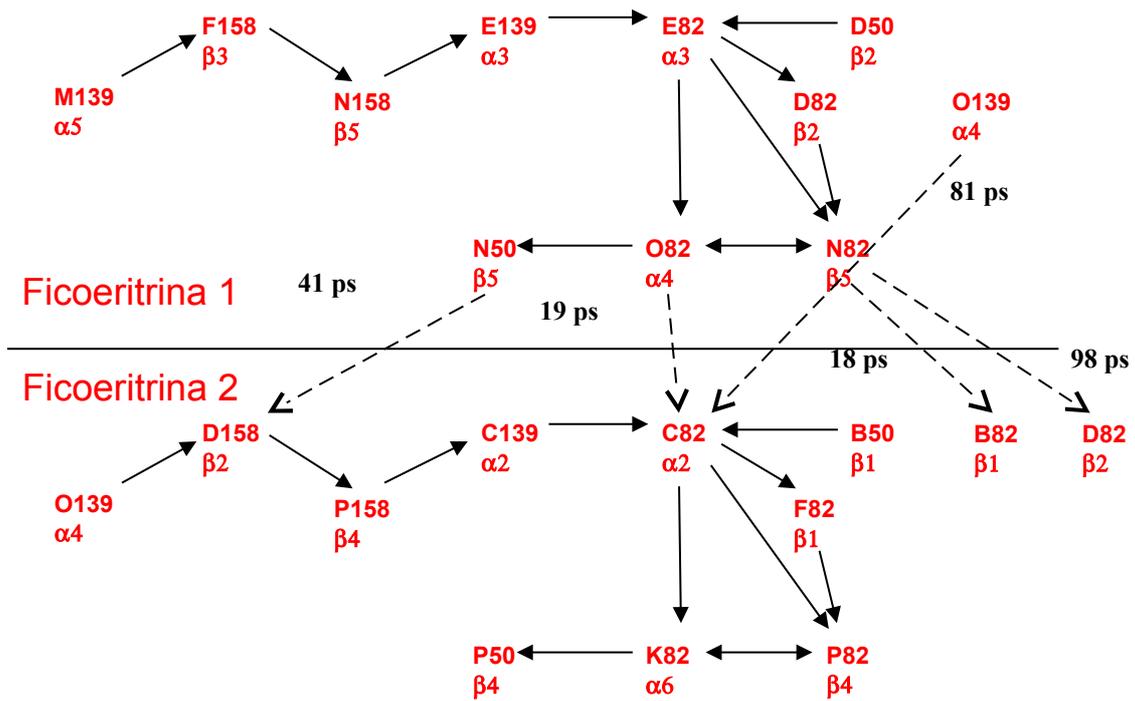


Figura 28. Vía preferencial de conducción de luz inter ficoeritrina para el segundo tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina.

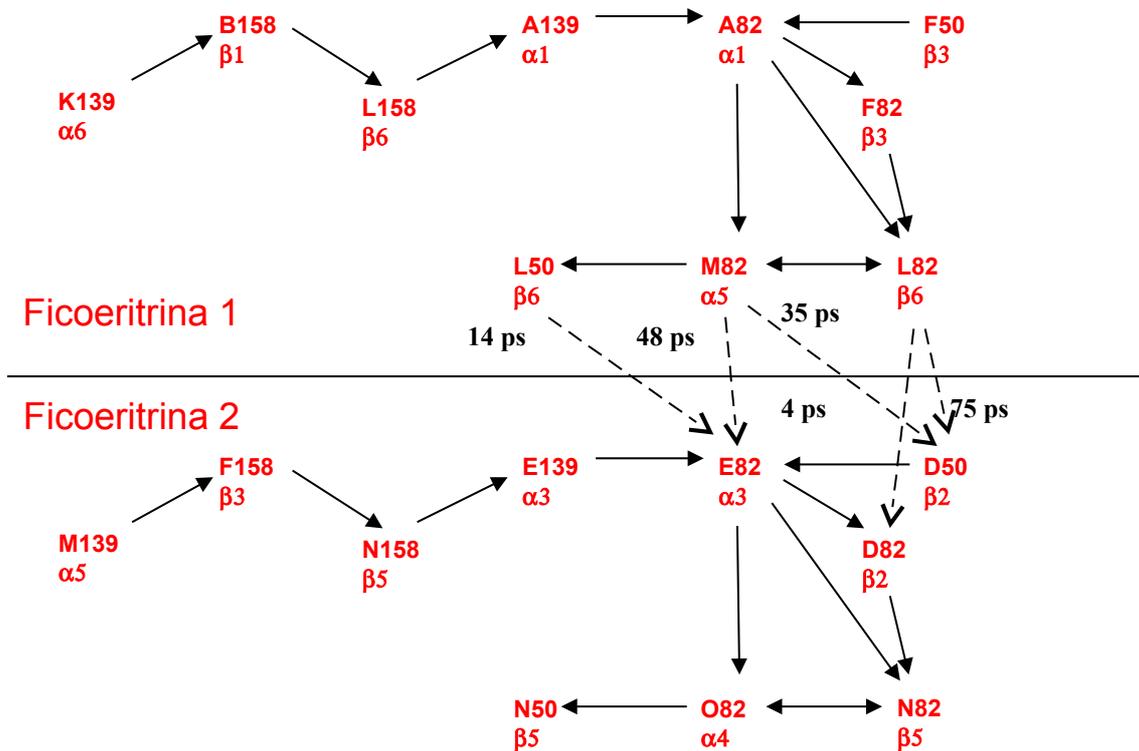


Figura 29. Vía preferencial de conducción de luz inter ficoeritrina para el tercer tercio de la ficobiliproteína. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 1” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 2” es la ficoeritrina receptora.

4.3.3 Vías preferenciales de luz inter ficoeritrina y ficocianina

4.3.3.a Vías inter ficoeritrina y ficocianina en el modelo rígido

El estudio se realizó utilizando la ecuación aproximada de la ley de Förster, tal como se describió en el capítulo de “metodología” de esta tesis. En base a la literatura,

se asumieron como posibles pares dador-aceptor a aquellos cuyo tiempo de duración del evento de transferencia de luz sea menor a 100 ps. La tabla 15 presenta los pares dador-aceptor que describen las vías inter ficoeritrina y ficocianina en el modelo rígido de varilla.

Dador	Aceptor	Distancia (Å)	Alfa	kappa	1/KT _{DA} (ps)	1/KT _{AD} (ps)
cye_K82	cyc_E84	30,33	130,33	1,13	39,12	58,63
cye_L82	cyc_B82	23,75	126,86	-0,41	67,04	100,48
cye_O82	cyc_C84	33,39	135,11	1,11	72,41	108,52
cye_P82	cyc_F82	24,44	130,30	-0,91	16,54	24,78
cye_N82	cyc_D82	22,49	124,16	-0,55	27,00	40,46
cyu_N50	cyc_D82	30,33	27,02	0,92	58,95	88,35
cyu_N50	cyc_E84	30,40	154,72	-0,91	61,65	92,40

Tabla 15. Se presentan los valores para cada par dador-aceptor que participan en las vías inter ficoeritrina y ficocianina. La nomenclatura usada es: cye = ficoeritrobilina; cyu = ficourobilina; cyc = ficocianobilina, cuya letra seguida es el identificador de cadena a la cual pertenece y el número a la cisteína a la cual está unido covalentemente. “Distancia” es la distancia que separa al par de cromóforos; “alfa” es el ángulo formado por los vectores que describen el momento dipolar de cada cromóforo; “kappa” es el factor de orientación determinado semiempíricamente; “1/KT_{DA} y 1/KT_{AD}” son los inversos de las constantes de transferencia para el paso dador-aceptor y aceptador-dador respectivamente.

De igual modo que en las vías inter ficoeritrina, se esquematizan las vías inter ficoeritrina-ficocianina partiendo como ficobiliproteína dadora a la ficoeritrina y como receptora a la ficocianina. Las figuras 30, 31 y 32 muestran las vías preferenciales de conducción de luz a partir de cada una de las 3 vías intra ficoeritrina.

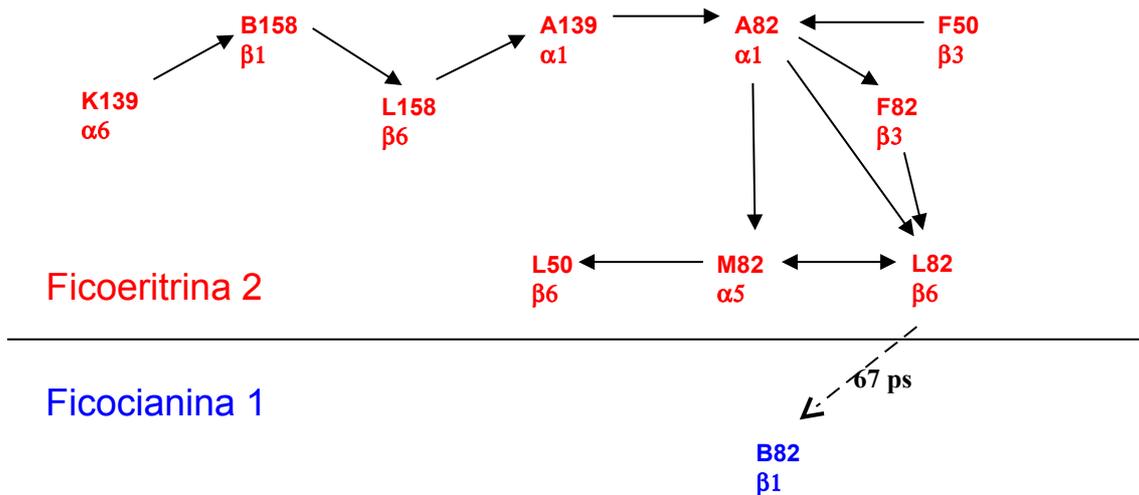


Figura 30. Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el primer tercio de ficoeritrina. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 2” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficocianina 1” es la ficocianina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina - ficocianina.

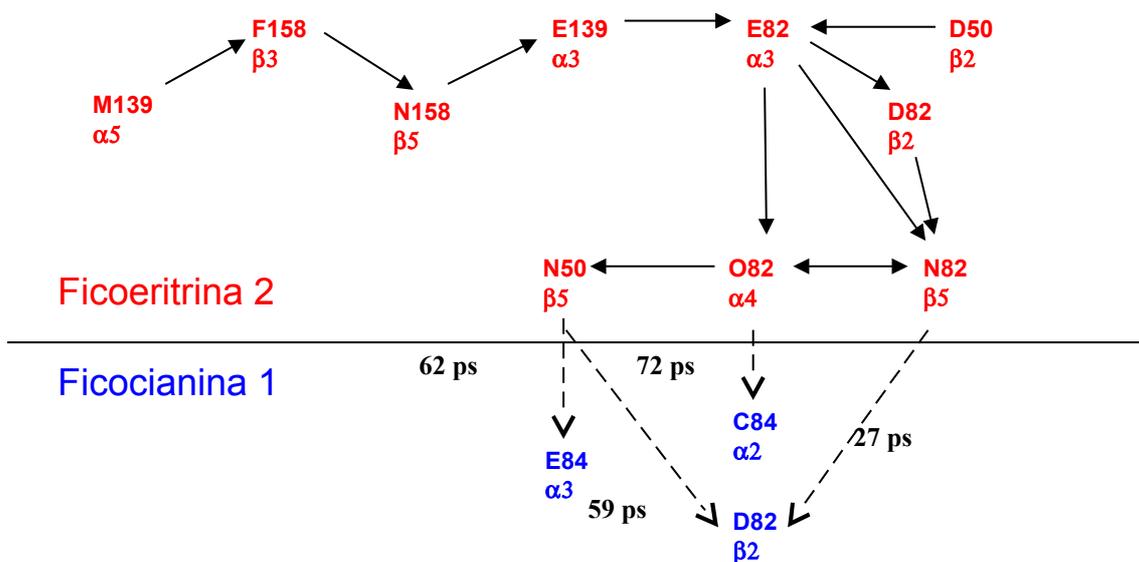


Figura 31. Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el segundo tercio de ficoeritrina. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 2” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficocianina 1” es la ficocianina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina - ficocianina.

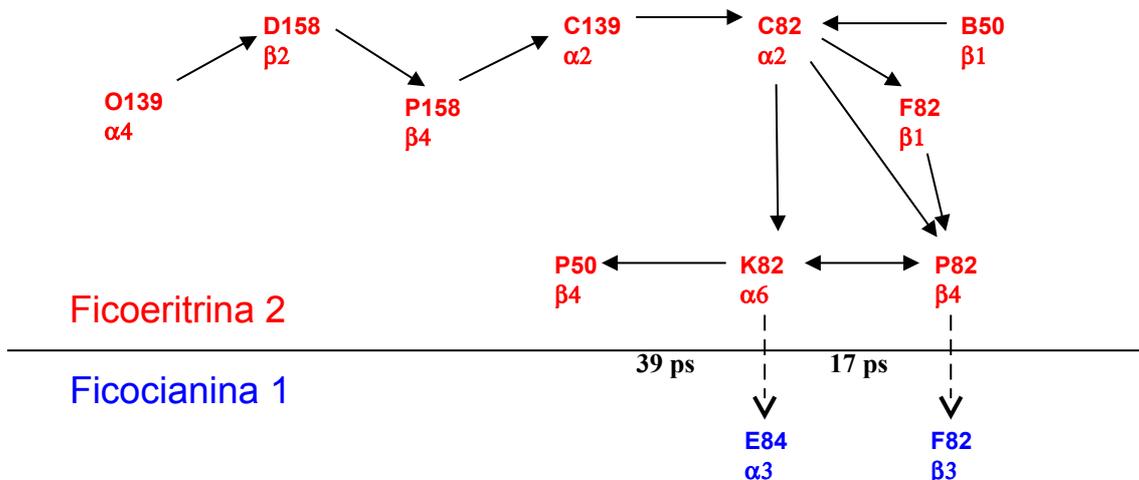


Figura 32. Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el tercer tercio de ficoeritrina. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 2” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficoeritrina 1” es la ficocianina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina - ficocianina.

4.3.3.b Vías inter ficoeritrina y ficocianina en el modelo post dinámica molecular.

Se determinaron las vías de conducción de luz en el modelo de varilla después de ejecutada la dinámica molecular. Se procedió de la misma manera efectuada en el caso de la varilla rígida, tal como lo señala el capítulo de metodología.

Dador	Aceptor	Distancia (Å)	alfa	kappa	1/KT _{DA} (ps)	1/KT _{AD} (ps)
cye_L82	cyc_B82	24,16	102,86	1,35	7,00	10,49
cye_L82	cyc_C84	34,56	79,02	-1,20	75,62	113,34
cye_N82	cyc_D82	24,22	71,20	1,33	7,27	10,89
cye_N82	cyc_E84	35,62	67,92	-1,59	51,53	77,24
cye_N158	cyc_F82	37,51	151,96	1,62	68,13	102,11
cyu_L50	cyc_B82	39,07	72,19	-1,61	87,54	131,21
cyu_N50	cyc_D82	34,10	75,42	-1,35	54,86	82,23
cyu_N50	cyc_E84	31,25	78,83	1,35	33,00	49,46

Tabla 16. Se presentan los valores para cada par dador aceptor que participan en las vías inter ficoeritrina y ficocianina. La nomenclatura usada es: cye = ficoeritrobilina; cyu = ficourobilina; cyc = ficocianobilina, cuya letra seguida es el identificador de cadena a la cual pertenece y el número a la cisteína a la cual está unido covalentemente. “Distancia” es la distancia que separa al par de cromóforos; “alfa” es el ángulo formado por los vectores que describen el momento dipolar de cada cromóforo; “kappa” es el factor de orientación determinado semiempíricamente; “1/KT_{DA} y 1/KT_{AD}” son los inversos de las constantes de transferencia para el paso dador aceptor y aceptor dador respectivamente.

En la tabla 16 se presentan los valores para los pares dador aceptor identificados en el modelo una vez acabada la dinámica molecular. Al igual que en el caso del modelo rígido, se eligieron como pares dador aceptor sólo a aquellos cuyo tiempo de duración del evento de transferencia de fotones era menor a 100 ps. El modelo post dinámica molecular, presenta a diferencia del modelo rígido, sólo 2 vías preferenciales de conducción de luz. Las figuras 33 y 34 muestran las vías preferenciales de conducción de luz a partir de las vías intra ficoeritrina.

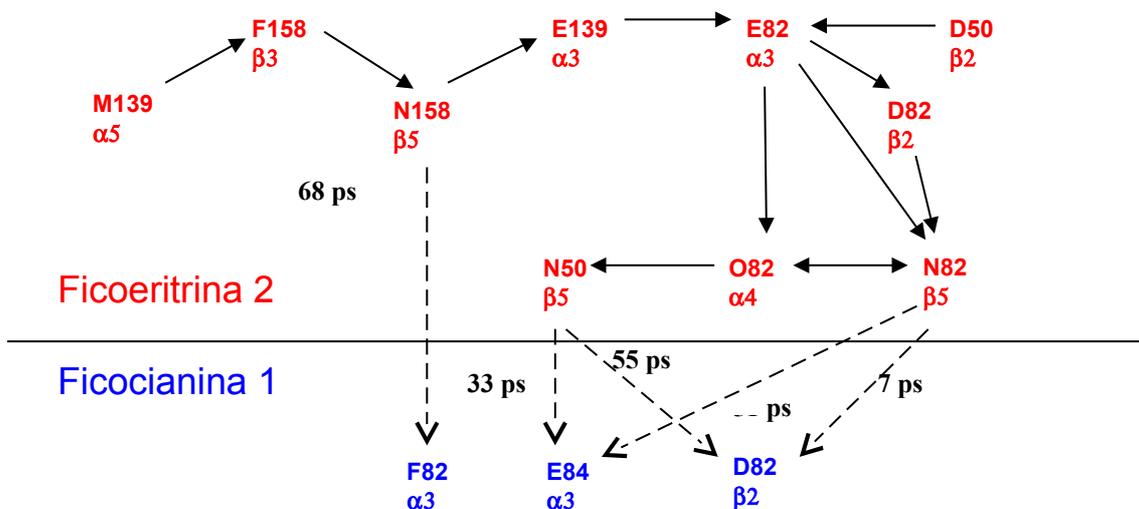


Figura 33. Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para un tercio de ficoeritrina. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 2” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficocianina 1” es la ficocianina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina - ficocianina.

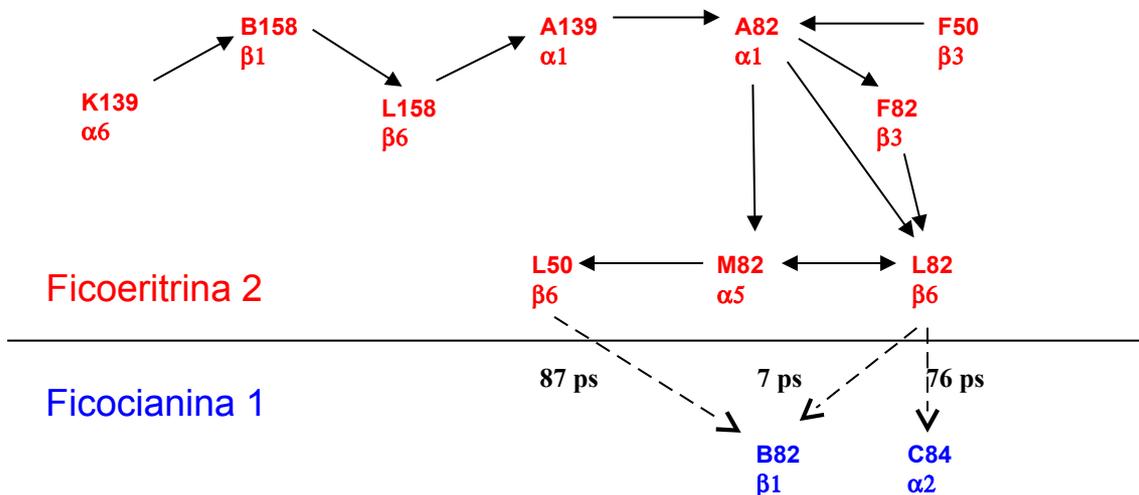


Figura 34. Vía preferencial de conducción de luz inter ficoeritrina y ficocianina para el segundo tercio de ficoeritrina. Se usa tanto la nomenclatura tradicional como la usada en el archivo de coordenadas PDB. “Ficoeritrina 2” corresponde a la ficoeritrina superior, espacialmente hablando, y se presenta como la ficoeritrina dadora. “Ficocianina 1” es la ficocianina receptora. Se muestran los tiempos de transferencia de luz entre los pares de cromóforos que participan en la comunicación inter ficoeritrina - ficocianina.

4.3.4 Vías preferenciales de luz intra ficocianina

Las vías preferenciales de conducción de luz tanto intra como inter ficocianina de *Gracilaria chilensis* han sido determinadas recientemente en nuestro laboratorio (tabla 17) (Almonacid, 2004) usando la expresión completa de la ecuación de Förster, dado a que existen en literatura los valores para cada uno de los parámetros usados en dicha expresión, determinados experimentalmente. Estos valores se utilizaron en conjunto con los determinados en este trabajo, para postular una vía preferencial completa de conducción al interior de una varilla formada por dos unidades de PE y dos unidades de PC. La figura 35 muestra un esquema que ilustra las 3 vías simétricas que describen la conducción de luz al interior de una ficocianina de *Gracilaria chilensis*.

Dador	Aceptor	Distancia	alfa	Kappa	1/KT _{DA} (ps)
$\alpha 84$	$\beta 82$	20,60	-1,25	4,22	6,50
$\beta 153$	$\beta 153$	26,10	1,51	5,64	5,64
$\alpha 84$	$\alpha 84$	25,60	1,27	15,00	15,00

Tabla 17. Vías de conducción de luz intra ficocianina. Las 3 vías simétricas son representadas en esta tabla resumen. El cromóforo es siempre ficocianoblina. “Distancia” es la distancia entre el par de cromóforos; “alfa” es el ángulo formado entre los vectores que describen el momento dipolar del par de cromóforos; “kappa” es el factor de orientación determinado semi empíricamente; “1/KT_{DA}” es la duración de evento de transferencia de fotones desde el dador al aceptor en pico segundos (ps). Adaptado desde Almonacid, 2004.

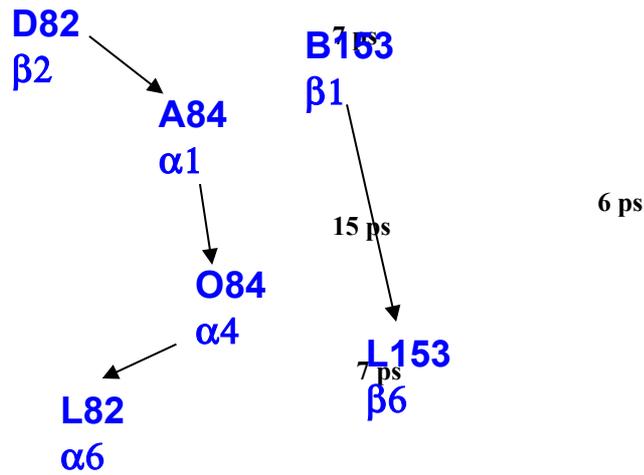


Figura 35. Esquema de una vía preferencial de conducción de luz intra ficocianina. Se observa que existen dos vías intra ficocianina por cada tercio, lo que nos indica la presencia de 6 vías intraficocianina, a diferencia de ficoeritrina que solo existen 3. Se señala además el tiempo de duración de la transferencia de luz entre los pares de cromóforos. Adaptado de Almonacid, 2004.

4.3.5 vías preferenciales de conducción de luz inter ficocianina

De igual modo que en el caso anterior, la información es extraída desde el estudio hecho por Almonacid en el 2004, y complementado por el estudio hecho por Bunster y colaboradores (trabajo en preparación), donde se usa un modelo rígido en el primer estudio, y luego uno con dinámica molecular, se postulan vías preferenciales de conducción de luz idénticas en ambos casos. En la tabla 18 se presentan los pares de cromóforos dador aceptor que describen las vías preferenciales de conducción de luz inter ficocianina, las cuales son 3 vías idénticas y simétricas.

Dador	Aceptor	Distancia	alfa	Kappa	1/KT _{DA} (ps)
β82	β 82	24,30	1,03	21,80	21,80
α84	β 82	34,50	-1,70	52,80	81,20
α84	α84	33,50	1,44	61,90	61,90
β153	β153	39,80	1,50	72,80	72,80
β82	α84	34,40	-1,69	83,20	54,10

Tabla 18. Vías de conducción de luz inter ficocianina. Las 3 vías simétricas son representadas en esta tabla resumen. El cromóforo es siempre ficocianoblina. “Distancia” es la distancia entre el par de cromóforos; “alfa” es el ángulo formado entre los vectores que describen el momento dipolar del par de cromóforos; “kappa” es el factor de orientación determinado semi empíricamente; “1/KT_{DA}” es la duración de evento de transferencia de fotones desde el dador al aceptor en pico segundos (ps). Adaptado desde Almonacid, 2004.

En la figura 36 se presenta un esquema de las vías preferenciales de conducción de luz inter ficocianina. Es interesante de apreciar cómo existen 2 vías independientes, una externa dada por los cromóforos β153, y una interna dada por el resto de ellos, sin nunca interferirse unas con otras.

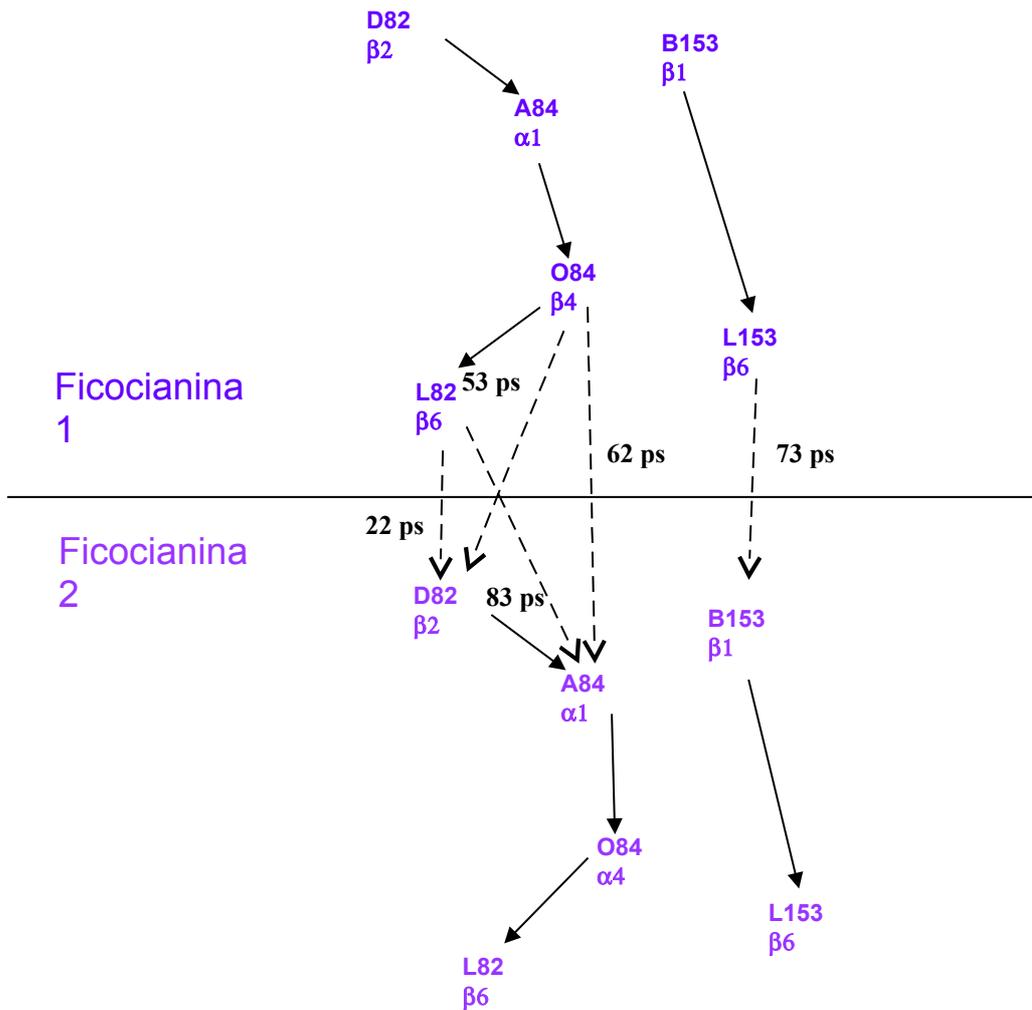


Figura 36. Esquema de las vías preferenciales de conducción de luz inter ficocianina. Se observa la presencia de 2 vías, una interna y una externa. “Ficocianina 1” corresponde a la ficobiliproteína dadora, y “Ficocianina 2” a la ficobiliproteína receptora. Se señala además el tiempo de duración de la transferencia de luz entre los pares de cromóforos. Adaptado desde Almonacid, 2004

5.- DISCUSIÓN

Elección del software de “docking” y generación del modelo rígido

La elección del software del “docking” utilizado en esta tesis, ZDOCK, resultó un importante punto en este trabajo, se realizó en base al evento CAPRI, que es una competencia que nos permite evaluar la eficiencia de los softwares para realizar “docking” de proteínas. Para último evento CAPRI, se utilizaron 7 complejos proteína-proteína de estructura conocida, pero cuyos datos no estaban publicados (generosamente donados por los autores mientras duraba el evento). Quince grupos de trabajo de los diecinueve participantes lograron predecir correctamente al menos 1 de los 7 complejos propuestos en el evento y ninguno de ellos pudo predecir el total de los complejos. Cinco grupos fueron capaces de predecir 3 complejos, seis grupos predijeron 2 complejos, y cuatro grupos predijeron sólo 1 complejo (Méndez y colaboradores. 2003).

Aunque existen softwares mejor evaluados, como ICM-DISCO (Fernández-Recio y colaboradores 2003), MolFit (Ben-Zeev y colaboradores 2003), SmoothDock (Camacho y colaboradores 2003), DOT (Law y colaboradores 2003). Ya que estos programas no estaban disponibles, el software de elección fue ZDOCK (Chen y colaboradores 2003), que logró predecir perfectamente 2 de los 7 complejos del evento CAPRI. Este programa además es atractivo y eficaz en sus funciones de evaluación, como son complementariedad geométrica, interacciones electrostáticas y desolvatación. La última función es muy importante, ya que superficies que interaccionan deben poder desolvatarse (Camacho y colaboradores 2003; Chen y colaboradores 2003), con el consiguiente gasto energético; predecir de antemano la factibilidad de desolvatar cierta parte de la superficie de la proteína nos permite acotar del área de búsqueda. El

programa ZDOCK evalúa la desolvatación con la metodología ACE (atomic contact energy), presentada en el capítulo de introducción, la que da muy buenos resultados. Sin embargo, el poder utilizar cálculos más exhaustivos en este punto, permitiría obtener resultados aun mejores. Este tipo de cálculos más exhaustivos consisten en buscar estructuración de solvente en la superficie de la proteína. El método consiste en hacer una simulación, en la cual nuestra proteína problema es puesta en una esfera de solvente y luego de un tiempo de simulación, que va a depender del sistema en particular, medimos la suma de momentos dipolares (vectores) de las aguas que rodean a la proteína. Al hacer esto, observaremos que si existe una estructuración de solvente en una porción de la proteína, el resultado de la suma vectorial de los momentos dipolares poseerá un valor considerable con un sentido definido, en tanto, en las porciones en las cuales no exista estructuración de solvente, la suma vectorial de los momentos dipolares será igual a cero, ya que al estar al azar se cancelarán unos con otros y por lo tanto, en estas últimas zonas es más factible que ocurra la desolvatación. Esta metodología, como ya se mencionó, posee un alto costo computacional, por lo cual no está integrada en ningún programa de “docking”. Si es un requerimiento su uso, es necesario acotar el área de búsqueda, restringiéndola sólo a aquella parte en la que se espera desolvatación. La metodología ACE es una aproximación estadística bastante buena, permitiendo que ZDOCK genere excelentes resultados (Chen y colaboradores 2003), que se han traducido en la compra de los derechos comerciales por parte de Accelrys Inc (<http://www.accelrys.com>). Otra ventaja de ZDOCK es que además está implementado en un servidor en la Universidad de Boston (<http://zdock.bu.edu>), lo que permite realizar los procesos de encaje o “docking”s, a una velocidad superior (36h de cálculo para un “docking” en un PC PIV de 2.0 GHz de velocidad del procesador y 1 GB en memoria ram, utilizando el 100% de ambos recursos, contra 40 minutos en el servidor).

El único inconveniente de este servidor es que sólo trabaja con un paso angular de 15°. Con las restricciones apropiadas resulta más que suficiente, si además se le suma una posterior dinámica molecular.

El uso de estructuras resueltas por difracción de rayos X nos entrega un dato experimental importante, pues se trabaja con la estructura real, y no con la de un modelo generado por homología, el cual podría presentar ciertos errores, y es por eso que no son recomendados para estudios de “docking”, tanto de proteínas como de ligandos.

La alta simetría presentada por la ficobiliproteínas es un dato importante a considerar, pues si bien conocemos que la conformación debe ser del tipo “back to back”, podemos tener distintas conformaciones si giramos axialmente entre 0 y 120°; dada la simetría de las ficobiliproteínas, cada 120° se repite la estructura. Por lo tanto, la información proveniente de los estudios de microscopía electrónica de ficobilisomas sobre el modo de interaccionar de las ficobiliproteínas en una varilla, es un dato experimental muy importante que permite acotar el área de búsqueda. Es necesario buscar el correcto ángulo de giro de una proteína sobre la otra. Estudios previos sobre complejos proteína-proteína (Vasker y colaboradores, 1994) señalan que debe existir una alta complementariedad entre las superficies de interacción de dos proteínas. Además, estudios realizados por Stec y colaboradores en 1999 demuestran la presencia de zonas cargadas positiva y negativamente en las ficobiliproteínas que coincidirían además con la disposición de mayor complementariedad. Estas evidencias apoyan el modelo formulado en esta tesis, ya que es consistente con los resultados del procedimiento de “docking”.

Los resultados de “docking” de las ficobiliproteínas indican que ZDOCK trabaja muy bien, prediciendo la orientación observada experimentalmente en los 3 casos, y además esta orientación está dentro de los 10 mejores resultados dado por el ranking del software. Los criterios de búsqueda seleccionan aquellos complejos que presentan el mayor empaquetamiento y el mayor número de aminoácidos hidrofóbicos presentes en la interfase. Esto puede estimarse ya que en complejos estables en el tiempo como ocurre con las ficobiliproteínas, se observa gran complementariedad geométrica e hidrofóbica. Esto explica la elección del complejo 308 en el “docking” ficoeritrina – ficoeritrina. El motivo de por qué ZDOCK predijo una solución buena en un ranking tan bajo puede radicar en los parámetros que usa ZDOCK para evaluar sus complejos, es decir, le otorga gran importancia a la desolvatación. Este parámetro pudo generar una superficie de interacción ideal errónea, que para el programa resulta lógica; en este caso el estudio de las superficies de interacción resulta muy importante y además presenta la necesidad de la información experimental como una herramienta adicional y necesaria para validar los resultados obtenidos mediante herramientas bioinformáticas.

Las proteínas en un “docking” son consideradas como cuerpos rígidos que interaccionan entre si, lo que desde un punto de vista computacional tiene grandes ventajas, pero desde el punto de vista real, está lejos de lo observado en la naturaleza. Esto hace necesario realizar una minimización energética del sistema, sobre todo debido a que los programas de “docking” tienden a forzar la interacción entre las proteínas, acercando más de lo debido algunos átomos (Krippahl y colaboradores, 2003). Esto hace que la energía del sistema sea muy alta, y nos entregue posibles falsos positivos con respecto a aminoácidos que se cree estén participando en la interfase del complejo proteína – proteína.

Creación de topologías de los cromóforos

Un importante logro en esta tesis fue la creación de las topologías de los cromóforos presentes en las ficobiliproteínas ficoeritrina y ficocianina. La disponibilidad de estas topologías posibilita en el futuro, cualquier tipo de estudio o simulación con la ficobiliproteína completa incluyendo los cromóforos.

La elaboración de las topologías se realizó en parte usando el servidor ProDrg (<http://davapc1.bioch.dundee.ac.uk/programs/prodrg/>; Schuettelkopf y colaboradores, 2004), y en su mayoría por análisis de la estructura cristalina de los cromóforos, debido a que ProDrg sólo entrega información para el campo de fuerza Gromos87, que no considera los hidrógenos de manera explícita, como lo hace OPLS. Esto obliga a complementar esta información con un estudio detallado de la estructura, así como de la hibridación presentada por cada átomo en la estructura del cromóforo.

Uno de los puntos más conflictivos es la asignación de cargas parciales, debido a que estas deben presentar aproximaciones tales que permitan incorporar la influencia del enlace peptídico así como la influencia de un medio proteico en las cercanías de las estructuras creadas, por lo que, posterior al cálculo *ab-initio* de estas cargas, se debe introducir ésta aproximación. Sin embargo, no existe consenso en la actualidad sobre cuál es la mejor aproximación post cálculo *ab-initio* de las cargas, por lo que se decidió asignar las cargas de acuerdo a lo indicado por ProDrg y por simple inspección de estructuras similares ya definidas en los campos de fuerza (por ejemplo, cadenas alifáticas, grupo Hemo, etc). Esta metodología es ampliamente usada por las personas

que crean librerías de topologías, como lo demuestran las publicaciones gratuitas de topologías de compuestos en la página de GROMACS (<http://www.gromacs.org>).

Elección del campo de fuerza utilizado

La inexistencia de información referente a campos de fuerza especializados para optimizar resultados de “docking” proteicos ha hecho necesario que se realice un estudio comparativo para así evaluar y elegir uno de estos que sea el más adecuado para este tipo de estudios. Como ya se ha visto, el software de mecánica molecular utilizado para realizar la dinámica molecular es GROMACS, el cual es muy versátil y rápido en comparación a otros. GROMACS posee 5 campos de fuerzas y estos fueron evaluados mediante un “docking dirigido” llamado así porque sus componentes están originalmente interaccionando y se conoce por ende la solución al problema. En este experimento se decidió trabajar con una solución con un grado de error, debido a que las superficies de interacción presentadas por los monómeros por separado ya son ideales (los aminoácidos poseen la conformación que presentan al estar interaccionando con otra proteína), lo que hace que se obtengan muy buenos resultados con el “docking” rígido, es decir, no sería un requisito realizar un posterior procedimiento de dinámica molecular. Al trabajar con un complejo que presenta una solución no óptima, lo que se busca es hacer que la dinámica reconstituya el complejo o al menos logre la orientación del mejor complejo (2.46 Å). En los experimentos de “docking no-dirigidos”, como es el caso de las ficobiliproteínas completas, los mejores resultados presentan ciertas imperfecciones, debido a que como los componentes han sido cristalizados por separado, sus superficies de contacto no presentan la conformación estérica óptima que presenta una superficie de un “docking” enlazado (Camacho y colaboradores 2003), por

lo que una etapa de dinámica molecular posterior permite corregir estas pequeñas imperfecciones.

Al observar los resultados dados por los 5 campos de fuerza, vemos que todos convergen a una estructura relativamente estable (figura 18) y además presentan un plateau de mínima energía (figura 19). Sin embargo, al hacer un análisis comparativo de las estructuras generadas luego de la dinámica molecular, vemos como OPLS/AA (ffoplsaa) produce los mejores resultados, pues fue el campo de fuerza, que si bien no reconstituyó el complejo cristalográfico, si logró optimizar la estructura, produciendo un desplazamiento de la subunidad β . Además, al comparar los Δ Energía, puede observarse que este campo de fuerza además presenta la mayor variación de energía potencial, lo que demuestra la generación de una estructura más estable que las generadas por los demás campos de fuerza. Cada uno de estos cálculos requiere 24h de cálculo computacional en un computador PC PIV de 2.0 GHz de velocidad del procesador y 1 GB de memoria ram, a excepción de Gromos96 para vacío (ffG43b1), que sólo requiere un par horas, debido al menor número de átomos considerado (no posee las aguas explícitas que tienen los otros campos de fuerza).

Una vez seleccionado el mejor campo de fuerza para realizar estudios de “docking”, es posible usarlo para hacer una serie de estudios de simulación molecular como los que se describen a continuación:

- **“minimizaciones energéticas”** para disminuir la energía del sistema creado por “docking” u optimizarlo estructuralmente, proceso en que los átomos no sienten ninguna fuerza (como si la simulación fuera realizada a 0 K), y que permite explorar

una pequeña zona de la hipersuperficie de potencial, pudiendo llegar a mínimos energéticos locales, que a veces pueden ser mayores que el mínimo global;

- "**simulaciones MonteCarlo**" (Abagyan y colaboradores, 1994; Steinbach, 2003), en que se crean y evalúan conformaciones al azar, dentro de un número de conformaciones posibles; esta metodología es útil para explorar varias posibilidades, permite encontrar un mínimo global, pero no permite generar una trayectoria del evento, por lo cual es difícil asegurar que se ha llegado a una conformación estable ;

- "**dinámicas moleculares**", técnica que permite encontrar el mínimo global y además genera una trayectoria que permite seguir la dinámica en el tiempo que dure la simulación, en el orden de los ps (pico segundos), pero que en algunos casos debe extenderse a los ns (nano segundos) debido a la alta energía que presente el sistema, o a que se obligue al sistema a llegar a una conformación deseada, para luego minimizar esta estructura (Steinbach, 2003). Una derivación de la dinámica molecular, es usar un protocolo conocido como "**simulated annealing**" (Filizola y colaboradores, 1997), el cual consiste en calentar el sistema y luego enfriarlo con el objeto de explorar en su totalidad la hipersuperficie de potencial, pues al aumentar la temperatura, aumentamos la energía al sistema, con lo cual es posible salir de mínimos locales muy pronunciados y llegar al mínimo global, cosa que se logra con el enfriamiento paulatino que permite ir explorando con más detalle la vecindad de los estados energéticos para una conformación dada.

Modelo de varilla post-dinámica molecular

La dinámica molecular aplicada a la varilla fue más corta en el tiempo, debido a que a los 120 ps el sistema converge a un punto donde no se presentan mayores variaciones, observándose un plateau que se mantiene constante hasta el final de la simulación (figura 20). Además, el uso de un protocolo de “simulated annealing” ayuda al sistema a encontrar el mínimo global, ya que al agregar temperatura, se entrega mayor energía al sistema lo que permite subir barreras energéticas de un mínimo local muy pronunciado, salir de éste y explorar toda la hipersuperficie de potencial. El uso de una temperatura constante para los cromóforos se debe a que en una primera simulación calentando todo el sistema, su conformación se alteró, ciclándose más de lo normal (figura 8), cosa que concuerda con la observación experimental que a temperaturas por sobre 30° C se pierden las propiedades espectroscópicas de las ficobiliproteínas (Bunster y colaboradores, 2000).

Los datos estructurales de la varilla obtenidos al finalizar la dinámica molecular muestran un aumento de su estabilidad al aumentar el número de puentes hidrógenos y puentes salinos, pero el número de aminoácidos interaccionando es menor, esto se explica porque el procedimiento de “docking” rígido, genera a veces distancias interatómicas cortas (Chen y colaboradores 2003), haciendo que interaccionen demasiado. Es justamente este tipo de interacciones las que se corrigen con la dinámica molecular posterior.

Se puede observar que la disposición espacial de los cromóforos en la antena rígida casi no varía con respecto a la posición observada en la varilla optimizada por la

dinámica molecular. Esto hace suponer entonces que sería posible realizar un modelo de conducción de luz realizando sólo “docking” rígido de las ficobiliproteínas, sin necesidad de una dinámica molecular. Sin embargo, si bien este modelo de conducción de luz presentaría las mismas vías preferenciales de conducción de luz que el modelo al que se le hizo una dinámica posterior, estructuralmente no sería válido. Al obviar la dinámica molecular generamos un modelo que posee superficies de interacción no optimizadas y contactos inter atómicos muy cortos de una proteína con respecto a la otra, por lo tanto se puede concluir que resulta necesario hacer la dinámica molecular para presentar un modelo, que además de explicar la alta eficiencia en la conducción de luz, sea válido desde el punto de vista estructural.

Vías preferenciales de conducción de luz intra ficobiliproteína

Las vías preferenciales de conducción de luz intra ficobiliproteínas fueron determinadas tanto para ficoeritrina como ficocianina usando la estructura cristalina de las mismas. Como ya se mencionó anteriormente, las vías de conducción intra ficocianina fueron determinadas en el año 2004 por Almonacid en el Laboratorio de Biofísica Molecular de la Universidad de Concepción, usando para ello la expresión completa de la ecuación de Förster (ecuación 1), debido a que se conoce la totalidad de los elementos que componen dicha ecuación. Además el factor de orientación κ fue calculado y no asumido, como se había mostrado en trabajos anteriores.

Un importante aporte de este trabajo es que, por primera vez se postulan vías preferenciales de conducción de luz al interior de ficoeritrina calculando constantes de transferencia de luz a través de la ley de Förster. Con anterioridad, sólo se habían usado aproximaciones basadas en distancias (Contreras-Martel, 2000), pero nunca se habían

calculado el valor de constantes de transferencia o tiempos de duración del evento de transferencia de fotones en un par dador aceptor.

Para poder hacer cálculos de constantes de transferencia, se utilizó la versión modificada de la ecuación que describe la ley de Förster (ecuación 3). Para usar esta ecuación, es necesario conocer el tiempo de vida media de fluorescencia del cromóforo dador, el factor de orientación κ , el radio de Förster del cromóforo dador, y la distancia entre el par de cromóforos. De la literatura se obtuvo el valor de tiempo de vida media de fluorescencia de ficoeritrobilina, 3.33 ns (Otto y colaboradores, 2003). Se usó este mismo valor para la totalidad de los cromóforos, inclusive ficourobilina, debido a que es similar al observado en el tiempo de vida media de fluorescencia de la ficoeritrina completa (Otto y colaboradores, 2003), lo que indica que ninguno de los cromóforos por separado tendrá un valor de vida media de fluorescencia mayor que este. Además, usando el valor de tiempo de vida media de fluorescencia de ficocianina en los primeros cálculos de constantes de transferencia, dió excelentes resultados en comparación a los observados luego con la expresión completa de la ecuación de Förster. El factor de orientación κ fue determinado en base a un método mecano cuántico semi empírico con el hamiltoniano PM3 (Frisch y colaboradores, 2002), para la determinación de los momentos dipolares, y luego calculado con un algoritmo matemático escrito en Maple7 (Waterloo Maple, Inc.). Esto es importante si se piensa que hasta la fecha sólo se habían asumido los momentos dipolares de los cromóforos, y más aun, nunca se habían realizado estudios a los cromóforos pertenecientes a ficoeritrina. La ecuación reducida, además integra el radio de Förster, parámetro que engloba todos los datos espectroscópicos que están explícitos en la ecuación completa de la ley de Förster. Dado que el fenómeno de FRET ocurre a distancias menores de 100 Å, y el radio de Förster se

define como la distancia a la cual la eficiencia de la transferencia es igual a la mitad del máximo (Stryer y colaboradores, 1967), se ha asumido este valor como en 50 Å, valor que ha sido utilizado en cálculos previos para describir, con esta expresión reducida de la ley de Förster, vías intra ficocianina (Duerring y colaboradores, 1991) antes de que se conocieran la totalidad de los parámetros espectroscópicos que permitieron usar la expresión completa de la ecuación.

La comparación de los resultados obtenidos con ambas metodologías fue similar, y estos son los principales argumentos para validar las vías postuladas con la ecuación reducida en ausencia de los parámetros espectroscópicos que permitan usar la expresión completa. No obstante, si analizamos la metodología utilizada y los resultados obtenidos, notaremos que el radio de Förster usado no es 50 Å, sino 61 Å. Se escogió este valor ya que con un radio de Förster de 50 Å no se logran conectar la totalidad de los cromóforos al interior de ficoeritrina, y existe evidencia experimental que demuestran que este valor puede oscilar entre 44 y 63 Å (Debreczeny y col. 1995). Se diseñó entonces una serie de cálculos en los cuales se definió como el radio de Förster correcto a aquel en el cual todos los cromóforos estén conectados al menos con uno más. El resultado de este pequeño experimento *in silico* reveló que con un radio de Förster 61 Å se logra la participación de la totalidad de los cromóforos, y no se observan cromóforos que no reciben ni entregan fotones, como ocurre con un valor de 50 Å, lo que desde el punto de vista funcional pierde sentido, debido a que si en la naturaleza se presentara esta situación, el sistema no presentaría la alta eficiencia observada.

Trabajos previos hechos en el Laboratorio de Biofísica Molecular (Contreras-Martel, 2000) mostraron que, basándose sólo en distancias, se presentaban 3 vías simétricas al interior de ficoeritrina, y que estas vías no corresponden a las esperadas que generen un tercio de ella, como sería los cromóforos que forman la estructura cristalina AB/KL (un tercio perfecto de la ficobiliproteína), sino que estos hacen participar a cromóforos de los tercios contiguos. Los cálculos realizados en esta tesis confirman esta aseveración, lo que nos indica entonces la importancia estructural de la forma de anillo de la proteína, y el por qué entonces no se presentan otros tipos de estructuras. Es importante señalar que el tiempo que duraría la transferencia de luz al interior de una ficoeritrina (35 ps aproximadamente), concuerda con los observados experimentalmente: 37 ps (Glazer y colaboradores, 1985), lo que confirma el modelo propuesto.

Vías preferenciales de conducción de luz inter ficobiliproteínas

Del mismo modo como fueron determinadas las vías intra ficobiliproteína fueron determinadas las vías inter ficobiliproteína. Para ficoeritrina, se observa que cada uno de los tres tercios que componen cada vía intra ficoeritrina posee conducción de luz hacia la ficoeritrina siguiente. No obstante, dichas vías hacia la siguiente ficoeritrina son distintas en cada tercio, esto es debido al ángulo generado entre las ficoeritrinas. Más aún, vemos que la dinámica molecular no es capaz de modificar este ángulo, lo que nos indica dos posibilidades: 1.- que la estructura es altamente estable con este grado de asimetría; o 2.- que el efecto de las ficocianinas hace que no sea posible la corrección de este ángulo. En este punto es bueno enfatizar la presencia de proteínas “linkers” (Glazer, 1989). Los "linkers" tienen por función ayudar a estabilizar al ficobilisoma y a sus componentes, y es por eso que tal vez estas proteínas podrían estar ayudando a

mantener una estructura más simétrica o compensar de alguna forma la asimetría. Sin embargo, se observa que los tiempos de transferencia están de acuerdo a los esperados para una ficobiliproteína (Glazer y colaboradores, 1985), y más aún, al analizar detalladamente los resultados podremos ver que se presentan 3 tipos de cromóforos: receptores, transmisores y dadores. La propiedad que un fotón puede entrar al sistema en cualquiera de estos, sin importar el tipo, demuestra la alta eficiencia del sistema, ahora representada por el modelo propuesto.

En el caso de la conducción ficoeritrina- ficocianina, a pesar de tener todos los datos para usar la ecuación completa de la ley de Förster para ficocianina, se debe trabajar con la ecuación reducida (ecuación 3). Esto se debe a que uno define las propiedades del cromóforo dador, y no las del receptor, y dado que el dador es un cromóforo de ficoeritrina, debemos usar la expresión reducida para así poder describirlo. Es interesante de señalar que, tanto en el modelo rígido como en aquel posdinámico, las vías son preferentemente unidireccionales desde ficoeritrina a ficocianina, a diferencia de lo que sucede entre ficoeritrina ficoeritrina, o ficocianina ficocianina, donde las vías son reversibles. Esto le da direccionalidad al sistema, y entrega la información que una vez que un fotón es entregado desde ficoeritrina a ficocianina, este no puede volver a atrás. No obstante, en este modelo es mucho más acentuado el ángulo formado entre las ficobiliproteínas, y es por ello que existe un menor número de vías, incluso al término de la dinámica molecular, sólo 2 de los 3 tercios de ficoeritrina aparecen participando en vías inter ficoeritrina ficocianina.

El modelo final se complementó con los datos obtenidos por Almonacid en el 2004, permitiendo disponer por primera vez de un modelo mínimo completo de conducción de luz a lo largo de una varilla del ficobilisoma de *Gracilaria chilensis*.

Estudios complementarios, usando la expresión completa de la ecuación de Förster, se realizarán siempre que se disponga de la totalidad de los datos espectroscópicos de cada cromóforo de ficoeritrina. Además como resultado de esta tesis es posible calcular los momentos dipolares usando métodos *ab initio* de mecánica cuántica, por lo tanto es posible formular en un futuro cercano un modelo de conducción de luz mucho más exhaustivo desde el punto de vista metodológico.

6.- CONCLUSIONES

1.- Se ha logrado estandarizar una técnica de “docking” proteico para ficobiliproteínas, estableciendo criterios cuantitativos de evaluación de los complejos generados.

2.- Se crearon las librerías de topologías de los cromóforos presentes en las ficobiliproteínas de *Gracilaria chilensis* para los diferentes campos de fuerza utilizados en los estudios de dinámica molecular.

3.- Se estableció un protocolo de optimización de complejos generados por procesos de “docking” mediante dinámica molecular.

4.- Se construyó un modelo mínimo de varilla de ficobilisoma de *Gracilaria chilensis* y se postuló vías de conducción de luz intra e inter ficobiliproteína. El modelo obtenido concuerda con las evidencias experimentales, por lo que constituye un modelo estructural y funcionalmente correcto.

7.- APÉNDICES

Apéndice 1: Instrucciones para el uso de ZDOCK

El programa ZDOCK así como sus programas anexos necesarios para realizar el “docking” de proteínas, no están integrados en el directorio /bin, por lo que deben ser ejecutados en la carpeta en la cual se encuentra el programa (por ejemplo: /home/maxi/zdock)

El primer paso es agregar los valores de ACE (atomic contact energy), parámetros para estimar la desolvatación en la superficie de cada proteína involucrada en el estudio de “docking” con el fin de acotar al área de búsqueda. Esto se realiza con el programa “mark_sur” del paquete ZDOCK, bajo el siguiente comando:

Comando 1:

```
maxi@localhost:./mark_sur proteina.pdb proteina_marcada.pdb
```

Este comando le dice al computador que lea el archivo denominado “proteína.pdb”, que le agregue los valores de ACE, de la superficie de Van der Waals y que reconozca los átomos que estén en la superficie. El resultado es un nuevo archivo, denominado “proteína_marcada.pdb”. Esto se debe hacer para ambas proteínas participantes del “docking”. Los valores de ACE permiten truncar los datos para acotar el área de búsqueda utilizando la información experimental disponible. Si sabemos qué aminoácidos no participan en la superficie de interacción, solo debemos cambiar su valor de ACE (columnas 55 y 56 del nuevo archivo PDB generado) por el valor 19. Esto hará que los resultados en los cuales estos aminoácidos aparezcan en la superficie de

interacción sean penalizados, con lo cual se le dice al programa que este resultado en particular no posee ningún significado biológico y debe ser descartado. Además, en la nueva versión que está disponible en el servidor de la Universidad de Boston, está la opción de que, además de decirle al programa qué aminoácidos no participan en la superficie de interacción, la de decirle cuales sí están en ella, garantizando que las observaciones hechas de manera experimental previamente sean incluidas en su totalidad a la hora de hacer la simulación de la interacción entre dos proteínas.

Una vez que ya tenemos nuestras proteínas listas, se debe ejecutar el programa ZDOCK propiamente tal, de la siguiente manera:

Comando 2:

```
maxi@localhost: ./zdock -N 500 -D -R proteina_marcada1.pdb  
-L proteina_marcada2.pdb -o resultados.out
```

Esta línea de comandos dice que ejecutemos ZDOCK con las opciones `-N` y `-D`. La opción `-N` indica cuantos resultados queremos en el archivo de salida, denominado en este caso “resultados.out”. En el ejemplo anterior, el valor de `-N` es 500, por lo cual sólo estarán los 500 primeros resultados, y no los 2000 resultados que entrega por defecto ZDOCK. Esto tiene importancia desde el punto de vista de espacio de disco duro, pues luego uno debe crear cada una de las estructuras, es decir, 2000 modelos por defecto, y si asumimos un valor de 1 MB por modelo, es necesario disponer de 2 GB de espacio libre en el disco duro para poder analizar los resultados. Luego, la opción `-D` es para hacer la búsqueda más exhaustiva, pues esto permite bajar el paso angular de 15°, que es el valor por defecto, a 6°. Este cambio, sin embargo, hace aumentar el tiempo de cálculo a 15 veces lo que demoraría el hacerlo a 15°. La orden `-R` y `-L` son para llamar a las proteínas que participarán del “docking”. `-R` será la receptora, es decir, la que esté

fija, e idealmente debe ser la de mayor tamaño para así disminuir el costo computacional que involucraría rotarla y trasladarla. –L entonces es la proteína ligando, la que será rotada y trasladada. Como ya se habrá deducido, la orden –o es para nombrar al archivo de salida que contendrá los resultados.

La duración del cálculo dependerá del tipo de equipamiento disponible; sin embargo el proceso puede monitorearse utilizando los comandos “top” o “ps” de Linux. Este software no permite detener el cálculo en estados intermedios. Al término del proceso se genera el archivo de salida “resultados.out”

Una vez finalizado el cálculo, es necesario generar las estructuras encontradas en el archivo “resultados.out”, este archivo incluye las coordenadas de los centros de masa de la proteína ligando y receptora en su encabezado, además del nombre de cada una de ellas (proteina_marcada1.pdb y proteina_marcada2.pdb). A partir de la línea 5 de este archivo se incluye una lista de los resultados desde el 1 al 2000 (1 al 500 en nuestro ejemplo) donde cada una de las líneas posee las coordenadas de la ubicación espacial de la proteína ligando, en base a su centro de masa. Con esta información, se llama a un nuevo programa del paquete ZDOCK para reconstruir las estructuras de los modelos encontrados en el archivo “resultados.out”. Esto lo hacemos con el comando “create.pl”, de la siguiente manera:

Comando 3:

```
maxi@localhost: ./create.pl resultados.out
```

El comando create.pl, es un script de Perl, por lo que debemos fijarnos que la ubicación de Perl sea la correcta en la primera línea del script. Para ello es necesario

escribir la ubicación correcta de Perl (la que podemos encontrar ya sea con el comando “whereis” o “locate”). Una vez ejecutado este “create.pl”, se crearán en el directorio actual los 500 complejos contenidos en el archivo resultados.out. Cada complejo es denominado complex.n, donde “n” es el número del ranking, o sea, el complejo N° 1 según ZDOCK será el archivo denominado complex.1, y el 500 será complex.500.

Apéndice 2: Instrucciones de uso para GROMACS

Gromacs a diferencia de ZDOCK está localizado dentro del directorio de ejecutables, por lo que puede ser llamado desde cualquier directorio creado para trabajar. Se recomienda que sea instalado en “rpm” (RedHat package manager) que es un formato precompilado que trae todo los requerimientos básicos para poder ejecutar la aplicación deseada.

1.- El primer paso es convertir las coordenadas del archivo pdb al formato gromos o gro. Es necesario además generar un archivo que contenga todas las topologías de este archivo de coordenadas. Esto lo hacemos con la siguiente orden:

Comando 4:

```
maxi@localhost: pdb2gmx -f proteina.pdb -p proteina.top -o
proteina.gro
```

En este ejemplo, el subprograma es “pdb2gmx” y las opciones `-f` indican el nombre del archivo de entrada; `-p` para señalar el nombre del archivo de salida de topologías, extensión `.top`; y `-o` es el nombre del archivo que contendrá las coordenadas en formato gromacs, con extensión `.gro`.

2.- En segundo lugar debemos generar una caja llena de solvente en la cual situaremos en el centro de ella a nuestra proteína:

Comando 5

```
maxi@localhost: editconf -f proteina -o -dc 0.5
maxi@localhost: genbox -cp out -cs -p proteina -o solvatada
```

La primera línea de comando llama al programa “editconf”, con las opciones `-f` para indicar el nombre del archivo de entrada en formato gro; la opción `-o` indica el nombre del archivo de salida, y en este caso se usará la opción por defecto que es `out.gro`; además se señala la opción `-dc`, que indica que la proteína debe ser centrada en la caja y que el solvente debe estar a una distancia no menor de 0.5 Å de los bordes de la caja. La segunda línea genera la caja propiamente tal, especificando el tipo (cúbica o u octaédrica) y el solvente utilizado. La opción `-cp` indica el nombre del archivo de entrada, que en este caso es `out` (el nombre que se dio como archivo de salida en la línea anterior); la opción `-cs` indica el tipo de solvente, que en este caso al no tener ningún valor utiliza el modelo de agua por defecto (`spc216`); la opción `-p` llama al archivo que contiene las topologías del sistema (archivo generado con el comando `pdb2gmx`); y por último la opción `-o` para indicar el nombre del archivo de salida, que contendrá las coordenadas de la proteína más el solvente seleccionado.

Para iniciar la simulación se necesita un archivo de comandos, denominado `archivo.mdp`, que puede generarse a partir de los que trae como ejemplo `gromacs`, lo que facilita mucho el trabajo. Estos archivos están en el directorio `/usr/local/share/gromacs/html`. El modo de dar estas instrucciones es de la siguiente manera:

Comando 6:

```
maxi@localhost: grompp -v -f em -c solvatada -o em -p
proteina
```

En esta línea de comando se llama al programa “grompp”, que se encarga de cargar todas las órdenes y deja todo listo para ejecutar la simulación. La opción `-v` despliega en pantalla de lo que va ocurriendo; la opción `-f` indica el nombre del

archivo de entrada de los parámetros de la simulación (en éste caso em.mdp); la opción `-c` entrega el archivo de coordenadas (solvatada.gro); la opción `-o` indica el nombre del archivo de salida (em.tpr) con las órdenes correspondientes para el comando que realizará la simulación (comando mdrun); por último, la opción `-p` indica el archivo que posee las topologías de nuestro sistema.

Este programa posee la capacidad de emitir señales de advertencia cuando detecta errores (warning). El error más frecuente corresponde a la carga neta del sistema que debe poseer carga cero (0), o lo más cercana a cero. Si existe una diferencia de carga, esta será informada con un warning, indicando la carga actual del sistema. Este problema se corrige usando el programa “genion” de acuerdo a:

Comado 7:

```
maxi@localhost: genion -s em -o em_ion -nname Cl -nn 15 -nq  
1 -g ion.log
```

En el ejemplo anterior, la opción `-s` indica el nombre del archivo de entrada, “em.tpr”, que generó como archivo de salida el programa “grompp”; la opción `-o` indica el nombre del nuevo archivo de coordenadas “gro”, que ahora posee además las coordenadas del contra ion agregado al sistema; la opción `-nname` es para indicar que será un ion negativo, en este caso Cl, también puede cambiarse por `-pname` si se trata de un ion positivo, seguido del nombre del ion; la opción `-nn` indica el número de iones que deben agregarse al sistema para neutralizar la carga existente (en este ejemplo agregan 15 cloruros); la opción `-nq` indica el valor absoluto de la carga del ion, que en este caso es 1; por último la opción `-g` genera un archivo log en el cual estará detalladamente todo lo realizado con el programa “genion”, inclusive la hora de uso. Al inicio de la ejecución de “genion” el programa interroga sobre los elementos del sistema

que deben sustituirse (agua por iones, generalmente). Una vez neutralizado el sistema se ejecuta nuevamente grompp, con el sistema más los contra iones. Esta etapa requiere editar el archivo de topologías, “proteina.top”, agregando en la última línea el nombre del ion (en este caso Cl) y al lado de él separado por al menos un espacio el número de iones, se debe restar este número al total de moléculas de solvente, denominadas SOL, para así mantener el número constante de elementos en la simulación. El programa grompp se vuelve a ejecutar nombrando como archivo de coordenadas en la opción -c, el nombre del archivo que contiene el sistema con los iones.

Ej: grompp -v -f em -c **em_ion** -o em -p proteina

La simulación se inicia con el programa “mdrun”, usando la sentencia:

Comando 8:

```
maxi@localhost: mdrun -v -s em -o em -c after_em -g emlog
```

El programa mdrun se ejecuta con la opción -v que detalla en pantalla lo que va sucediendo; la opción -s que indica el nombre del archivo de entrada, “em.tpr” en este caso; la opción -o que indica el nombre del archivo de salida que contendrá la trayectoria de la simulación, con la información correspondiente a las diferentes estructuras que se vayan formando a cada momento; la opción -c que indica el nombre del archivo de salida, correspondiente a un archivo de coordenadas formato “gro” de la última estructura generada por la simulación; por último, la opción -g que genera un archivo “log” que posee toda la información generada durante la simulación.

El protocolo que se utiliza con el software GROMACS es general a todo tipo de cálculo. La diferencia de cada simulación va a depender entonces del archivo de parámetros mdp (Se sugiere trabajar tomando como base el archivo mdp que viene

como ejemplo dentro de la carpeta “tutor. En este archivo se debe especificar el tipo de integrador a usar, es decir la indicación sobre el tipo de simulación a efectuar. Si se usa el integrador “steep” esta orden le indica al computador que la simulación debe ser una minimización energética utilizando “steepest descent”. Para realizar una dinámica molecular, se requiere cambiar el integrador a “md” y modificar algunos parámetros en el archivo mdp, especialmente los referente a temperatura y velocidades iniciales, parámetros que no se requieren para un steepest descent.

La ejecución de un steepest descent tarda solo unos segundos, y puede seguirse la simulación en pantalla. La ejecución de una dinámica molecular, requiere un tiempo de cálculo mucho mayor, dependiendo del procesador (días o semanas). Esto hace que al ejecutar el programa “mdrun” que comenzará la simulación, sea necesario el siguiente comando:

Comando 9:

```
maxi@localhost: mdrun -v -s pr -e pr -o pr -c after_pr -g  
prlog >& pr.job &
```

En este ejemplo, se observa que el nombre del archivo de parámetros es “pr.mdp”; todos los archivos de salida poseen este prefijo. Se recomienda llamar estos archivos de acuerdo a la simulación a realizar: en una minimización energética es conveniente llamar los archivos “em” (energy minimization), en el ejemplo recién citado el prefijo “pr” corresponde a una dinámica de pre equilibrio. En el “comando 9” para el programa “mdrun” se ha agregado la opción -e, que indica a “mdrun” que debe generar un archivo de salida que guarde toda la información referente a la energía del sistema a lo largo de la simulación. Este dato permite el análisis de la simulación hasta llegar a un mínimo energético. La segunda opción que se ha agregado al sistema está dada por el signo “>”. Este símbolo es usado en UNIX para indicar que lo se genere se

guarde en el archivo que se señale a continuación de este signo. En este ejemplo, el archivo es pr.job. Los símbolos “&” indican al computador que ejecute estas órdenes en segundo plano, guardando el proceso en el archivo pr.job. El comando “tail” permite saber lo que esta sucediendo en la simulación, y el tiempo de simulación restante. El siguiente comando ejecuta “tail”:

Comando 10:

```
maxi@localhost: tail -f pr.job
```

La opción -f indica que archivo debe leer. Como su traducción lo indica este comando lee la “cola” del archivo que se especifique en la opción -f.

A continuación se entrega el archivo mdp usado para la dinámica por 200 ps con simulated annealing de la varilla rígida del ficobilisoma de *Gracilaria chilensis*:

```

; User Maxi (236)
; septiembre 5 del 2004
; Input file
;
title           = MD antenna
cpp             = /lib/cpp
constraints     = all-bonds
integrator      = md
dt              = 0.002 ; ps !
nsteps         = 100000 ; total 200 ps.
comm_mode      = none
nstcomm        = 1
nstxout        = 500
nstvout        = 1000
nstfout        = 0
nstlog         = 500
nstenergy      = 500
nstlist        = 10
ns_type        = grid
rlist          = 1.0
rcoulomb       = 1.0
rvdw           = 1.0
optimize_fft   = yes
; Berendsen temperature coupling is on in two groups
Tcoupl         = berendsen
tc-grps        = Protein SOL CYC CYE CYU ASM NA+
tau_t          = 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t          = 288 288 283 283 283 288 288
; Energy monitoring
energygrps     = Protein SOL CYC CYE CYU ASM NA+
; Isotropic pressure coupling is now on
Pcoupl         = berendsen
Pcoupltype     = isotropic
tau_p          = 0.5
compressibility = 4.5e-5
ref_p          = 1.0
; Parametros para el simulated annealing
annealing      = single single single single single single
single
annealing_npoints = 4 4 2 2 2 4 4
annealing_time  = 0 10 150 190 0 10 150 190 0 200 0 200 0 200 0
10 150 190 0 10 150 190
annealing_temp  = 288 303 303 288 288 303 303 288 283 283 283 283
283 283 288 303 303 288 288 303 303 288
; Generate velocities is on at 288 K.
gen_vel        = yes
gen_temp       = 288.0
gen_seed       = 173529

```

Figura 37. Archivo “mdp” para realizar dinámica molecular con protocolo de simulated annealing usado en la varilla rígida de esta tesis.

Otras opciones para cada programa mencionado así como el modo de uso de los distintos programas de análisis pueden encontrarse en el manual de GROMACS

(<http://www.gromacs.org>; Van der Spoel, 2003)

Apéndice 3: Edición de archivos para crear topologías de compuestos nuevos

Una dato a considerar es que el servidor ProDrg (<http://davapc1.bioch.dundee.ac.uk/programs/prodrg/>) genera topologías para la mayoría de los compuestos en el campo de fuerza gromos (ffgm_x o gromacs).

Los distintos campos de fuerza que posee GROMACS traen por defecto las definiciones topológicas para un número limitado de compuestos químicos para las simulaciones correspondientes. Estos compuestos son los 20 aminoácidos, agua, algunas sustancias orgánicas como alcanos, alquenos y alquinos, y sustancias como ATP, NADPH; ligandos típicos que son comunes a muchas proteínas y enzimas, además de metales, iones y algunos gases.

A estas librerías de compuestos es posible agregar nuevos compuestos de interés para el investigador, como se describe a continuación.

En GROMACS, los archivos que definen las características de los compuestos están ubicados en el directorio denominado “top”, el cual se encuentra bajo la ruta /usr/local/share/gromacs. Para cada uno de los campos de fuerza, 7 archivos deben modificarse independientemente, más un octavo que es común a todos:

- aminoacid.dat
- ff??.atp
- ff??.rtp
- ff??.hdb
- ff??.bon.itp

- ff??nb.itp
- ff??-n.tdb
- ff??-c.tdb

En esta lista, los símbolos “??” indican que en esa posición se introducen nombres de los distintos campos de fuerza. “aminoacid.dat”, corresponde a una base de datos que contiene los nombres de los distintos compuestos definidos en todos los campos de fuerza, en esta ubicación se agrega el nuevo compuesto a esta base datos con el nombre que posee en el archivo pdb que analizaremos.

Ff??-n.tdb y ff??-c.tdb generan la información para indicar si nuestro compuesto es el N terminal o el C terminal de una cadena polipeptídica. Si no se ubica en alguna de estas posiciones su edición puede omitirse.

El archivo ff??.atp posee la información de los distintos tipos de átomos definidos para el campo de fuerza en particular. Un mismo átomo puede poseer distintas definiciones en este archivo; por ejemplo, el carbono es un átomo con muchas definiciones: ha sido definido participando como metano, CH₂, CH₃, alqueno, alquino, en un anillo de 6 miembros, de 5 miembros, etc. Lo mismo sucede con el oxígeno y los estados de oxidación de algunos metales. Aquí están definidas casi la totalidad de las circunstancias en la cual se podría encontrar un átomo, sin embargo, es posible además agregar tipos nuevos, correspondientes a un tipo nuevo de enlace.

Para agregar un nuevo compuesto solo es necesario editar 4 archivos:

- ff??.rtp

- ff??.hdb
- ff??.bon.itp
- ff??.nb.itp

ff??.rtp define cómo es el compuesto físicamente, es decir indica qué átomos están enlazados, los ángulos de enlaces, y define el tipo de átomos. El segundo archivo, ff??.hdb, indica a que átomos debemos agregar hidrógenos. Ff??.bon.itp indica las interacciones enlazantes y ff??.nb.itp las interacciones no enlazantes.

Ejemplo de edición para glicina en el campo de fuerza ffgmx (gromos87)

a) edición de archivo ffgmx.rtp:

Este archivo define los compuestos desde el punto de vista de enlaces, ángulos y tipos de átomos. Cada compuesto debe definirse uno a continuación del otro separado por una línea en blanco. Dentro de la definición de cada compuesto no deben quedar líneas en blanco. Este archivo posee la siguiente forma:

```
[ GLY ]
[ atoms ]
  N      N    -0.280    0
  H      H     0.280    0
  CA     CH2   0.000    1
  C      C     0.380    2
  O      O    -0.380    2
[ bonds ]
  N      H
  N      CA
  CA     C
  C      O
  -C     N
[ impropers ]
  N      -C    CA     H
  -C     -CA   N      -O
```

Figura 38. Formato de un compuesto en el archivo rtp.

Vemos que el aminoácido glicina es primeramente nombrado de la manera en la cual es llamado en el archivo pdb, GLY, dentro de paréntesis cuadrados. A continuación son indicados los átomos, luego los enlaces y por último los ángulos impropios que se dan en este aminoácido.

Dentro del campo átomos, en la primera columna se encuentran los átomos que componen el aminoácido, que deben corresponder a los listados en el archivo pdb a excepción de los hidrógenos que son agregados por GROMACS, pero si deben estar indicados en esta columna; y en la segunda columna se describe a que tipo de átomo corresponde. La obtención del tipo de átomo la sacamos del archivo ffgmx.atp, donde debemos buscar la definición que mejor se ajuste a nuestro átomo en particular. La tercera columna muestra la carga parcial que posee este átomo, y la cuarta columna al grupo par cargado, donde se ve que los dos primeros son un par, pues sus cargas suman cero, y así sucesivamente. La obtención de la carga parcial para un residuo nuevo puede ser uno de los problemas más grandes, pues en la actualidad existen diversos y contradictorios métodos de obtención de las cargas parciales, todos basados en cálculos ab-initio de mecánica cuántica (Frisch y colaboradores, 2002), llamados single point, debido a que no optimizan la estructura. Se deben usar bases de la familia Hartree Fock, en especial la HF/6311G** (Pigache y colaboradores, 2003), y usar aproximaciones para ser usadas en campos de fuerza, ya que se debe emular ciertas distribuciones de carga cuando interactúan con otros compuestos o átomos. Para esto pueden utilizarse metodologías como ESP/RESP y CHELPG (Frisch y colaboradores, 2002; Pigache y colaboradores, 2003). Resulta apropiado revisar los compuestos ya definidos extraer la información necesaria para el compuesto que queremos definir.

Enlaces. El segundo campo corresponde a los enlaces, donde se observan sólo 2 columnas. La primera indica el átomo que está enlazado con el de la columna dos. Se utiliza el nombre del átomo, no el tipo de átomo.

Ángulos impropios. El tercer campo hace referencia a los ángulos impropios y contiene en cada línea los cuatro átomos que forman un ángulo dihedro impropio. Esta información puede obtenerse por observación de la estructura de nuestro compuesto. El signo menos (-) indica los átomos que conforman parte del esqueleto de la proteína, a excepción del nitrógeno que no debe llevar este signo. Con esta información GROMACS deduce qué tríos de átomos están formando un ángulo, y cuáles son los dihedros propios. Para otros campos de fuerza, como OPLS/AA, en este archivo se debe detallar la información de ángulos, y dihedros propios en campos apartes señalado los átomos que los conforman.

b) edición del archivo ffgmx.hdb:

Este archivo posee la información para agregar los distintos hidrógenos a los residuos definidos en el archivo rtp. La forma de este archivo es así:

```
GLY  1
      1      1      N      -C      CA
```

Figura 39. Formato de un compuesto en el archivo hdb

En este archivo, la primera línea debe llevar el nombre del aminoácido o compuesto, seguido del número de adiciones. El número de adiciones (no es el número de hidrógenos agregados) es el número de átomos a los que se les deben agregar

hidrógeno. En las líneas sucesivas se indican estas adiciones. Cada adición se divide en 3 partes, dada por las 5 o 6 columnas que componen cada línea.

La primera columna, indica cuántos hidrógenos deben ser agregados a este átomo.

La segunda columna, indica el tipo de hidrógeno que corresponde a uno de los 7 tipos siguientes:

- 1 un hidrógeno planar, por ejemplo en anillos o enlace peptídico
- 2 un hidrógeno simple, por ejemplo el del un hidroxilo
- 3 dos hidrógenos planares, por ejemplo los visto en un NH_2
- 4 dos o tres hidrógenos tetraédricos, por ejemplo los del CH_3
- 5 un hidrógeno tetraédrico, como el del CCH
- 6 dos hidrógenos tetraédricos, como los del CH_2
- 7 dos hidrógenos de agua, HOH

Las 3 o 4 columnas restantes indican los nombres de los átomos que estén juntos al átomo que obtiene el hidrógeno, y se les denomina átomos control. El primer átomo que se nombra es aquel al que se le agrega el hidrógeno. Si son 3 o 4 átomos en total va a depender del tipo de adición, donde el tipo 1 lleva 3 átomos, y el resto 4, a excepción del 7.

Para glicina, sólo hay una adición, correspondiente al hidrógeno que debe ser agregado al nitrógeno del enlace peptídico (tipo 1), y que el átomo que lo recibe es el N, los átomos control son el C y el Ca.

c) edición del archivo ffgmxbon.itp:

Este archivo contiene la información separada por bloques en las cuales se definen las distintas interacciones. Así es posible escoger un caso cualquiera para ejemplificar el modo en el que deben ir las órdenes y cómo es posible obtener algunos valores.

```
[ bondtypes ]
; i   j   func      b0      kb
  C   O   1   0.12300  502080.
  C   OM  1   0.12500  418400.

[ angletypes ]
; i   j   k   func      th0      cth
  HO  OA   C   1   109.500  397.480
  HO  OA  CH1  1   109.500  397.480

[ dihedraltypes ]
; i   l   func      q0      cq
NR5* NR5   2   0.000  167.360
NR5* NR5*  2   0.000  167.360

[ dihedraltypes ]
; j   k   func      phi0      cp      mult
  C   OA   1   180.000  16.736  2
  C   N   1   180.000  33.472  2
```

Figura 40. Formato del archivo ffgmxbon.itp

Este archivo contiene los parámetros de todas las interacciones enlazantes, divididas en 4 grupos: tipos de enlaces, tipos de ángulos, y tipos de ángulos dihedros propios e impropios. Cada uno posee el tipo de átomo que participa (no el nombre del

átomo), el valor de la medida y una constante para este valor. Cada uno de los bloques que componen este archivo se describen a continuación.

El primer bloque hace referencia a los tipos de enlace, donde en la primera y segunda columna se nombra el tipo de átomo que participa del enlace, la tercera columna corresponde al tipo de función, que para enlaces es 1 en ffgmx y OPLS, y 2 para los basados en Gromos96. La cuarta columna lleva el valor de la longitud del enlace en nanómetros (nm). La quinta columna posee el valor de la constante de este enlace en $\text{kJ mol}^{-1}\text{nm}^{-2}$.

El segundo bloque hace referencia a los tipos de ángulos. Está compuesto por 6 columnas, de las cuales las 3 primeras indican los tipos de átomos que forman este ángulo; la cuarta columna indica la función de esta interacción, “1” para ffgmx y OPLS, y “2” para los basados en Gromos96. La quinta columna indica el valor en grados sexagesimales del ángulo theta (θ). Por último la sexta columna indica el valor de la constante de este ángulo en $\text{kJ mol}^{-1}\text{rad}^{-2}$ para ffgmx y OPLS, y kJ mol^{-1} para los basados en Gromos96.

El tercer bloque es para definir a los ángulos dihedros impropios. Posee 5 columnas, de las cuales las dos primeras hacen referencia a los tipos de átomos que están en las posiciones “i” y “l”. La tercera columna indica la función, que siempre es 2; la cuarta columna indica el valor del ángulo ξ , formado por los planos i,j,k y j,k,l. La quinta columna indica el valor de la constante de este ángulo, expresada en $\text{kJ mol}^{-1}\text{rad}^2$.

El cuarto bloque define los ángulos dihedros propios. Posee 6 columnas, de las cuales las 2 primeras indican los tipos de átomos en las posiciones “j” y “k”. La tercera columna indica la función, que siempre es 1. La cuarta columna define la medida del ángulo φ , formado por los planos i,j,k y j,k,l , la cual será 0 ó 180, dependiendo si los átomos i y l están al mismo lado vistos a través del eje formado por los átomos j y k (forma *cis*) o en lado opuestos (forma *trans*). La quinta columna indica el valor de la constante de este ángulo, expresada en kJ mol^{-1} . La sexta columna describe la multiplicidad, siendo ésta la distinta combinación de enlaces formados a partir de los átomos j y k .

d) edición del archivo ffgmxnb.itp:

Al igual que el archivo anterior, está dividido en bloques. Se describe un ejemplo general para indicar la composición de cada bloque.

```
[ atomtypes ]
;name      mass      charge  ptype      c6          c12
O          15.99940   0.000   A    0.22617E-02  0.74158E-06
OM         15.99940   0.000   A    0.22617E-02  0.74158E-06

[ nonbond_params ]
; i      j  func      c6          c12
O      O    1  0.22617E-02  0.74158E-06
O      OA   1  0.22617E-02  0.13807E-05

[ pairtypes ]
; i      j  func      cs6          cs12
O      O    1  0.22617E-02  0.74158E-06
O      OM   1  0.22617E-02  0.74158E-06
```

Figura 41. Formato del archivo ffgmxnb.itp

Este archivo está formado por 3 bloques; el primero de ellos hace referencia a los tipos de átomos. En este bloque deben ser definidos: el tipo de átomo; su masa, que la obtenemos del archivo ffgmx.atp; su carga, especificando si es neutro o no; el tipo de

partícula, en este caso “A” denotando que son átomos; y por últimos tenemos los parámetros de Van der Waals V (c6) y W (c12). Estos coeficientes se calculan de la siguiente manera:

Ecuación 4:

$$V_{ii} = C_i^{(6)} = 4 \epsilon_i \sigma_i^6 \text{ [kJ mol}^{-1} \text{ nm}^6 \text{]}$$

$$W_{ii} = C_i^{(12)} = 4 \epsilon_i \sigma_i^{12} \text{ [kJ mol}^{-1} \text{ nm}^{12} \text{]}$$

Donde ϵ es la constante dieléctrica y σ es el radio de Van der Waals.

El segundo bloque define parámetros no enlazantes entre dos átomos i j. Posee 5 columnas, donde las dos primeras indican los átomos que están participando de la interacción; la tercera columna indica la función, que siempre es 1; y la cuarta y quinta corresponde a los valores de los coeficientes de Van der Waals V y W, que se calculan de la siguiente manera:

Ecuación 5:

$$V_{ij} = C_{ij}^{(6)} = (C_i^{(6)} C_j^{(6)})^{1/2}$$

$$W_{ij} = C_{ij}^{(12)} = (C_i^{(12)} C_j^{(12)})^{1/2}$$

En que los valores de $C^{(6)}$ y $C^{(12)}$ se calculan de acuerdo a la regla descrita para el primer bloque.

El tercer bloque define las interacciones 1 – 4, es decir aquellas separadas por 3 enlaces. Describe lo mismo que lo observado en el segundo bloque, y el modo de calcular los parámetros V y W es igual.

Este procedimiento permite editar los archivos necesarios para crear un nuevo compuesto definido en uno de los campos de fuerza de GROMACS, en este caso el campo de fuerza ffgmx. Para la edición de las topologías para otros campos de fuerza se repite el procedimiento, con algunos cambios en el orden o en la ubicación de algunos bloques.

Apéndice 4: Cálculo de constantes de transferencia de luz de acuerdo a la Ley de Förster

Esta etapa es realizada íntegramente bajo sistema operativo Windows utilizando los programas Gaussian98, GaussView, y la herramienta matemática Maple7. Todos estos programas están instalados en las máquinas de cálculo del Laboratorio de Biofísica Molecular.

Obtención de los μ . El primer paso es la obtención de los momentos dipolares de los cromóforos que contienen las ficobiliproteínas. Para ello se deben extraer las coordenadas de cada cromóforo desde el archivo .pdb. Esto se hace abriendo el archivo .pdb que contiene las coordenadas atómicas de un hexámero de Ficobiliproteína en un procesador de texto. Se extraen las coordenadas de cada cromóforo a partir del archivo .pdb, y se crea un nuevo archivo .pdb para cada uno de ellos. Es importante notar que las coordenadas atómicas de los cromóforos pueden encontrarse en la posición respectiva que les corresponde dentro de la cadena polipeptídica, al final de la cadena polipeptídica en la cual se encuentran, o al final de todas las cadenas polipeptídicas. Otras veces no todos los átomos figuran juntos en el archivo .pdb, es decir, a veces una porción del cromóforo aparece en una parte del archivo .pdb y el resto en una posición distinta. Hay que tener estas precauciones en mente. Un cromóforo normal posee 43 átomos.

Luego se utiliza GaussView para abrir el archivo .pdb de cada cromóforo. Este programa permite agregar los átomos de hidrógeno necesarios (antes de abrir un archivo el programa pregunta automáticamente: “Would you like to add hydrogens to your structure?” a lo cual debemos responder que sí). Por defecto los archivos se

muestran en el estilo wireframe, un formato que dificulta visualizar átomos. Una visualización adecuada se obtiene cambiando la opción “Display Format” en menú “View” y dentro de ella “molecule”, donde se debemos cambiar la opción “wireframe” por “Balls & Stick”. Es común que aparezcan inconsistencias en la hibridización de los átomos pesados en la estructura de los cromóforos. Para agregar átomos de hidrógenos, se debe seleccionar la opción “Add Valence” y luego hacer “click” sobre el átomo al cual se quiere agregar un hidrógeno. Para borrar un hidrógeno basta con seleccionar la opción “Delete Atom” y luego hacer click sobre el átomo que se quiere borrar.

Estas mismas herramientas se usan para modelar la porción aromática de un cromóforo o protonar un Nitrógeno o Carboxilo. Aquí es también útil ir al menú View y seleccionar la opción Labels. El número de cada átomo se va a mostrar en el mismo orden en que aparecen en el archivo .pdb de entrada. Finalmente se debe guardar el archivo con extensión .gjf

Estos archivos gjf deben ser modificados para especificar el tipo de cálculo que uno quiere realizar en Gaussian98. Es necesario abrir el archivo .gjf de cada uno de los cromóforos en un editor de texto y borrar las líneas iniciales que por defecto dicen:

```
%mem=6MW
%nproc=1
%chk=153B.chk
# rhf geom=connectivity
```

Figura 42. Encabezado de un archivo .gjf

En vez de las líneas anteriores, se escriben 2 líneas que definen el tipo de cálculo que se desea y la memoria que se va a utilizar en el cálculo:

```
%mem=200MB
#T PM3 sp scf=(maxcycle=1000)
```

Figura 43. Encabezado correcto en un archivo .gjf para realizar cálculos SP (single point)

Además, se debe reemplazar la línea que viene a continuación (Title Card Required) por un título para el cálculo.

Finalmente se deben borrar todas las conectividades que posea el archivo, procurando dejar dos líneas en blanco al final del archivo.

Para ejecutar los cálculos en Gaussian98 se puede abrir Gaussian y dentro de él el archivo que se desee ejecutar. Luego se debe presionar la opción Run y especificar el nombre del archivo de salida que llevará la extensión .out.

Cuando se posee múltiples archivos .gjf a los cuales se desea conocer el momento dipolar, es conveniente ejecutarlos todos a partir de un solo archivo .bat desde DOS. Para ello, se debe crear un archivo .bat en la misma carpeta donde se poseen los archivos .gjf. Este archivo ejecutable debe poseer en la primera línea el archivo raíz donde se encuentra Gaussian98, y a continuación ejecutar Gaussian98 no interactivo (g98.exe) sobre cada una de las moléculas que se tienen, en este ejemplo: 153B, 153D y 153F.

```
path=c:\g98w
c:\g98w\g98.exe 153B.gjf
c:\g98w\g98.exe 153D.gjf
c:\g98w\g98.exe 153F.gjf
```

Figura 44. Ejemplo de un archivo .bat para ejecución seriada de archivos .gjf

Esto genera un archivo con extensión .out para cada cromóforo, el cual posee información referente al momento dipolar y a las coordenadas estándar asignadas por Gaussian98 a cada cromóforo.

El cálculo de kappa (κ) a partir de los momentos dipolares, así como el valor de las constantes de transferencias se realiza con el software Maple7. Este software asigna a cada cromóforo en el archivo PDB su correspondiente momento dipolar calculado en Gaussian98, calcula kappa, y luego, usando los valores de tiempo de vida media de fluorescencia y radio de Förster determina las constantes de transferencia. Para lograr todo esto, es necesario generar un archivo denominado arbitrariamente como nitrogenos-dipolo.txt, el cual contiene las coordenadas de los 4 cromóforos en el archivo .out más el momento dipolar. Los datos que se extraigan de cada archivo .out deben escribirse en un archivo de texto en el mismo orden en que aparecen. Es decir, para cada cromóforo tendremos una lista de 5 vectores: los 4 primeros correspondientes a la posición de los nitrógenos y el último para el momento dipolar. La lista que tendremos será de la siguiente manera:

```
-6.092470 -0.926771 0.105649
-1.932569 0.321522 -0.328373
1.009972 0.574141 0.206445
5.266833 -0.194070 -0.159424
1.000000 0.000000 0.000000
```

Figura 45. Líneas que describen a un cromóforo en el archivo nitrógenos-dipolo.txt

Luego, los datos de todos los cromóforos se reúnen en un solo archivo, el cual, por tanto, poseerá 480 vectores en el caso de la varilla generada en esta tesis (96 cromóforos). Es importante señalar que el orden en que se dispongan los cromóforos debe ser igual al orden escogido para definir las propiedades experimentales de ellos que se usan para los posteriores cálculos en Maple.

Posteriormente se extraen las coordenadas de los 4 nitrógenos de cada cromóforo, pero ahora desde el archivo PDB. Esta lista debe ser en el mismo orden que la lista anterior y debe ir inmediatamente a continuación (son 384 líneas más). Por lo tanto obtendremos un archivo de 884 líneas.

Este archivo generado es útil sólo si se usa ya sea la versión simplificada de la ecuación de Förster, o la versión completa, pero no podemos mezclar ambas. Por lo tanto, es recomendable generar archivos nitrógenos-dipolo.txt diferenciados para cada caso (por ejemplo uno para PE-PE, que usará la ecuación reducida; y uno para PC-PC que usará la ecuación completa, por poseer una información experimental completa).

Para la ecuación reducida, debe usarse el archivo Maple radio-mod.mws, disponible en el Laboratorio de Biofísica Molecular, dentro del cual se encuentra claramente detallado cada parámetro y el tipo de cálculo que realiza cada sección. Este programa permite modificar el radio de Förster y tiempo de vida media, así como el tipo de resultados a desplegar. Este archivo radio-mod.mws en su primera línea indica donde se encuentra el archivo nitrogenos-dipolo.txt, por lo tanto la ruta debe estar correctamente especificada. Los resultados se despliegan en el orden que son entregados a los cromóforos en el archivo nitrógenos-dipolo.txt, es decir, 1 corresponde al primer

cromóforo, y 10 al décimo, así con todos, por lo que es conveniente tener muy claro en que orden van los cromóforos para luego tener una correcta asignación de pares dadores-aceptores calculados en Maple7.

Para el uso de la ecuación de Förster completa, se requiere usar el archivo PC-PC.mws. Este archivo es similar al presentado recientemente, pero posee parámetros de tiempo de vida media de fluorescencia para cada cromóforo en particular; integrales de acoplamiento entre los espectros de emisión y absorción de los pares de cromóforos; y valores para el rendimiento cuántico de cada cromóforo. Los resultados son entregados de la misma forma; por lo tanto es necesario conocer correctamente el orden de los cromóforos en el archivo nitrógenos-dipolo.txt.

BIBLIOGRAFÍA

- Abagyan, R.; M. Trotoev; D. Kuznetsov. "ICM: a new method for structure modeling and design: applications to "docking" and structure prediction from the distorted native conformation". J. Comp. Chem., 1994, 15:488-506
- Almonacid, D "Vías preferenciales de transferencia de luz intra-ficocianinas e inter-ficocianinas en ficobilisomas", 2004 tesis para optar al título de bioquímico, Universidad de Concepción
- Almonacid, D.; M. Bunster y A. Matamala "Cálculo de la vía preferencial de conducción de luz en ficobilisomas" (en preparación)
- Anderson, L. K. and C. M. Toole "A model for early events in the assemble pathway of cyanobacterial phycobilisomes"
- Archakov, A; V. Govorum; A. Dubanov; Y. Ivanov; A. Veselovsky; P. Lewi; P. Janssen. "Protein-protein interactions as a target for drugs in proteomics", Proteomics, 2003, 3: 380-391
- Ben-Zeev E.; M. Eisenstein. "Weighted Geometric "docking": Incorporating External Information in the Rotation-Translation Scan" PROTEINS: Structure, Function, and Genetics, 2003, 52: 24-27
- Bunster, M.; C. Contreras-Martel; C. Bruna and J. Martínez-Oyanedel "R-ficoeritrina de *Gracilaria chilensis*: Estabilidad e interacción entre subunidades". Boletín de la Sociedad Chilena de Química, 2000, 45: 303-309
- Bunster, M.; J. Tellez and A. Candia "Characterization of phycobiliproteins present in *Gracilaria chilensis*". Boletín de la Sociedad Chilena de Química. 1997, 42: 449-455
- Camacho, C.; D. Gatchell. "Successful Discrimination of Protein Interactions" PROTEINS: Structure, Function, and Genetics, 2003, 52: 92-97
- Camacho, C; SR. Kimura; C. DeLisi; S. Vajda. "Kinetics of desolvation mediated protein-protein binding" Biophys. J., 2000, 78:1094-1105
- Camacho, C.; S. Vajda. "Protein-protein association kinetics and protein "docking"" Current opinion in Structural Biology, 2002, 12:36-40
- Chen, R.; L. Li; Z. Weng. "ZDOCK: An Initial-Stage Protein-"docking" Algorithm" PROTEINS: Structure, Function, and Genetics, 2003, 57: 80-87
- Chen, R.; Z. Weng "docking" unbound proteins using shape complementarity, desolvation and electrostatics", PROTEINS: Structure, Function, and Genetics, 2002, 47: 281-294

- Contreras-Martel, C. “Resolución de la estructura terciaria de R-Ficoeritrina de *Gracilaria chilensis*”, 2000, tesis para optar al grado de Doctor en Ciencias Biológicas área Biología Celular y Molecular, Universidad de Concepción.
- Contreras-Martel, C.; J. Martínez-Oyanedel, M. Bunster, P. Legrand, C. Piras, X. Vernede and J. C. Fontesilla-Camps “Crystallization and 2.2 A resolution of R-phycoerythrin from *Gracilaria chilensis*: a case of perfect hemihedral twinning”. *Acta crystallographica*, 2001, D57: 52-60
- Debreczeny M., K. Sauer, J. Zhou, D. Bryant “Comparison of Calculated and Experimentally Resolved Rate Constant for Excitation Energy Transfer in C-Phycocyanin. 1. Monomers” *J. Phys. Chem.* 1995, 99: 8412-8419
- Ducret, A; W. Sidler, G. Frank and H. Zuber “The complete amino acid sequence of R-phycoerythrin- α and β subunits from the red alga *Porphyridium cruentum*”. *Eur. J. Biochem.* 1994, 221: 563-580
- Ducret, A.; W. Sidler, E. Wehrli, G. Frank and H. Zuber “Isolation, characterization and electron microscopy analysis of a hemidiscoidal phycobilisome type from the cyanobacterium *Anabaena* sp. PCC 7120”. *Eur. J. Biochem.* 1996, 236: 1010-1024
- Elcock, A.; D. Sept, and J. A. McCammon “Computer Simulation of Protein-Protein Interactions”. *J. Phys. Chem. B*, 2001, 105: 1504-1518
- Fernández-Recio J.; M. Totrov; R. Abagyan. “ICM-DISCO “docking” by Global Energy Optimization With Fully Flexible Side-Chains” *PROTEINS: Structure, Function, and Genetics*, 2003, 52: 113-117
- Filizola M.; N. Centeno; J. Pérez. “Computational Study of the Conformational Domains of Peptide T”. *Journal of Peptide Science*, 1997, 3: 85-92
- Frisch, M. J.; G. W. Trucks; H. B. Schlegel; G. E. Scuseria; M. A. Robb; J. R. Cheeseman; V. G. Zakrewski; J. A. Montgomery; R. E. Stratmann; J. C. Burant; S. Dapprich; J. M. Millam; A. D. Daniels; K. N. Kudin; M. C. Strain; O. Farkas; J. Tomasi; V. Barone; M. Cossi; R. Cammi; B. Mennucci; C. Pomelli; C. Adamo; S. Clifford; J. Ochterski; G. A. Petersson; P. Y. Ayala; Q. Cui; K. Morokuma; D. K. Malick; A. D. Rabuck; K. Raghavachari; J. B. Foresman; J. Cioslowski; J. V. Ortiz; A. G. Baboul; B. B. Stefanov; G. Lui; A. Liashenko; P. Piskorz; I. Komaromi; R. Gomperts; R. L. Martin; D. J. Fox; T. Keith; M. A. Al-Laham; C. Y. Peng; A. Nanayakkara; M. Challacombe; P. M. W. Gill; B. Johnson; W. Chen; M. W. Wong; J. L. Andres; C. Gonzalez; M. Head-Gordon; E. S. Replogle; J. A. Pople. “Gaussian 98, Revision A.11.4” Gaussian Inc., Pittsburgh PA, 2002
- Gillbro T., A. Sandtrom, V. Sundtrom, A. Holzwarth “Polarized absorption picosecond kinetics as a probe of energy transfer in phycobilisomes of *Synechococcus* 6301” *FEBS* , 1983, 162(1): 64-68

- Glazer, A. “Directional Energy Transfer in a Photosynthetic Antenna”. *Journal Bio. Chem*, 1989, 264(1), 1-4
- Glazer, A, W. S. Yeh, S. Webb, J. Clark “Disk to Disk Transfer as the Rate-limiting step for energy flow in phycobilisomes” *Science*, 1985. 227: 419-423
- Guex, N.; M. Peitsch; T. Schwede; A. Diemand “Swiss-PdbViewer v3.7” <http://www.expasy.org/spdbv>, 2003
- Halperin, I.; B. Ma, H. Wolfson, and R. Nussinov “Principles of “docking”: An Overview of Search Algorithms and a Guide to Scoring Functions” *PROTEINS: Structure, Function, and Genetics*, 2002, 47:409–443
- InsightII. <http://www.accelrys.com/insight/>
- Janin, J.; K. Henrick, J. Moult, L. T. Eyck, M. J. E. Sternberg, S. Vajda, I. Vakser, and S. J. Wodak “CAPRI: A Critical Assessment of PRedicted Interactions” *PROTEINS: Structure, Function, and Genetics*, 2003, 52: 2–9
- Jiang, T.; J. Zhang, W. Chang and D. Liang “Crystal structure of R-Phycocyanin and possible energy transfer pathways in the phycobilisome”. *Biophysical Journal*, 2001, 81: 1171-1179
- Krippahl, L; J. Moura; P. Nuno-Palma “Modeling Protein Complexes with BIGGER”. *PROTEINS: Structure, Function, and Genetics*, 2003, 52: 19–23
- Laskowski, R. MacArthur, M. Smith, D. Jones, D. Hutchinson, G. Morris, L. Moss, D. Thornton, J. “PROCHECK v3.5.4” <http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html>
- Law D.; L. Eyck; O. Katzenelson; I. Tsigelny; V. Roberts; M. Pique; J. Mitchell. “Finding Needles in Haystacks: Reranking DOT Results by Using Shape Complementarity, Cluster Analysis, and Biological Information” *PROTEINS: Structure, Function, and Genetics*, 2003, 52: 33-40
- Linux Mandrake. <http://www.mandrakelinux.com>
- Lundell, D. Williams, R. Glazer, A. “Molecular architecture of a light-harvesting antenna. In vitro assembly of the rod substructures of *Synechococcus* 6301 phycobilisomes”. *Journal. Bio. Chem.* 1981, 256: 3580-3592
- Méndez R.; R. Leplae; L. De María; S. Wodak. “Assessment of Blind Predictions of Protein-Protein Interactions: Current Status of “docking” Methods” *PROTEINS: Structure, Function, and Genetics*, 2003, 52: 51–57
- Maiorov V. N. and G. M. Crippen “Size-Independent Comparison of Protein Three-Dimensional Structures” *PROTEINS: Structure, Function, and Genetics* 1995, 22: 273-283

- Otto H, T. Lamparter, B. Borucki, J. Hughes, M. P. Heyn. “Dimerization and Inter-Chromophores Distance of Cph1 Phytochrome from *Synechocystis*, as Monitored by Fluorescence Homo and Hetero Energy Transfer” *Biochemistry* 2003: A-K
- Pigache A., P. Cieplak, F. Dupradeu “RED version 1.0” <http://www.upicardie.fr/labo/lbpd/RED>
- PROCHECK v3.5.4 Server. <http://biotech.ebi.ac.uk:8400/>
- Protein-Protein Interaction Server V. 1.5, enero 15 2003. <http://www.biochem.ucl.ac.uk/bsm/PP/server/index.html>
- Samsonoff, W.; R. MacColl “Biliproteins and phycobilisomes from cyanobacteria and red algae at the extremes of habitat” *Arch. Microbiol*, 2001, 176: 400-405
- Schimer, T.; R. Huber; M. Schneider; W. Bode; M. Miller; M. L. Hackert “Crystal structure analysis and refinement at 2.5 Å of hexameric C-phycoyanin from the cyanobacterium *Agmenellun quadruplicatum*. The molecular model and its implications for light harvesting” *J. Mol. Biol.*, 1986, 188:651-676
- Schuettelkopf, A. W.; and D. M. F. van Aalten “PRODRG - a tool for high-throughput crystallography of protein-ligand complexes” *Acta Crystallographica*, 2004, D60: 1355-1363
- Stec B.; R. Troxler; M. Teeter “Crystal structure of C-phycoyanin from *Cyanidium caldarium* provides a new perspective on phycobilisome assembly” *Biophysical Journal*, 1999, 76:2912-2921
- Steinbach, P. “Introduction to Macromolecular Simulation” http://cmm.info.nih.gov/intro_simulation/course_for_html.html, 2003
- Stryer L, R. Haugland “Energy Transfer: A Spectroscopic Ruler” *P. N. A. S.* 1967, 58: 719-726
- van der Spoel, D.; E. Lindahl; A. Hess; R. van Buuren; E. Apol; P. J. Meulenhoff; D. P. Tieleman; A. L. T. M. Sijbers; K. A. Feenstra; R. van Drunen; and H. J. C. Berendsen “Gromacs User Manual version 3.2”, www.gromacs.org, 2004
- Vasker, I; C. Aflalo “Hydrophobic “docking”: a proposed enhancement to molecular recognition techniques” *PROTEINS: Structure, Function, and Genetics*, 1994, 20: 320-329
- Wang, X.; L. Li, W. Chang, J. Zhang, L. Gui, B. Guo, D. Liang “Structure of C-phycoyanin from *Spirulina platensis* at 2.2 Å resolution: a novel monoclinic crystal form for phycobiliproteins in phycobilisome”. *Acta Crystallography*, 2001, D57: 784-792

- Zhang, C.; G. Vasmatazis; JL Cornette; C. DeLisi. “Determination of atomic desolvation energies from the structures of crystallized proteins” J. Mol. Biol., 1997, 267:707-726
- ZDOCK server. <http://zdock.bu.edu>
- Zolla, L.; M. Bianchetti and S. Rinalduci “Functional studies of the Synechocystis phycobilisome organization by HPLC on line with a mass spectrometer” Eur. J. Biochem. 2002, 269: 1534-1542